Forthcoming in *Philosophical Studies*

How Fallacious is the Consequence Fallacy?

Wai-hung Wong & Zanja Yudell

In Williamson (2007), Timothy Williamson argues against "the tactic of criticizing confidence in a theory by identifying a logical consequence of the theory (not itself a logical truth) whose probability is not raised by the evidence" (232-233). He dubs it the *consequence fallacy*. In this paper, we will show that Williamson's formulation of the tactic in question is ambiguous. On one reading of Williamson's formulation, the tactic is indeed a fallacy, but it is not a commonly used tactic; on another reading, it is a commonly used tactic (or at least more often used than the former tactic), but it is not a fallacy. The two readings of Williamson's formulation of the tactic are:

(T1) Arguing that the probability of a theory is not raised by the evidence by identifying a logical consequence of the theory whose probability is not raised by the evidence.

(T2) Arguing that a theory is not made likely to be true by the evidence by identifying a logical consequence of the theory that is not made likely to be true by the evidence.[1]

Williamson's use of the phrase 'probability is not raised' suggests (T1), while his use of the word 'confidence' suggests (T2). (T1) is a fallacy, but Williamson's argument for its fallaciousness, though ingenious, is not completely satisfying. We will explain the weaknesses in his argument and try to improve on it. We will also show that (T2) is not a fallacy and explain why it is a more commonly used tactic than (T1).

Williamson's argument for the fallaciousness of the consequence fallacy — now understood as (T1) — is brief. Consider a theory h and evidence e. Assume that e is evidence for h in the sense that it raises the probability of h, although it does not make h certain. In symbols,

$Pr(h) < Pr(h \mid e) < 1$. Now the material conditional $e \rightarrow h$ is a logical consequence of h. Moreover, $e \rightarrow h$ is less probable given e, since e's truth eliminates the possibilities in which $e \rightarrow h$ is true whether or not h is true. Hence, $Pr(e \rightarrow h \mid e) < Pr(e \rightarrow h)$. Thus, there is always at least one logical consequence of a theory whose probability goes down (and hence is not raised) in the face of evidence that raises the probability of the theory. And so, it would be fallacious to conclude that $Pr(h \mid e) < Pr(h)$ on the basis of $Pr(c \mid e) \leq Pr(c)$ for some c that is a logical consequence of h.

It is clearly true that one cannot conclude that the probability of h goes down because the probability of $e \rightarrow h$ goes down. However, that particular logical consequence might have a unique form that makes it behave in a peculiar manner. One might wonder if merely excluding certain kinds of logical consequences will allow one to effectively conclude from the non-raised probability of all *other* kinds of logical consequences that the probability of the theory itself is also lowered. Williamson has merely shown that for certain logical consequences of hypotheses the inference is inappropriate, but that does not prove that for all or most logical consequences such an inference is inappropriate.

Williamson says that it is a fallacy to criticize a theory by identifying "a logical consequence" of that theory whose probability is not raised by the evidence. This statement is, again, ambiguous. On one disambiguation, what Williamson says is a fallacy is to criticize a theory if the probability of *any* logical consequence is not raised by the evidence. On a second disambiguation, what he says is a fallacy is to criticize a theory if the probability of *this* logical consequence (where perhaps this consequence is of a certain kind) is not raised by the evidence. Given his argument, it is clear that Williamson intends the first interpretation. But one wonders if the second interpretation might prove to be a valid principle of reasoning rather than a fallacy, if only the right types of consequences can be identified. For example, consider a consequence of a theory h that is logically *equivalent* to the theory, such as (h & e) $\vee$ h. If e raises the probability of h, it raises the probability of (h & e) $\vee$ h. Hence if the evidence does not raise the probability of (h & e) $\vee$ h, it does not raise the probability of h. How common are logical consequences like these?

