**Counterfactual:** A counterfactual assertion is a conditional whose antecedent is false and whose consequent describes how the world *would have been* if the antecedent had obtained. The counterfactual takes the form of a subjunctive conditional: “If P had obtained, then Q would have obtained”. In understanding and assessing such a statement we are asked to consider how the world would have been if the antecedent condition had obtained. For example, “If the wind had not reached 50 miles per hour, the bridge would not have collapsed” or “If the Security Council had acted, the war would have been averted.” We can ask two types of questions about counterfactual conditionals: What is the meaning of the statement, and how do we determine whether it is true or false? A counterfactual conditional cannot be evaluated as a truth-functional conditional, since a truth-functional conditional with false antecedent is ipso facto true. (That is, “if P then Q” is equivalent to “either not P or Q”.) So is it necessary to provide a logical analysis of the truth conditions of counterfactuals if they are to be useful in rigorous thought.

There is a close relationship between counterfactual reasoning and causal reasoning. If we assert that “P caused Q (in the circumstances C_i)”, it is implied that we would assert: “If P had not occurred (in circumstances C_i) then Q would not have occurred.” So a causal judgment implies a set of counterfactual judgments. Symmetrically, a counterfactual judgment is commonly supported by reference to one or more causal processes that would have conveyed the world from the situation of the antecedent to the consequent. When we judge that the Tacoma Narrows Bridge would not have collapsed had the wind not reached 50 miles per hour, we rely on a causal theory of the structural dynamics of the bridge and the effects of the wind in reaching the consequent.

How do we assign a truth value to a counterfactual statement? The most systematic answer is to appeal to causal relations and causal laws. If we believe that we have a true causal analysis of the occurrence of Q, and if P is a necessary part of the set of sufficient conditions that bring Q to pass—then we can argue that, had P occurred, Q would have occurred. David Lewis (1973) analyzes the truth conditions and logic of counterfactuals in terms of possible worlds (possible world semantics). A counterfactual is interpreted as a statement about how things occur in other possible worlds governed by the same laws of nature. Roughly: “in every possible world that is relevantly similar to the existing world but in which the wind does not reach 50 miles per hour, the bridge does not collapse.” What constitutes “relevant similarity” between worlds is explained in terms of “being subject to the same laws of nature.” On this approach we understand the counterfactual “If P had occurred, Q would have occurred” as a statement along these lines: “P & [laws of nature] entail Q”. This construction introduces a notion of “physical necessity” to the rendering of counterfactuals: given P, it is physically necessary that Q.


Encyclopedia of Social Science Research Methods, edited by Michael Lewis-Beck (University of Iowa), Alan Bryman (Loughborough University), and Tim Futing Liao. Sage Publications.