

**If Counterfactuals were Assessable, then a Method for
Assessment would Exist.**

Cassandra Maki
November 19, 2004

Counterfactual conditionals are a common occurrence in philosophical debate, the field of science, and in everyday conversation. Students, for example, can often be heard saying things such as “if I didn’t have to study, I’d go to the bar with you.” However, there is no systematic way in which counterfactuals are currently assessed, which raises the question of whether or not they even can be assessed. If philosophers found a convincing method of assessment for counterfactuals, they would be able to oppose (or support) arguments with more strength.

In this paper, I look at different possibilities for assessing counterfactual conditionals. In the absence of finding a suitable method, I will try to flush out the problem so it is clear that we do need a systematic method for assessing counterfactuals. One way that I hope to show works is the assessment of counterfactuals as condensed arguments.

Conditional statements have the form “if A, then B.” Counterfactuals are also in the form “if A, then B,” the difference is that the ‘A’ always expresses a false statement. Counterfactual statements are conditionals that hypothesize the way the world would be if one actual fact were different. Counterfactuals assert that A would imply B if A were the case.

The evaluation of counterfactual statements presents a problem because of their nature. We can presume that it is true that “if it were raining, my shirt would be wet,” but it is difficult (a) to state why it would be true and (b) to assess counterfactual statements that are not as intuitive as the above example.

Scientists may use counterfactuals as hypotheses for experiments. For example, suppose a scientist is setting up an experiment. He is trying to figure out if a certain

molecule is of a certain type, call it type T. If it is, it will react to light in the same manner as other molecules of type T. Before the experiment he says to his associate, “if this molecule were of type T, then it would expand if subjected to light.” Then, the two scientists would set up an experiment to see if the molecule is in fact a T.

Another way in which scientists use counterfactuals is, in a sense, by definition. A scientist might say that an item is soluble just in case ‘if I were to put this item in water, it would dissolve.’

To look at this issue from another angle, we can look at counterfactuals used in philosophy. Instead of acting as a hypothesis for an experiment or as a means of expressing or explaining a definition, counterfactuals in philosophy are often used as part of a thought experiment to outline a certain point or create a counterexample. The idea behind them is if someone can come up with one counterexample to a claim, the claim is false. For example, suppose someone makes the claim “if x is a person, then x is conscious.” A counterexample to this claim would be a person that is sleeping, for although x is not conscious, most would still claim that x is a person.

In order to better understand where counterfactuals can be important in philosophy and why it would be helpful to assess them, consider the following example from Michael Zimmerman:

(B) Jones wants to kill Smith. Jones intends to kill Smith. Right as Jones pulls the trigger, a bird flies in the way of his aim, causing a dead bird and a living Smith. (S) If the bird had not flown in front of the gun, Jones would have killed Smith. The claim, then, is to say that Jones is just as guilty (in degree) for killing Smith as he would have been if the shot had hit.

Counterexamples such as this often force people to reassess the theories to which they once clung. In this example, the theory that people wanted to cling to was that a

person is guilty for *actually* committing a crime, not just for intending to commit a crime. This example made people question that idea.

This case forces people to look at the justice system, or perhaps just our own moral code, and question whether or not people with the same degree of blame should be held morally accountable in the same way. It also shows that luck should not play a part in assigning moral responsibility¹. It was a matter of luck that the bird flew in the way, not a matter of choice on Jones' part. So, according to Zimmerman, Jones is just as morally responsible as if he had actually murdered Smith.

While I actually agree with the claim that Jones is just as blameworthy in degree for not killing Smith as he would have been if he had, it is difficult to assess whether or not the counterfactual (S) is true. While (S) brings up a good point and seems intuitively true, how are we able to assess such a claim? Because we cannot assess the counterfactual, this leaves only a reason to question the original theory, not to dismiss it. The reason that (S) is so important as opposed to the other sentences in this claim is because it is the only sentence that does not have a determined truth-value.

Now we are left with the task of actually finding a way to assess counterfactuals. This must be done because counterfactuals, while ambiguous at times, are important. As in example (B), they can make people question theories that they had previously never examined. However, I believe they could hold even more weight, and be justified in doing so, if there was a more mechanical, more systematic way in which to assess them. As it stands, counterexamples are weakened by their unassessability.

¹ Since this paper only briefly mentions luck and its involvement with moral responsibility, it masks a long, important discussion on the matter. One could discuss everything from it being a matter of luck that the bird flew to it being a matter of luck that Jones and Smith ever met. However, this discussion is outside of the scope of this paper; it is only mentioned to show Zimmerman's view. For a more in depth discussion on the role luck plays in ethical behavior and actions, see Thomas Nagel's "Moral Luck."