As it turns out, these kinds of logical consequences are fairly special. It can be shown that for any logical consequence of a theory such that the consequence is of strictly *greater* probability (which consequences logically equivalent to the theory cannot be), there is evidence that will not raise the probability of the consequence while raising the probability of the theory. In such a case, Pr(c) > Pr(h), and Pr(c & h) = Pr(h) (because c & h is logically equivalent to h, on the assumption that c is a logical consequence of h). Let the evidence be c $\rightarrow$ h. It can be shown that, first, Pr(h | c $\rightarrow$ h) > Pr(h), and second, Pr(c | c $\rightarrow$ h) < Pr(c).

Here is the proof of the first inequality: Pr(h | c $\rightarrow$ h) = Pr(h & (c $\rightarrow$ h))/Pr(c $\rightarrow$ h) = Pr(h)/Pr(c $\rightarrow$ h). Pr(c $\rightarrow$ h) = 1 – Pr(c & ~h). Because c is a logical consequence of h but not equivalent to it, Pr(c & ~h) = Pr(c) – Pr(h) > 0, so Pr(c $\rightarrow$ h) = 1 – (Pr(c) – Pr(h)) < 1. Pr(h | c $\rightarrow$ h) = Pr(h)/Pr(c $\rightarrow$ h) = Pr(h)/(1– (Pr(c) – Pr(h))), and hence Pr(h | c $\rightarrow$ h) > Pr(h).

Here is the proof of the second inequality: Pr(c | c $\rightarrow$ h) = Pr(c & (c $\rightarrow$ h))/Pr(c $\rightarrow$ h) = Pr(c & h)/Pr(c $\rightarrow$ h) = Pr(h)/Pr(c $\rightarrow$ h) (because c is a logical consequence of h, Pr(c & h) = Pr(h)). Pr(h)/Pr(c $\rightarrow$ h) = Pr(h)/(1-Pr(c & ~h)) = Pr(h)/(1 – (Pr(c) – Pr(h)) = Pr(c)/(1 + (1 – Pr(c))(Pr(c)/Pr(h) – 1)). Because 1 > Pr(c) > 0, 1 – Pr(c) > 0. Because Pr(c) > Pr(h), Pr(c)/Pr(h) – 1 > 0. So 1 + (1 – Pr(c))(Pr(c)/Pr(h) – 1) > 1 and hence Pr(c)/(1 + (1 – Pr(c))(Pr(c)/Pr(h) – 1)) < Pr(c).

So, for any logical consequence of a theory that is of higher probability than the theory, no matter the logical form of the two statements, there is evidence that raises the probability of the theory but lowers the probability of the consequence. Again, however, it seems that we have relied on a special logical form of a statement; in this case, it is the evidence that takes a special form. Can we rehabilitate the consequence fallacy by restricting the form of the evidence?

We don't think the consequence fallacy — understood as (T1) — can be saved from fallaciousness in this way. Consider that h and c & ~h are contraries, and hence cannot simultaneously be true, and consider that, since c is a logical consequence of h, c is logically equivalent to $h \vee (c \,\&\, {\sim}h)$. It follows that $\Pr(c) = \Pr(h) + \Pr(c \,\&\, {\sim}h)$, and $\Pr(c \mid e) = \Pr(h \mid e) + \Pr(c \,\&\, {\sim}h \mid e)$. Hence, even if $\Pr(h \mid e) > \Pr(h)$, as long as $\Pr(h \mid e) - \Pr(h) \leq \Pr(c \,\&\, {\sim}h) - \Pr(c \,\&\, {\sim}h \mid e)$, then $\Pr(c \mid e) \leq \Pr(c)$. Note that since h and c & ~h are contraries, there is in general no barrier to e raising the probability of h while lowering the probability of c & ~h.

So (T1) is a fallacy. But (T2) is not. It is not a fallacy to argue that a theory h is not made *likely to be true* (a theory is likely to be true if its probability is higher than 0.5) by evidence e on the basis of the fact that c, a logical consequence of h, is not made likely to be true by e.

