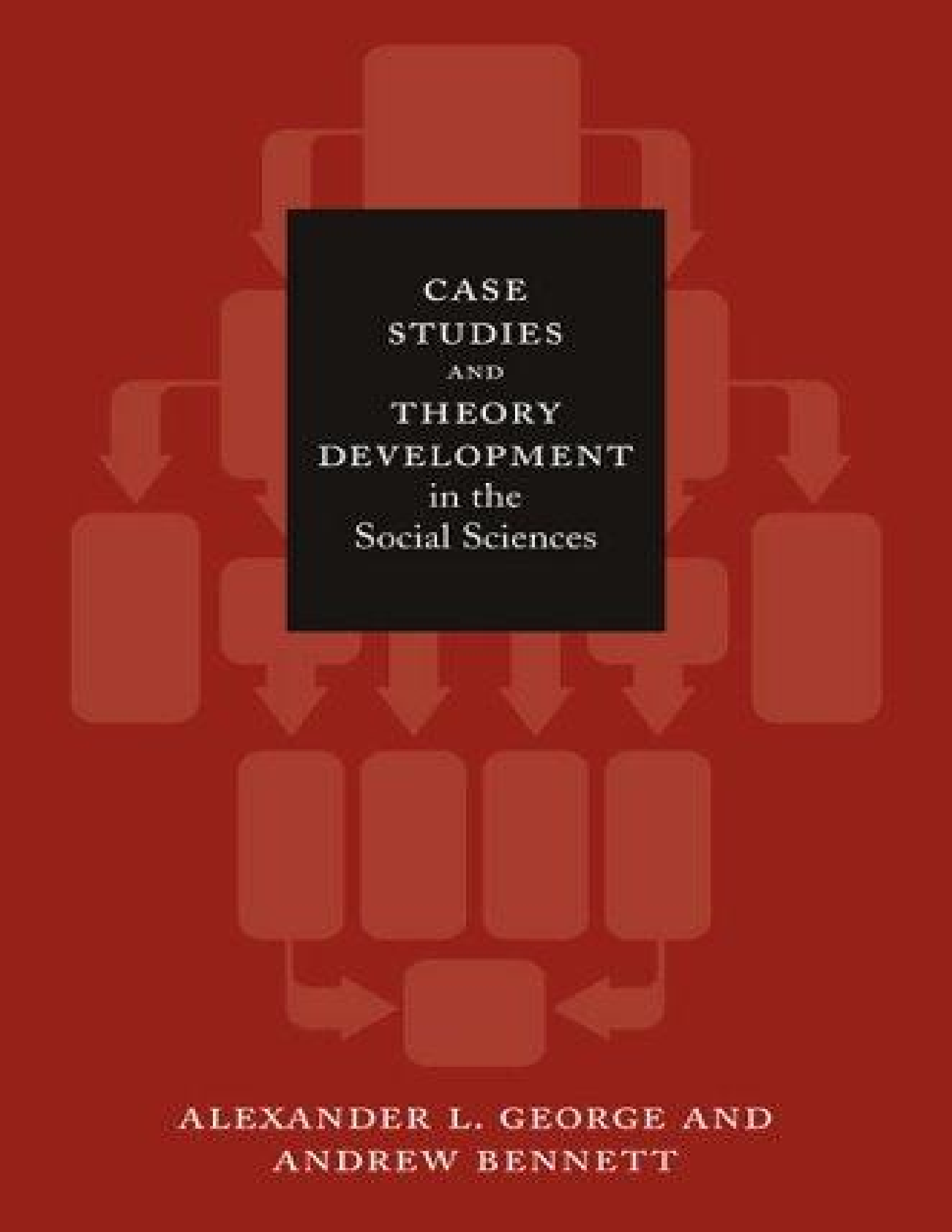

CASE
STUDIES
AND
THEORY
DEVELOPMENT
in the
Social Sciences
ALEXANDER L. GEORGE AND
ANDREW BENNETT

Table of Contents

Chapter 9

The Congruence Method

The congruence method occupies a special place in our conception of how a single case or a small number of cases can be used for theory development. As we noted in Chapter 8, the method of controlled comparison requires the investigator to find two cases similar in every respect but one. Since this requirement is difficult to meet, an alternative approach is often needed—one that does not attempt, as a controlled comparison does, to achieve the functional equivalent of an experiment. The alternative we propose is the within-case method of causal interpretation, which may include congruence, process-tracing, or both, and which does not operate according to the structure or causal logic of experiments. This chapter discusses the congruence method, and we turn to process-tracing in Chapter 10.

The essential characteristic of the congruence method is that the investigator begins with a theory and then attempts to assess its ability to explain or predict the outcome in a particular case. The theory posits a relation between variance in the independent variable and variance in the dependent variable; it can be deductive or take the form of an empirical generalization. The analyst first ascertains the value of the independent variable in the case at hand and then asks what prediction or expectation about the outcome of the dependent variable should follow from the theory. If the outcome of the case is consistent with the theory's prediction, the analyst can entertain the possibility that a causal relationship may exist. Of course, the finding of mere consistency between a theory's predictions and case outcomes may not be significant, and in this chapter we discuss several questions that can guide researchers as they assess the significance of preliminary findings.[376]

The congruence method has several attractive features. The investigator does not have to trace the causal process that leads from the independent variable to the case outcome; so the method does not require a great deal of data about the case being studied. Because the congruence method does not use process-tracing, it does not require a search for data that might establish a causal process from independent to dependent variables. (However, process-tracing can be combined with the congruence method to assess whether the congruence between independent and dependent variables is causal or spurious and also to enrich theories that only posit a relationship between independent and dependent variables and have nothing to say about the intervening variables and causal process that connect them.)