Now, some may claim that counterfactuals do not need to be assessed, that they serve their purpose without a standard mode of assessment. Imperatives and commands, for example, are statements that cannot be assessed logically since they have no truth-value, yet we use them constantly and without controversy. There is one major difference that is being overlooked, however. Imperatives and commands are not being used to rebuke arguments. While they have a place in everyday speech, they do not hold the weight that is put on counterfactuals to help justify or break down theories or long-held beliefs. In fact, they have little to no place in logic or argumentation at all.

In order to further illustrate the problem, let us try to evaluate some counterfactuals. Since counterfactuals are a type of conditional statement, one may evaluate counterfactuals with the same methods used for evaluating conditional statements. One way in which conditional statements are assessed are as material conditionals. A material conditional is true if either its antecedent is false or its consequent is true. If we apply this methodology to counterfactuals we have a problem, since all counterfactuals have false antecedents. This view from classical logic would render all counterfactuals true, which is surely not the case. For example, this criterion would show “if Nigel would have been President, we would have never gone to war,” and “if Nigel would have been President, we would have definitely gone to war,” both to be true. Obviously this doesn’t work, since the consequents are contradictory statements.

Another way conditional statements could be assessed is by looking at them as strict conditionals of modal logic. Modal logic assesses the necessary and possible truth and falsity of propositions by appealing to possible worlds. The easiest way to think of this is to imagine the way the world *could* be. For example, we can imagine a world in

which the birds all fly north for the winter. This is a possible world. All of the ways that the world *could* be designate possible worlds. In modal logic, a strict conditional is true if in all worlds in which the antecedent is true, the consequent *must* also be true.

If counterfactuals are assessed as strict conditionals, another problem occurs. Consider, for example:

(A) If fewer people smoked, fewer people would contract lung cancer.

This seems to be true, since so many people currently contract lung cancer from smoking cigarettes. However, we can imagine a possible world in which hardly anyone smokes, but using computers causes lung cancer. This seems to show that the counterfactual could be false; it is possible for the antecedent to be true and the consequent to be false.

We can even take the liberty of reducing the possible world scenarios to those worlds that are almost exactly like the actual world. For example, we could assert that we will only examine possible worlds in which the natural laws are the same. In this case we can still imagine a world where hardly anyone smokes, but where all the buildings are made out of asbestos. Even though this world has been tailored to mimic our own in its natural laws, etc, we can still see how (A) can be false. Since we were trying to assess A as a strict conditional, although (A) seems intuitively true, we have not found an assessment that would show it to be necessarily true.

Some may argue that counterfactuals should not be assessed in this way because reaching for an assessment of necessary truth is too strong. Another possibility, in regard

to Modal Logic, is to try to assess counterfactuals by looking at whether the consequent is true in the nearest possible world in which the antecedent is true. However, with this there is another problem—determining what the ‘nearest possible world’ is. In the asbestos case, all of the natural laws are the same; the buildings are just lined with different material. If that example does not work, how *do* we determine the nearest possible world?

If counterfactual statements cannot be assessed as conditional statements, another method of assessment needs to be fashioned. Without a systematic mode of assessment, counterfactuals are simply ambiguous statements whose ‘truth-value’ is based on intuition. However, there are possibilities for counterfactuals beyond conditional assessments.

In *Counterfactuals and Causal Law*, J. L. Mackie proposes that counterfactuals could be seen as condensed arguments. He states, “how can we describe non-imaginatively an imaginary situation? Only by inference from something else that we know or believe. And yet, to state a counterfactual is not to give an inference in full...A counterfactual, then, is a condensed and incomplete argument” (Mackie, 68). In other words, behind the words of every counterfactual is an argument that is not fully spelled out. I agree.

However, even after stating that counterfactuals are condensed arguments and filling in an argument, Mackie still says that they are “incomplete arguments”. He even says “we can understand it without being able to complete the argument” and “to advance a counterfactual is not to say that there is such an argument available” (Mackie, 68). These two ideas seem to be inconsistent, or at least in conflict with one another. On one

hand he is trying to convince his audience that counterfactuals are condensed arguments, on the other he is saying it cannot be done.

This is where I think he is wrong. I will show that counterfactuals *can* be made into full arguments, and, while I cannot look at every counterfactual, perhaps it will be made clear that this method can be expanded to encompass other counterfactuals than the two I have chosen. Furthermore, I think that for each counterfactual there is an argument available because each counterfactual simply represents an argument.