To see why (T2) is not a fallacy, we have to notice, first of all, that $\Pr(c \mid e) \geq \Pr(h \mid e)$. Since c is a logical consequence of h, there cannot be a world in which h is true but c is false — for every world in which h is true, there is a world in which c is true (the very same world). Hence, if one thinks of the probability of a proposition as the ratio of the measure of the worlds in which the proposition is true as compared with the measure of the total worlds, it is clear that $\Pr(c) \geq \Pr(h)$. For the exact same reason, $\Pr(c \mid e) \geq \Pr(h \mid e)$. Since $\Pr(c \mid e) \geq \Pr(h \mid e)$, if $\Pr(c \mid e) \leq 0.5$, then $\Pr(h \mid e) \leq 0.5$. (For those who think a formal proof is needed, here is the proof of $\Pr(c) \geq \Pr(h)$: $\Pr(c) = \Pr(c \,\&\, h) + \Pr(c \,\&\, {\sim}h) = \Pr(h) + \Pr(c \,\&\, {\sim}h)$ (because c is a logical consequence of h, c & h is logically equivalent to h). $\Pr(h) + \Pr(c \,\&\, {\sim}h) \geq \Pr(h)$, so $\Pr(c) \geq$

Pr(h). And the proof of Pr(c | e) ≥ Pr(h | e): Pr(c | e) = Pr(c & e)/Pr(e) = (Pr(c & e & h) + Pr(c & e & ~h))/Pr(e) = (Pr(e & h) + Pr(c & e & ~h))/Pr(e) = Pr(h & e)/Pr(e) + Pr(c & ~h & e)/Pr(e) = Pr(h | e) + Pr(c & ~h | e) ≥ Pr(h | e).)

According to (T2), we can infer from the fact that c is not made likely to be true by e (i.e. Pr(c | e) ≤ 0.5) to the conclusion that h is not made likely to be true by e (i.e. Pr(h | e) ≤ 0.5). There are two ways in which c is not made likely to be true by e. In the first way, both Pr(c) ≤ 0.5 and Pr(c | e) ≤ 0.5. It does not matter whether Pr(c | e) is greater or smaller than Pr(c), that is, it does not matter whether e raises the probability of c or not, for given that Pr(c | e) ≥ Pr(h | e) and that Pr(c | e) ≤ 0.5, we can conclude that Pr(h | e) ≤ 0.5 even if Pr(c | e) > Pr(c). In the second way, Pr(c) > 0.5 and Pr(c | e) ≤ 0.5, that is, e lowers the probability of c to such an extent that c is not likely to be true given e. Again, we can conclude that Pr(h | e) ≤ 0.5. (T2) is thus applicable not only to cases in which the evidence does not raise the probability of the consequence above 0.5, but also to cases in which the evidence lowers the probability of the consequence to 0.5 or below.

We have claimed that (T2) is a more commonly used tactic than (T1). Although it is difficult, if possible at all, to demonstrate the truth of the claim, we can explain why it is reasonable to think that the claim is true. It is tempting to offer the following simple explanation: it is reasonable to think that (T2) is a more commonly used tactic than (T1), at least among philosophers, because (T1) is a fallacy and (T2) is not. This may not, however, be a good explanation if (T1) is a fallacy that is easy to commit even for philosophers. There is, in any case, a better explanation. It is reasonable to think that (T2) is a more commonly used tactic than (T1) because the tactic in question is one of criticizing *confidence* in a theory. Our claim is, more precisely, that (T2) is more commonly used than (T1) as a tactic of criticizing confidence in a theory. It is not difficult to see why it is reasonable to think that this claim is true once we see how the expression ‘confidence in a theory’ should be understood.

The most reasonable understanding of ‘confidence in a theory’ is that one has confidence in a theory only if one believes that the theory is likely to be true. It does not seem to make sense for anyone to say ‘I have confidence in the theory, but I don’t believe that the theory is likely to be true’. If one understands ‘confidence in a theory’ this way, then one would use (T2) rather than (T1) — even if one does not know that (T1) is fallacious — to criticize confidence in a theory because it is (T2) rather than (T1) that is *clearly* about whether a theory is made likely to be true by the evidence.