The congruence method offers considerable flexibility and adaptability. It can contribute to theory development in several ways; it can be employed in a disciplined-configurative type of case study, a plausibility probe, or in a crucial case (or tough test) of an existing theory.[377] The theory employed in the congruence method may be well-established and highly regarded, or it may be formulated or postulated by the investigator for the first time on the basis of a hunch that it may turn out to be important.

Often, however, available theories lack clarity and internal consistency so that they cannot make specific predictions and thus cannot be tested in any rigorous way. Nonetheless, investigators often succumb to the temptation to attribute predictive or explanatory power to such theories, leading to spurious or inconclusive tests of loosely formulated theories. The priority is not to test such theories, but to refine them if possible so that they can be tested. The congruence method may contribute to such refinement and development. An investigator may be able to clarify and refine a theory through its use in case studies, making it more nearly testable. As noted in Chapter 4, an investigator must establish the level of concreteness and differentiation with which variance in the dependent variable will be measured. How well this task is performed may well determine whether one can find congruence between the independent variable in the theory and outcomes on the dependent variable. This point is demonstrated later in this chapter.

A final attractive feature of the congruence method is that it can be used either as a within-case method or, when coupled with a counterfactual case, as a form of controlled comparison. The latter possibility is discussed later in this chapter.[378]

An important general standard for congruence tests is "congruity": similarities in the relative strength and duration of hypothesized causes and observed effects.[379] This does not mean that causes must resemble their effects or be on the same scale, and researchers must avoid the common bias toward assuming this should be the case. For example, there is a temptation to assume that large or dramatic effects must have large and dramatic causes, but this is not necessarily true. Researchers must take into account theoretical reasons why the effects of hypothesized causes might be amplified, diminished, delayed, or sped up (through expectations effects). Once this has been done, it is possible to address the question of whether the independent and dependent variables are congruent; that is, whether they vary in the expected directions, to the expected magnitude, along the expected dimensions, or whether there is still unexplained variance in one or more dimensions of the dependent variable.

Although consistency between a theory's predictions and case outcomes is often taken as providing support for a causal interpretation (and, for that matter, for assessing deductive theories generally), researchers must guard against unjustified, questionable imputation of a causal relationship on the basis of mere consistency, just as safeguards have been developed in statistical analysis to deal with the possibility of spurious correlation.

There are several ways in which this problem can be addressed. The investigator can employ process-tracing to attempt to identify a causal path (the causal chain) that depicts how the independent variable leads to the outcome of the dependent variable. (We note the close connection between process-tracing and causal mechanisms in Chapter 7.)

The usefulness of combining the congruence method with process-tracing was demonstrated in the innovative study by Yuen Foong Khong, *Analogies at War.* Earlier examples of the use of process-tracing in case studies to elaborate (or assess) the causal standing of an explanation first derived by applying a deductive theory include the studies by Vinod

Aggarwal in *Liberal Protectionism,* and by David Yoffie in *Power and Protectionism: Strategies of the Newly Industrializing Countries.*[380] (The studies by Khong and Aggarwal are discussed later in this chapter.)

Another way in which the investigator can attempt to deal with the limitations of the congruence method is to provide a plausible or convincing argument that the deductive theory or empirical generalization being employed is powerful and well validated, that it fits the case at hand extremely well, and that it is not rivaled by competing theories or at least is better than conceivable alternative theories. By invoking the superior standing of the theory employed or by resorting to process-tracing, the investigator may be satisfied that the within-case approach suffices and need not be buttressed by across-case comparisons.

When an investigator lacks confidence in the results of the congruence method employed in the within-case mode, he or she may supplement it by making use of counterfactual analysis. That is, the investigator invents a new case that is presumably similar to the original case in every respect but one (keeping in mind the limitations of counterfactuals discussed in Chapter 8).

The next section discusses the concepts of spuriousness, causal priority, and causal depth, three possible relationships between independent and dependent variables that researchers should consider as they assess preliminary findings that the outcome in a case is congruent with a theory. The two sections that follow provide more specific advice on how researchers can assess whether a finding is spurious and whether the independent variable is a necessary condition for the outcome of the dependent variable. We then discuss how the congruence method can be used to assess the causal role of beliefs in decision-making, highlighting the difficulty of ascertaining how decision-makers come to their decisions and noting how several scholars have coped with this challenge. Finally, we consider how the congruence method can be used to add to studies of deductive theories that put a "black box" around decision-making and strategic interaction, emphasizing the usefulness of process-tracing as a way to strengthen results by identifying a causal process that could lead from the independent to the dependent variable.

Spuriousness, Causal Priority, and Causal Depth

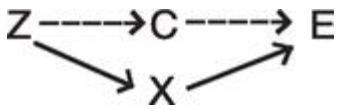
To assess the possible causal significance of congruity in a case, the researcher should ask two questions inspired by the logic of experiment. First, is the consistency spurious or of possible causal significance? Second, is the independent variable a necessary condition for the outcome of the dependent variable, and how much explanatory or predictive power does it have? The latter question is important, since a condition may be necessary but still contribute little to the explanation or prediction of the outcome in question.