In order to utilize counterfactual statements to their fullest extent, those who use them should consider them condensed arguments. This procedure would be applied as follows:

1. An argument is given by A.
2. A counterexample including a counterfactual is given by B.
3. In order to fully understand the strength that the counterfactual could have on her argument, A examines the argument that is summarized by the counterfactual.
4. This argument (from the counterfactual) is then treated like any other argument in regard to validity, soundness, and/or strength.
5. If the argument is strong, A must rethink her original argument. If it is weak, she will wait for a better counterexample.

This can be applied to the examples that were discussed earlier. If (A) were expanded, for example, it would read:

- P1) Smoking causes lung cancer.
- P2) If P1, then other things being equal, people smoking helps to increase the number of people with lung cancer.
- C1) Other things being equal, people smoking helps to increase the number of people with lung cancer.
- P4) If C1, then if fewer people smoked, fewer people would contract lung cancer.
- C2) Therefore, if fewer people smoked, fewer people would contract lung cancer.

This argument is valid. P1 is based on known facts, research, and statistics. P2 simply means that, irrespective of anything else that causes lung cancer, smoking contributes to the number of people that acquire cancer. P4 is based on the idea smoking

contributes to the number of people acquiring cancer; not smoking would contribute to the number of people not acquiring cancer.

Another way to expand a counterfactual into an argument would be to make the antecedent premise one and the consequent the conclusion. This would show how one would logically form the connection between the antecedent and the consequent by setting up an argument that leads from one to the other. However, just setting this sort of argument up leads to problems of its own. For example, the antecedent would have to be reworded in order to be a premise, which would change the meaning of the antecedent. For that reason, I will not approach that method in this paper.

To more clearly explain the method I am trying to propose, let us look at another example. (B) can be expanded in the same matter as (A).

P1) Jones had Smith in his sights.

P2) Jones is an accurate shot.

P3) Jones' gun is in working order.

P4) If P1, P2, and P3, then if Jones had pulled the trigger, he would have shot Smith.

C1) If Jones had pulled the trigger, he would have shot Smith.

P5) A bird flew in front of Jones at the exact moment that he pulled the trigger.

P6) If C1 and P5, then if the bird had not flown in front of the gun, then Jones would have killed Smith.

C2) If the bird had not flown in front of the gun, then Jones would have killed Smith.

This argument is valid. Premises 1 through 4 simply describe the story. They set up the scene to show that Jones would have been able to shoot Smith under normal circumstances. P5 is also just another fact of the story. This argument shows how the facts of a scene can lead us to accept a counterfactual. By spelling the facts out, by way of an argument, it becomes clear why we should be inclined to accept the conclusion, even though it is contrary to fact.

Unfortunately, there may be a problem with expanding counterfactuals into arguments. In the two arguments that were expanded in this paper, a counterfactual ends up as the consequent of a conditional. The problem with this is the original problem, which is that we cannot assess the truth-value of the counterfactual, which of course means that we cannot assess the truth-value of that particular conditional. In other words, since there is a counterfactual in the consequent of a conditional statement, the conditional cannot be assessed as true or false. And, since the conditional cannot be assessed, the argument can never be evaluated as sound. So, although we can say that the arguments are valid, we cannot say that they are sound. However, I think that expanding counterfactuals in this way allows us to see more clearly why we should accept them. In a sense, it makes the assessment more educated instead of just intuitive.

Taking counterfactuals and expanding them into assessable arguments takes examples that seem intuitive and gives them logical weight. Intuitions are oftentimes the things that make us begin to question, but the intuitions of different people are often very different. Because of this, it is useful to find a more objective backing for our ideas. As is shown with (A), an intuition that leans toward true can be solidified with a strong, valid argument. In light of the possibility of solidification (or the opposite), this method of assessment should be implemented when dealing with counterfactuals.

A possible objection I see with this method is that, in essence, people already do this. While this, in a way, defends the expansion of counterfactuals for assessment, it also makes it essentially needless in practice. When faced with a counterfactual, we say that they are intuitively true or false because we mentally run through the line of reasoning that would make the consequent follow from the antecedent.

After looking at an example like B, I think it is clear that most people do not sit and go through that line of reasoning when faced with a counterfactual. However, I believe that implementation of this type of methodology will help people when they are faced with a counterfactual that is not intuitively one way or the other (or perhaps one that is, intuitions can certainly be wrong), since there will be a systematic way of breaking the counterfactual down.