Is this how Williamson himself understands ‘confidence in a theory’ as he uses the expression in his argument? That the answer should be ‘Yes’ can be seen from his remark that “all this is compatible with *a high degree* of *legitimate* confidence in both h and e → h” (232; italics added). It obviously does not make sense for anyone to say ‘I have a high degree of legitimate confidence in the theory, but I don’t believe that the theory is likely to be true’, for if the person does not believe that the theory is likely to be true, her degree of confidence should not be high, nor should she see her confidence as legitimate.

If this is how ‘confidence in a theory’ is understood, an effective way of criticizing a person’s confidence in a theory would be to criticize her belief that the theory is likely to be true — if that belief is false or unjustified, she should not have confidence in the theory. This is precisely what we do when we criticize a person’s confidence in a theory by identifying a logical consequence (of the theory) whose probability is not raised, or is lowered, by the evidence such that it is not likely to be true. The criticism is not that the probability of the theory is thereby not raised, or is lowered, by the evidence (which, as Williamson rightly argues, is a fallacy), but that the probability of the theory cannot be so high that it is likely to be true (even when the probability is raised). This is the tactic of (T2).

By contrast, we don't see how (T1) could ever be used to criticize confidence in a theory. If all we point out is that the theory has a logical consequence whose probability is not raised by the evidence, this point is compatible with the possibility that the probability of the logical consequence is still greater than 0.5 given the evidence, and hence compatible with the possibility that the probability of the theory is greater than 0.5 given the evidence (even if the probability of the theory is lowered by the evidence). Because of this compatibility, (T1) could not be used to criticize confidence in a theory even if it were *not* a fallacy to infer from the fact that the probability of a logical consequence of the theory is not raised by the evidence to the conclusion that the probability of the theory is not raised. It is not the mere raising or lowering of the probability of the consequence of a theory that is relevant to our confidence in the theory, but the *value* of the probability raised or lowered.

Of course, a person can believe or accept a theory without having confidence in it in the above sense. But such a person is not likely to be our opponent in a debate in which we argue that the theory should not be accepted. This may be another explanation of why it is reasonable to think that (T2) is a more commonly used tactic than (T1).

To further support our claim that (T2) is a more commonly used tactic than (T1), let us consider a version of the argument for epistemological skepticism formulated by Anthony Brueckner (1994: 831-832), which is effectively of form (T2), not (T1). Let *P* be some arbitrary proposition about the external world, such as 'I have hands', and let *SK* be some radical skeptical hypothesis that is incompatible with P, such as 'I am a brain in a vat':

(1) If I have justification for believing that P, then I have justification for believing that ~SK.

(2) If my evidence for believing that ~SK does not favor ~SK over SK, then I lack justification for believing that ~SK.

(3) My evidence for believing that ~SK (my sensory evidence) does not favor ~SK over SK.

(4) I lack justification for believing that ~SK.

(5) I lack justification for believing that P.

(6) I do not know that P.

According to Brueckner, the above argument employs two epistemic principles:

(EP1) For all S, ϕ, ψ, if S's evidence for believing that ϕ does not favor ϕ over some incompatible hypothesis ψ, then S lacks justification for believing that ϕ.

(EP2) For all S, ϕ, ψ, if S has justification for believing that ϕ, and ($\phi \rightarrow \psi$), then S has justification for believing that ψ.

On the understanding that justification is conducive to truth,[2] it is reasonable to understand both (EP1) and (EP2) in terms of probability: 'does not favor ϕ over ψ' can be understood as 'does not make ϕ rather than ψ likely to be true' and 'has justification for believing that ϕ' can be understood as 'has evidence that makes ϕ likely to be true'.[3] The employment of (EP1) and (EP2) in the above argument can then be understood to be the same as (T2): arguing that P is not made likely to be true by the evidence by identifying a logical consequence of P, namely, ~SK, that is not made likely to be true by the evidence.[4]

To offer a contrast with the above skeptical argument, which uses (T2), let us consider the following argument to which, Williamson suggests, the skeptic may apply (T1):

(1**) It appears to me that I have hands.