Except for tests of deterministic theories stated in terms of necessity and sufficiency, a single congruence test is not strong enough to provide confirmation or falsification of theories.[381] More than one theory may be equally congruent with the outcome, or the outcome may be caused by other factors not identified by any of the theories considered. Researchers must be sensitive to the issues of spuriousness, causal priority, and causal depth in judging the strength of inferences made on the basis of congruence tests. A few comments on each of these three issues are needed. *Spuriousness* occurs when the observed congruence of the cause C and effect E is artificial because both C and E are caused by some third factor Z (whether or not Z has been identified in a competing theory):

Alternatively, the putative cause C lacks *causal priority* if C is necessary for E, but C is itself only an intervening variable wholly or largely caused by a necessary prior variable Z. In this instance, both Z and C are necessary for E, but C has no independent explanatory value:

Z–→C–→E

A third possibility is that C can be defined as lacking *causal depth* if a third variable Z would have brought about E even in the absence of C. In this instance, it does not matter whether or not Z is related to C. In other words, Z has greater causal depth because it appears to be necessary and sufficient for E, and Z may act through C or through some other variable X. In contrast to the example of causal priority, C is not in this instance a necessary condition for E.[382]



Thus, the appearance of congruence, especially when only one or primarily one theory is considered, cannot support an inference of causality, nor does the lack of congruence deny a possible causal role. Moreover, even if a congruence test suggests that a variable played a causal role in a given case, this does not mean that this theory proposes causal factors that are necessary, sufficient, or causal in any sense in other cases where contextual and conjunctive variables are different.

These problems of spuriousness, causal priority, and causal depth underscore that congruence tests by themselves may be inconclusive when several competing theories are involved. In such circumstances, for causal relations short of necessity or sufficiency, congruence tests are very difficult unless all the effects of the theories in question have been established with precision and confidence through previous testing. The problem is that alternative theories may focus on the same independent variables but point to different causal mechanisms that relate these variables to the observed outcome. The theories compete in logic, but may or may not make different predictions on the outcome. Theories may also be complementary, addressing different variables without contradicting one another logically. Such complementary theories may either reinforce or counteract one another's predicted effects.

A real-world example, drawn from Andrew Bennett's research on the rise and fall of Soviet military interventionism in the Third World in the 1970s and 1980s, illustrates these issues. The "Reagan Doctrine" and "Soviet economic stringency" explanations for Soviet retrenchment in the 1980s are complementary, and they both pointed toward an increased likelihood of Soviet retrenchment. Retrenchment occurred, but the congruence method alone cannot tell us if both explanations were important factors, if only one was primarily responsible for the outcome, or if neither was causal and the result was driven by other variables.[383]

Now consider the problem of competing explanations. A competitor to the Reagan Doctrine theory is the "hard-line reactive theory," which holds that the Reagan Doctrine aid, rather than speeding up Soviet retrenchment, galvanized a hard-line coalition in the Soviet Union and delayed the retrenchment in Soviet foreign policy.[384] These competing views on the effects of the Reagan Doctrine complement the economic stringency view and are consistent with the outcome of retrenchment. The difference is that the Reagan Doctrine theory suggests that U.S. aid to Afghan rebels, in addition to Soviet economic constraints, led to the Soviet withdrawal from Afghanistan; the hard-line theory could suggest that Soviet economic constraints, despite the delays and hedging caused by the hard-line coalition, caused the Soviet withdrawal. These competing versions can be tested for congruence with the timing, nature, and completeness of the Soviet withdrawal.

This example also illustrates why it is important not to summarily dismiss explanations that seem inconsistent with the outcome. In this case, trends in Soviet forces for power projection appeared to be inconsistent with the Soviet retrenchment, as these forces actually grew through much of the 1980s. However, the strengthening of these forces might help explain why Soviet retrenchment did not take place sooner or more precipitously.

Bennett's research on Soviet interventionism also employed an additional kind of congruence test. The research objective was to test a relatively new theory, learning theory, as an explanation for patterns of Soviet military intervention. This required first establishing whether there was any unexplained variance after accounting for the combined effects of more

established theories. Bennett thus canvassed these theories and assessed their individual and collective congruence with both the rise and fall of Soviet interventionism. Bennett concluded that these theories collectively provided a more complete explanation of the rise of Soviet interventionism in the 1970s than of its fall in the 1980s (which is consistent with the fact that many analysts in the late 1970s expected such interventionism to continue to increase). This test suggested that it was not possible to reject out of hand that a learning explanation might account for some of the variance in Soviet policies.

Multivariate congruence testing can be complex, but it is also a familiar form of historical analyses and arguments. One historian may argue that the structure of the international system and the bipolar distribution of power between the United States and the Soviet Union made the Cold War inevitable. Another may argue that the Cold War arose from not just the distribution of power, but also from the specific domestic political dynamics in the United States and Soviet Union and despite the lack of any immediate danger of a military invasion by one superpower against the other. A third might argue that this balance of contributing and counteracting forces underdetermines the emergence of the Cold War unless one takes Stalin's personality into account.

Two injunctions can help clarify such debates. First, it is important to consider a wide range of potentially causal factors, to specify the predicted contributing and counteracting effects of each, and to identify where underlying causal arguments are complementary and competing. Second, it is useful to guard against the bias of what has been termed "explanatory overdetermination."[385] When called upon to predict events, theorists and experts often give underdetermined accounts, yet when these same observers are asked to explain past events, their accounts make these events seem overdetermined. For example, almost no scholars predicted the collapse of the Soviet Union and the end of the Cold War, but afterwards many scholars pointed to numerous, seemingly overdetermining "causes" of these outcomes. Careful use of congruence testing, and inclusion of all the candidate theories, might instead lead to the conclusion that these outcomes were underdetermined, or at least that their timing and particular course could have been quite different if a few variables had been changed.

We now discuss how researchers can assess their preliminary findings of congruity.

How Plausible is the Claim of Congruity?

The possibility that consistency between the values of the independent and dependent variable in a given case is not spurious—and possibly causal—gains a measure of support if the relationship can be supported by a general law or statistical generalization. For example, a causal inference drawn from the observed consistency between an independent cognitive variable(s) such as the actor's *belief* and some aspect of that individual's *behavior* can be supported by psychological theories of cognitive balance that call attention to the fact that individuals generally (at least under certain conditions) strive to achieve consistency between their beliefs and their actions. This, of course, is a very general theory. If more specific generalizations or theories could be adduced, the imputation of a causal relation would be strengthened. Typically, the stronger and more precise the version of a more general theory, the more confidence we ought to attach to claims that consistency is not spurious.[386]

Is the Independent Variable a Necessary Condition for the Outcome of the Dependent Variable?

If the consistency identified appears to be causal and not spurious, the investigator may attempt to assess whether the independent variable is a necessary condition for the outcome in question. This question, of course, may be difficult to resolve. Efforts to do so will require the investigator to move beyond within-case analysis. Ideally, one would try to find other cases in which the same type of outcome occurred in the absence of that independent variable. If such a case(s) were discovered, then the independent variable could not be regarded as a necessary condition.[387]

When one or more comparable cases are not available, then the investigator can resort to analytical imagination to think of hypothetical cases that might help to judge whether the same type of outcome might occur in the absence of that independent variable. In other words, the investigator resorts to counterfactual analysis and mental experiments in an effort to create a controlled comparison.[388] Disciplined use of analytical imagination will at least provide a safeguard against the temptation to move too quickly and confidently from the earlier judgment that consistency was not spurious to the further inference that the independent variable is a necessary condition for the occurrence of that type of outcome.[389] If the grounds for regarding the independent variable as a necessary condition are shaky or dubious, as is often likely to be the case, then it is advisable to claim no more than that the type of independent variable in question appears to favor—make more likely—the occurrence of a certain type of outcome. In other words, the independent variable is a contributing cause, though neither necessary nor sufficient.

Analysts should also address the question, "Is the independent variable that is causally related to this particular outcome of the case also consistent with *other* possible outcomes?" In the analysis of a single case, history

provides only one outcome of the dependent variable. Accordingly, it is easy to overlook the possibility that other outcomes, had they occurred, might *also* have been consistent with the value of that independent variable. Once again, if the investigator cannot locate cases in which the independent variable with the same value was accompanied by diverse outcomes, he or she can resort to disciplined imagination to assess this possibility. To do so, the investigator should immerse himself or herself in the rich details of the historical case being examined; this may enable him or her to envisage with greater confidence that the outcome might well have gone in different directions even with the independent variable held constant, had variation occurred in other operative independent variables. If there is reason to believe this might have been so, the investigator must assign weaker general predictive and explanatory power to the independent variable in question. It should be noted that broadening the assessment of the causal status of the independent variable (or theory) in question requires that the investigator take into account that other independent variables in the case may have played a role in producing that outcome.

Still another question can be asked: "Is it possible to conceive of any outcomes of the historical case that would *not* have been consistent with the independent variable?" Investigators should attempt to identify outcomes that would be inconsistent with the independent variable and associated conditions because this highlights the need to construct falsifiable theories. By immersing oneself in the historical case, the investigator might envisage a number of other possible outcomes interestingly different from the historical outcome that would also have been consistent with the implications of the independent variable. If so, then the independent variable (of the deductive or empirical theory in question) may be part of the explanation, but its ability to discriminate among alternative outcomes and its predictive power are much weakened.[390] On the other hand, if the investigator cannot envisage other outcomes that could also plausibly occur in the case in question, then there would be reason to attribute stronger predictive power to the independent variable or theory of which it is a part.

Similarly, if all or many of the conceivable outcomes would be consistent with the theory, then its explanatory power may be limited or negligible. Conversely, if other outcomes might have occurred that were not consistent

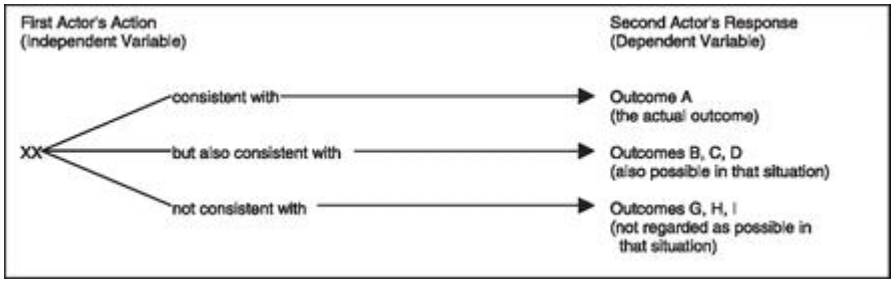
with the theory, then the investigator has additional presumptive evidence of the explanatory power of the theory at least for the actual or the other conceivable outcomes identified.

A hypothetical example will illustrate and clarify how questions of this kind, which attempt to replicate the logic of controlled experiment, can contribute to making more refined and more valid causal interpretations in single-case analysis.