(2**) If it appears to me that I have hands, then I have hands.

(3**) I have hands.

(2**) is supposed to be a material conditional. In that case, (2**) is a logical consequence of (3**). Williamson suggests that “[a] skeptic about the external world may hold that our evidence raises the probability of (1**) but not of (2**)”, and then points out that it is a fallacy “to argue on that basis that, given our evidence, we are not entitled to high degrees of confidence in (2**) and (3**)” (Williamson 2007: 234).

If all the skeptic holds is that the evidence raises the probability of (1**) but not of (2**), it is clear that he cannot conclude on that basis that the probability of (2**) is not high enough for (2**) to be likely to be true. Since he cannot draw that conclusion, he cannot conclude that the probability of (3**) is not high enough for (3**) to be likely to be true *even if the probability of (3**) went down with that of (2**)*. In other words, even if (T1) were not a fallacy, it would not allow the skeptic to conclude that we should not have confidence in (3**).

However, there is no reason to think that the skeptic would use (T1). Suppose the skeptic holds that the evidence does not raise the probability of (2**) above 0.5. Now even if he takes into consideration the fact that (T1) is a fallacy and acknowledges the possibility that the probability of (3**) goes up when that of (2**) goes down, he still has to conclude that the probability of (3**) must be lower than 0.5 (because $Pr((2^{**})| e) \geq Pr((3^{**}) | e)$), and hence that we should not have confidence in (3**). To argue this way is to apply (T2), which is a valid tactic of criticizing confidence in a theory.

Although Williamson’s discussion of the consequence fallacy is brief, it would be a significant contribution to philosophical methodology if he had identified a fallacious tactic of criticizing confidence in a theory that is commonly used by philosophers. If Williamson’s target is (T1), his argument is not very significant because (T1) does not seem to be a commonly used tactic by philosophers. If his target is (T2), then his argument is unsound. Now even if (T2) is not commonly used, it is important to distinguish it from (T1), for it is, unlike (T1), not a fallacious tactic. Indeed, since this paper has made clear that (T2) is a sound tactic, it may serve the purpose of encouraging the use of it.

References

BonJour, L. (1985). *The Structure of Empirical Knowledge*. Cambridge, Mass.: Harvard University Press.

Brueckner, A. (1994). 'The Structure of the Skeptical Argument', *Philosophy and Phenomenological Research*, 54: 827-835.

Carnap, R. (1962). *Logical Foundations of Probability*, 2nd edition. Chicago: The University of Chicago Press.

Huemer, M. (2001). *Skepticism and the Veil of Perception*. Lanham, Md.: Roman & Littlefield.

Williamson, T. (2007). *The Philosophy of Philosophy*. Oxford: Wiley-Blackwell.

NOTES

[1] The distinction between the probability of a theory's being raised by the evidence and a theory's being made likely to be true by the evidence is basically the same as the distinction between 'confirmation as increase in firmness' and 'confirmation as firmness' that Carnap makes in the preface to the second edition of *Logical Foundations of Probability* (see Carnap 1962: xv-xiv).

[2] As Laurence BonJour points out, "if finding epistemically justified beliefs did not substantially increase the likelihood of finding true ones, then epistemic justification would be irrelevant to our main cognitive goal and of dubious worth" (BonJour 1985: 8).

[3] Cf. "The radical skeptic I will be confronting holds that (some significant class of) common sense beliefs are *not at all* justified, which is to say: there is no reason to believe that they are true; it is no more rational to think they are true than to think they are false. According to one theory of probability, this is the same as to say that the beliefs are no more *likely* to be true than false." (Huemer 2001: 20; original italics)

[4] Many epistemologists understand the skeptical argument to employ both (EP1) and (EP2), though their versions of the skeptical argument may not be exactly the same as the version we consider here. Brueckner's final reconstruction of the skeptical argument, however, employs only (EP1) (see Brueckner 1994: 833).