In our hypothetical example, the first actor takes an action (independent variable XX*)* that appears to have a particular impact on the second actor's behavior (outcome A). The investigator finds that independent variable XX (but not YY or ZZ) is consistent with outcome A. The investigator now asks whether XX can explain and predict *only* outcome A. Or would outcomes B, C, and D—outcomes that did not occur in this case—also have been consistent with XX? If so, while XX may be part of the explanation, its explanatory (and predictive) power is diminished since other explanatory variables are needed to round out the explanation of why the second actor's response was A (and not B, C, or D). These interpretations of the explanatory power of XX are summarized in Figure 9.1.

A more refined analysis is possible. Suppose that although outcome A differs in interesting respects from outcomes B, C, and D, all four outcomes share a certain characteristic—for example, that all are conciliatory responses by the second actor to the first actor's action (though the precise nature of the conciliatory response varies). Suppose further that out-comes G, H, and I are all hard, refractory responses to the first actor's behavior. If so, then XX acquires added explanatory and predictive power of a quite useful kind, for it discriminates between conciliatory and refractory responses (though not by itself between variants of a conciliatory response).[391]

Figure 9.1. Possible Outcomes of an Independent Variable.



From this hypothetical example we turn to a more general discussion of using the congruence mode to assess the causal role of an actor's beliefs in his or her decision-making.

Use of the Congruence Method to Assess the Causal Role of Beliefs in Decision-Making

Specialists who focus on decision-making approaches in the study of foreign policy have long emphasized the importance of cognitive variables.[392] Attention has centered on how decision-makers' general beliefs about international politics can affect their choices of policy. However, important methodological issues arise in attempting to assess the role that such beliefs play in two different phases of the process of decision-making: the processing of information and analysis that *precedes* the decision taken, and the *actual choice of policy.* The foregoing discussion of the congruence method is relevant for addressing these issues.

General support for the assumption that a policymaker's beliefs about international politics influence his or her decisions is provided by cognitive consistency theory. But an individual's beliefs and behavior are not always consistent with one another for various reasons. While a decision-maker's beliefs play an important role in information processing that precedes actual choice of action, variables other than these beliefs affect the choices made. For example, the policymaker's decisions will likely be influenced by the need to obtain sufficient support for whatever policy he or she decides upon, by the need for compromise, by domestic or international constraints on the leader's freedom of action, etc. These factors may run in a direction that significantly modifies or is contrary to his or her preferred option.

It is more useful, therefore, to regard an individual's general beliefs as introducing two types of *propensities,* not determinants, into his or her decision-making: diagnostic propensities, which extend or restrict the scope and direction of information processing and shape the decision-maker's diagnosis of a situation; and choice propensities, which lead him or her to favor certain types of action alternatives over others (but which may give way or be altered in response to decisional pressures).

Thus, psychological consistency theory cannot by itself provide robust support to conclusions from congruence method studies of the role of beliefs in decision-making. Causal interpretations in such studies must be disciplined by the methodological questions noted above.

STEPHEN WALKER'S STUDY OF HENRY KISSINGER

Confidence that consistency between an individual's beliefs and actions is of causal significance is enhanced if it is encountered repeatedly in a sequence of decisions taken by an actor over a period of time. This observation played an important role in Stephen Walker's pioneering study of the role of Henry Kissinger's beliefs in his negotiations with North Vietnamese leaders.[393] In this study, Walker developed highly systematic and explicit methods for employing the congruence procedure. He also addressed the important question of whether Kissinger's actions were better explained by situational or role variables than by his beliefs. Walker advanced a plausible argument that Kissinger's operative beliefs were idiosyncratic in important respects and not easily accounted for by situational or role variables. That is, the set of Kissinger's beliefs and his policy actions consistent with those beliefs probably would not have been displayed by anyone else in his position. Walker noted that the Nixon administration's policy on Vietnam was controversial and that there were policy preferences that competed with Kissinger's. Moreover, the position of national security adviser that Kissinger occupied at that time was not precisely defined. This permitted the incumbent considerable latitude. For these and other reasons, Walker concluded, Kissinger's role in the prolonged bargaining process with North Vietnamese leaders exemplifies both "action indispensability" and "actor indispensability" as defined by Fred Greenstein.[394]

KHONG'S STUDY OF HISTORICAL ANALOGIES

The causal role of beliefs in decision-making was the subject of an exemplary study by Yuen Foong Khong.[395] Khong decided to focus not on operational code beliefs, as Stephen Walker had, but rather on the role historical analogies play in policymaking. Khong confronts the nettlesome problem of how the analyst can decide whether historical analogies are used by policymakers merely to *justify* decisions they take or whether analogies actually have a causal impact on the information processing that precedes decisions and the choice of a policy option. Drawing on Alexander George's "Causal Nexus" paper, Khong assesses the role of several historical analogies held by top-level U.S. policymakers at critical junctures of the Vietnam crisis: the February 1965 decision to initiate slow-squeeze graduated air attacks on North Vietnam and the July 1965 decision to expand substantially the deployment of U.S. combat forces.

In analyzing these two decisions, Khong examines three historical analogies of previous crises that U.S. policymakers were familiar with: Munich, the Korean War, and Dien Bien Phu. He finds evidence in historical materials and from interviews that each of these analogies was present in the minds of U.S. policymakers in 1965. However, by means of an ingenious and complex research strategy that uses both the congruence method and process-tracing, Khong concludes that the Korean analogy played the most influential role in U.S. decisions to use slowly graduated air attacks and then to put in large-scale ground forces.

Only a brief account of the essence of his rich analysis can be presented here. First, Khong built on the distinction mentioned above between diagnostic propensities and choice propensities that are implicit in the beliefs held by policymakers by distinguishing six different but closely related diagnostic tasks. (Although he labels all six tasks as "diagnostic," they do include choice propensities; in effect, he collapses the distinction

between diagnostic and choice propensities.) Khong emphasizes that historical analogies are often used by policymakers to perform diagnostic tasks.

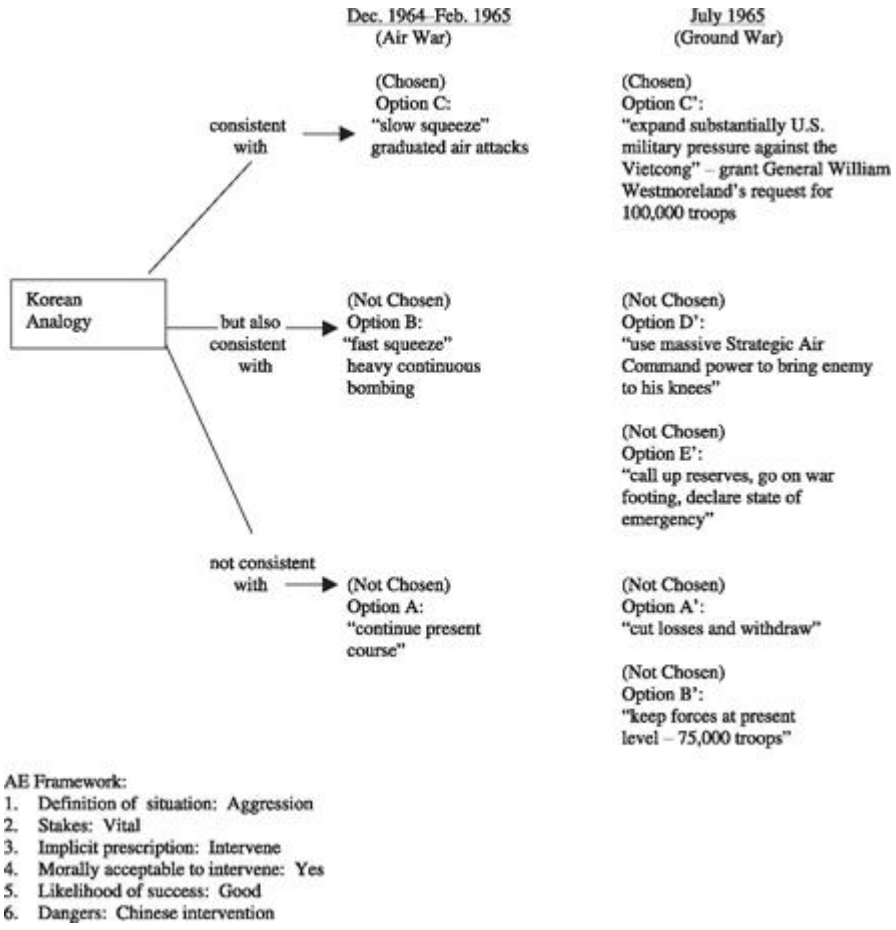
His six diagnostic tasks are: a definition of the new situation, facilitated by comparing it with a past one; a judgment of what is at stake; an implicit prescription as to how the new situation should be dealt with—i.e., the "solution" to the problem or type of policy response needed; an assessment of the moral acceptability of the implied prescription; an assessment of the likelihood of its success; and an estimate or warning of the dangers and risks of the implicit policy should it be adopted.

Khong labels this set of diagnostic tasks the Analogical Explanation (AE) Framework. He converts these six diagnostic tasks into a set of general standardized questions to be asked of each of the historical analogies; these are a central feature of his research design.[396] The answers to these questions satisfy the data requirements for comparing the role the analogies played in information processing. The study, therefore, constitutes an explicit example of the method of structured, focused comparison: it is only by asking the same general questions of each case that systematic comparison becomes possible.

Khong establishes the implications that each of the three historical analogies had for these diagnostic tasks via process-tracing by making a careful analysis of the available historical record and through interviews with U.S. policymakers. He then employs the congruence method to assess the implications of each analogy's answer to the six diagnostic tasks for the various policy options that were being considered at the time.

The question for Khong, then, was which of the various policy options under consideration were consistent with the diagnostic implications of the analogy and which were not. Khong employs a version of the congruence method discussed earlier in this chapter for each of the historical analogies. We reproduce in Figure 9.2 his analysis for the Korean analogy.[397]

Figure 9.2. The Lessons of Korea and the Option Chosen.



Having established the answers to the diagnostic tasks each analogy suggested, Khong then looks for congruity between an analogy's diagnosis and the policy options that were under consideration by policymakers. According to Khong's analysis, the Korean analogy's answers to the six diagnostic tasks were highly consistent with the policy decision actually taken from December 1964 to February 1965 period to employ a "slow squeeze" version of graduated air attacks. But it was also consistent with a policy option calling for heavy, continuous bombing that was not taken. This left unanswered for the moment why the "slow squeeze" version of air attacks was chosen. A further challenge for analysis was raised by Khong's finding that the Munich analogy had identical implications for these two policy options. Similar results emerged when the congruence method was used to compare the implications of the Korean and Munich analogies for the various policy options under consideration in July 1965.

Thus, as Khong notes, both historical analogies supported the case for either of the two options. But Khong argues persuasively that the Korean analogy was more influential in the two decisions of February and July. He arrives at this conclusion by attributing decisive importance to the different answers the two analogies provided for the sixth diagnostic task. The Korean analogy carried a strong fear that resort to the stronger of the two options in both February and 1965 would trigger Chinese intervention in the Korean War. This particular vision of the Korean War was deeply etched in the historical memory of U.S. policymakers in 1965. Khong cites ample evidence from archival and interviews in support of this observation.

In contrast, the Munich analogy did not warn of the dangers of making a hard response to aggressions by the Japanese and Germans in the 1930s. Although the Munich analogy could account, as did the Korean analogy, for the rejection of the nonintervention options in 1965, it was unable to suggest why, among the intervention options, the least hard one was selected.[398]

In this excellent study, Khong has shown how an imaginative, disciplined research design that combines congruence and process-tracing methods can be used to confront the extremely complicated, difficult task of distinguishing between a justificatory role and an information processing function of historical analogies in foreign policy decision-making. His study is the most rigorous and disciplined treatment we know of for dealing with the theoretical and methodological issues associated with determining whether historical analogies are being used by policymakers solely to justify their decisions or whether the analogies play a genuine causal role in the information processing that leads to the decisions taken. Khong states his conclusions with appropriate cautions, noting a number of limitations and questions that remain, but he has raised the discussion of this difficult problem to a new level of analytical sophistication.[399]

RITTBERGER'S STUDY OF GERMANY'S POST-UNIFICATION FOREIGN POLICY A study organized by Volker Rittberger also employed both the congruence method and process-tracing, this time to assess competing theories for predicting German foreign policy after the unification of the two Germanies. [400] The bulk of literature on this question

predicted that post-unification German foreign policy would be dominated by the question of whether its improved power position should lead to a significant change in its foreign policy. The research question posed in Rittberger's study was whether there would be continuity or significant change in post-unification foreign policy. Three theories were formulated and submitted to a carefully constructed empirical test: neorealism (and a modified version of it that introduced variation in security pressures); utilitarian liberalism; and constructivism (which holds that state actors follow a logic of appropriateness whose behavior is shaped by international and societal norms).

To conduct an empirical test of these three theories, the authors selected four issue areas that provide a representative cross-section of German foreign policy and that include both issues of "high politics" and "low politics." These are German security policy within NATO; German constitutional policy vis-à-vis the European Union; German foreign trade policy within the European Union and the General Agreement on Tariffs and Trade (GATT); and German human rights policy within the United Nations. The research design included a before-after component that enabled the authors to evaluate the extent to which post-unification Germany changed its foreign policy behavior. Three independent variables were included in the research design: power position, domestic interests, and social norms. The methodology of structured, focused comparison was employed in a series of case studies, each consisting of one or more observations of *post*-unification policy on a particular issue and one or more observations of *pre*-unification foreign policy on the same issue.

The congruence procedure was the centerpiece of the research design. The degree of consistency between a theory's predictions and the observed values of the dependent variable was regarded as the most important indication of its explanatory power. This test was employed in a differentiated manner that took into account tough tests and easy ones, dealt with instances in which several theories made correct predictions, and evaluated evidence based on additional observable implications a theory was able to make. These additional observable implications were studied via process-tracing, except for the implications of neorealism, which does not lend itself to the process-tracing procedure.

Post-unification German foreign policy was found to display a mixture of continuity and change. The evaluation of each theory called attention to its successful and unsuccessful predictions. The study found that the eight cases examined strongly disconfirmed neorealism. The modified variant of neorealism did better. Social norms associated with constructivist theory turned out to yield the best explanation of post-unification German foreign policy, capturing both cases of continuity and change as well as hard and easy tests. Liberalism's explanatory power seemed to depend on the policy network structure that dominated in a particular issue.[401]

Use of the Congruence Method in Studies of Deductive Theories that "Black Box" Decision-Making and Strategic Interaction

The congruence method can be useful also in the studies that work with deductive theories that "black box" decision-making or strategic interaction. Such studies employ a deductive theory to make predictions of outcomes in a single case or in a number of cases too few to permit statistical analysis. The research objective is often to test the performance of the deductive theory in question or to identify and bound its scope. If its performance proves to be inadequate—i.e., a number of incorrect predictions occur that can not be attributed to measurement errors—then one must ask whether the internal structure or contents of the theory are flawed and in need of reformulation. If so, the congruence method may be used to develop and refine the provisional theory.

These uses of the congruence method have been applied in international relations studies that work with structural-realist, rational choice, or game theories, all of which involve black box decision-making and strategic interaction, and also in studies that directly examine internal decision-making processes and the dynamics of strategic interaction. Use of the congruence method (though it is not known by this name) also is employed in small-n case studies that focus on theories of macro-political processes.

What is involved in using the congruence method in research projects which, as an initial simplification, black box or set aside internal processes of decision-making or strategic interaction? The first step is to formulate a version of the general deductive theory being employed—whether it be structural realism, rational choice, or game theory—that addresses more specifically the phenomenon being studied. This first step can be noted in studies such as those by Barry Posen, Vinod Aggarwal, David Yoffie, and Bruce Bueno de Mesquita.[402]

A second step is to identify historical cases whose outcomes will enable the investigator to apply the congruence method to test, assess, or refine the theory's predictive and explanatory power. Selection of cases is a critical decision in research design and it is discussed in detail in Chapter 4. Suffice it to note here that the investigator must avoid "selection bias" and be clear about whether a representative sample of the universe of cases of the phenomenon is necessary to satisfy the research objective and to reach an acceptable statement of the nature and scope of the findings. It is a common misunderstanding to assume or to insist that all small-n studies must somehow satisfy the requirement of a representative sample, and that the findings of a small-n study must be capable of projecting a valid probability distribution of outcomes for the entire universe.[403]

A third step is to match the predictions and expectations of the theory with the outcomes of the cases to see if they are consistent. If consistency is noted, then the investigator should address the several questions that were discussed earlier in this chapter regarding the causal significance that can be properly inferred from congruence. Outcomes not consistent with the predictions and expectations of the theory should receive special attention. How can one account for these discrepant cases? How can the possibility of measurement error be correctly assessed, and how can that be distinguished from the possibility that the internal composition and logic of the deductive theory are faulty?

A fourth step is possible and we strongly recommend that it be undertaken. Process-tracing of the case should be employed for several purposes: to help assess whether the consistency noted is spurious or causal; to identify any possible intervening causal process that connects the deductive theory with the case outcomes; and to provide an explanation for deviant cases that the theory failed to predict correctly. Process-tracing was used for these purposes by Aggarwal, Yoffie, and Posen, but not by Bueno de Mesquita.

AGGARWAL'S STUDY OF TRADE REGIMES

Aggarwal's study was one of a number of studies in political economy undertaken by Ph.D. students at Stanford under the direction of Robert Keohane (with Alexander George serving as a second reader). The starting point for all these studies was the assumption that the best way to study problems of trade relations between the United States and its weaker trading partners (and also to study the development and possible transformation of international trade regimes) was to adapt structural realist theory for the specific issue-area and actors involved. (This assumption was substantially modified as students encountered the problem of developing causal inferences and explanations for outcomes of trading episodes.) The initial research design focused on the relative power advantage the United States possessed which, according to structural realist theory, should lead to outcomes favorable to the United States. When such favorable outcomes in trading episodes occurred, it might be assumed that realist theory provided an adequate explanation and could have predicted these outcomes.

However, Aggarwal realized that mere congruence of outcomes with the general predictions and expectations of structural-realist theory did not necessarily provide a reliable explanation—that it was not an adequate test of the theory. Hence, Aggarwal engaged in process-tracing of each trading episode to ascertain whether he could identify a causal process that supported the role attributed to the structural variable. He felt it necessary to proceed in this fashion since it was not possible to undertake a large-N statistical study for this purpose. In addition, to understand and explain a number of those deviant cases in which the outcomes were *not* as favorable to the United States as its relative power advantage would have predicted, Aggarwal undertook a detailed analysis of the dynamics of the trading interaction and engaged in process-tracing to identify how the actors' decision-making and their strategic interaction in bargaining with each other might have led to an outcome not predicted by the theory.

Hence, Aggarwal was not satisfied to settle for the familiar fall-back position that structural realism is a probabilistic theory that does not claim to predict all cases successfully. Instead, Aggarwal attempted to explain discrepant cases and, if possible, to enrich and differentiate the theory. He referred to these cases as anomalies and argued that in the absence of a large number of cases to permit statistical analysis, "a second approach, known as 'process-tracing,' is an effective and potentially superior substitute. In process-tracing, the decision-making procedure in a negotiation is systematically analyzed with an eye to identifying the degree to which participants appear to respond to international systemic or other constraints."[404]

As the third and fourth steps emphasize, one should not be satisfied merely with a finding of consistency. Since the data required for adequate process-tracing are often not available, the checks regarding the causal significance of consistency noted earlier should be undertaken.

Congruence and Structural-Realist Theory

Studies that use structural-realist theory to predict outcomes are in special need of supplementary process-tracing or other checks. Kenneth Waltz's structural-realist theory is not a fully developed deductive theory; it can make only very general probabilistic predictions, since it does not quantify its probabilistic claims. Strictly speaking, a finding that the outcomes of cases are consistent with probabilistic predictions is not an adequate basis for assuming a causal relationship exists *unless* other explanations for the outcomes are considered and eliminated. And even when support for some kind of causal relationship can be mustered, one must still establish whether the independent variable is either a necessary or sufficient condition for the outcome in question, and how much it contributes to a full explanation of the outcome.

In other words, partial, incomplete deductive theories based on structural realism often lack "operationalization"—i.e., the fine-tuning and specification of the theory that permits case-specific rather than general probabilistic prediction of outcomes for each of the cases examined. The only fully operationalized variant of a structural realist theory of which we are aware is that developed by Bruce Bueno de Mesquita in *The War Trap.*

In striking contrast to *The War Trap* is the case that Christopher Achen and Duncan Snidal offered for rational deterrence theory. They made no effort to formulate the level of specification and refinement of the theory needed to make concrete predictions; therefore, the theory they provided was a quite primitive and nonfalsifiable deductive theory. That is, any outcome—whether deterrence succeeded or failed in particular cases — would be "explainable" by the vague rational deterrence theory they espoused. Even more disconcerting in the argument these authors made on behalf of the superiority of a rational deterrence theory was their failure to

address the requirements of a full-fledged, operationalized deductive theory.[405]

Even when operationalized, deductive theories may fail to identify or provide a satisfactory account of the causal mechanism that links the theory to the outcomes in question. Proponents of deductive theories based on rational choice or game theory might say that a causal mechanism is implicit in the internal logic of such deductive theories and needs no further explication or demonstration if the theory generates successful predictions. Yet some proponents of rational choice theory have recently emphasized the need to couple and integrate the rational choice framework with detailed case studies that make use of process-tracing in order to establish intervening causal processes.[406] We stated earlier that the congruence method applies not only to theories that focus on the causal role of beliefs in decision-making but, as has now been discussed, also to deductive theories associated with the structural realist theory of international relations and more generally to rational choice and game theories.