

# Comments on Tim Willenken’s “Against Epistemic Conservatism”

Eleonora Cresto

CONICET (Argentina)

[eleonora.cresto@gmail.com](mailto:eleonora.cresto@gmail.com)

## 1. Introduction

Tim seeks to defend the view, which he calls *liberalism*, according to which sometimes  $S$  can acquire additional rational support to believe that  $p$  on the basis of  $E$  even if  $S$  has no  $p$ -independent justification to believe that each of  $p$ 's underminers are false. To this effect he discusses a well-known argumentative pattern, together with recent attempts to show that positing the transmission of justification (or warrant) along this pattern sometimes conflicts with basic Bayesian claims – provided we agree to build some bridges between propositional justification and probabilities. In what follows I will dub the aforementioned pattern “the I-II-III pattern” (following Crispin Wright’s terminology back in the eighties; cf. Wright (1985)) and I will refer to the attempt to show an incompatibility between Bayesianism and the assurance of warrant transmission as “the Cohen-White argument”, or, for short, “the C-W argument” (cf. Cohen (2005), or White (2006)). Tim identifies three crucial assumptions for the C-W argument to run, and rejects one of them (the “Auxiliary Thesis about Independent justification”). The rationale for the rejection – we are told – can actually be spelled out as a direct argument against *conservatism* (the negation of liberalism).

I would like to say at the outset that Tim and I are pretty much on the same page on most issues discussed in the paper; if anything, my suggestions here will point to the possibility to deepen the path he is already embarked on. I will divide this commentary in three further sections. In section 2 I will raise a different concern on the C-W argument. I’m very sympathetic to the claim that “both sides have got it wrong by overstating what epistemic conclusions we can read off of the Bayesian formalism” (Willenken (2009), p. 11) – but I tend to be even more radical. Tim seems to think that,

if the auxiliary assumptions are accepted, the C-W argument should be taken to be compelling. I want to contend that this is not necessarily the case.

In section 3 I will voice some worries on the intended scope of the paper. The paper advertises itself as an attempt to defend epistemic liberalism (in the sense mentioned above), very broadly conceived; in particular, Tim seems eager to remain neutral towards Jim Pryor's specific variety of liberalism, which is usually referred to as a particular type of *dogmatism* (cf. Pryor (2000)). However, we might suspect that the chosen strategy forces the author to commit himself to a stronger position, as not all possible versions of liberalism require that we take pains to deactivate the C-W argument first.

Finally, in section 4 I will examine the contention that the Auxiliary Thesis about Independent Justification is illegitimate, as well as his direct argument against conservatism. I will suggest that his position would benefit from a more robust account on how probabilities and epistemic justification relate to each other. The worry is that, lacking such an account, the alleged examples against conservatism might be found to be question begging.

## 2. A violation of Bayesian theorems?

Let me begin by recalling a few general (and rather trivial) facts about probabilities and Bayesian confirmation. Suppose that  $H$  entails  $H'$ . Then  $P(H) \leq P(H')$ , and hence  $P(H/E) \leq P(H'/E)$ . But it could still be the case that  $E$  confirms  $H$  but not  $H'$ , as Roger White reminds us in his (2006). In other words, we could well have:

$$1) P(H') > P(H'/E) \geq P(H/E) > P(H)$$

Supporters of the C-W argument claim that equation 1) holds for typical examples of the I-II-III pattern, where line I is  $E$ , line II is  $H$ , and line III is  $H'$  (notice that (1) can be maintained even if  $H'$  is not strictly *deducted* from  $H$ ). Equation 1) yields  $P(E/-H') > P(E/H)$ ,<sup>1</sup> and the underminer of line II (let's call it  $U$ ) is usually taken to be  $-H'$ . Moreover, line II is usually taken to imply line III. Thus, supporters of the C-W argument typically assume that

$$2) P(E/U) = P(E/-H') = P(E/H) = 1 > P(E/H')$$

---

<sup>1</sup>  $P(H') > P(H'/E) \Rightarrow P(-H'/E) > P(-H') \Rightarrow P(-H'/E)/P(-H') > P(H'/E)/P(H') \Rightarrow P(E/-H') > P(E/H)$ .

I tend to think that one or more terms in equation 2) should be contested.

Just to be clear, let me recall here that the C-W argument does not say that Bayesianism will force us to conclude that transmission of warrant fails *in all cases*, but only that it will fail for paradigmatic examples of the I-II-III pattern. I think we can argue that the situation is just the opposite: even though equation 2) holds in some cases, paradigmatic examples of the pattern do not satisfy it, and hence they do not license the conclusion that transmission of warrant is bound to fail – not even assuming the legitimacy of Auxiliary Theses of the sort identified by the paper. Let us see this with more care.

Consider the following statements:

**E** (I am having perceptual experience as of hands)

**Q** (I have hands)

**PERC** (I have **Q** on the basis of my perceptual experience as of hands)

**U** (I am a deceived brain in a vat – without hands – having perceptual experience as of hands)

Using such statements, Tim builds an instance of the I-II-III pattern (he actually borrows the example from Pryor (2007)):

I) **E**

II) **PERC**

III)  $\neg$ **U**

Presumably, line III amounts to something like “I am not a deceived brain in a vat – without actual hands – having perceptual experience as of hands”, but, as we will see in a moment, it is not so clear how to understand the negation in this claim.

Following the lessons of the C-W argument, we are told that

$$3) P(\text{PERC}/E) < P(\neg U)$$

(Tim writes “less or equal than”, but we can strengthen it to a strict inequality). The C-W argument takes this result to mean that “if *PERC* is going to be justified after having experience *E*,  $\neg U$  had better be independently justified” (Willenken (2009), p. 12-13). Now, a first observation I would like to make here is that we can arrive at a symmetric result about  $\neg$ *PERC*. Given that *PERC* entails *E*,  $P(\text{PERC}/E) > P(\text{PERC})$ , so  $P(\neg \text{PERC}) > P(\neg \text{PERC}/E)$  (which intuitively says that *E* lowers our confidence in the fact that we have acquired *Q* for the wrong reasons, so to speak – e.g., through “envattement”). As *U* entails  $\neg$ *PERC*,  $P(U/E) \leq P(\neg \text{PERC}/E)$ , hence  $P(\neg \text{PERC}/E) \geq P(U/E)$ . Hence:

$$4) P(U/E) < P(\neg\text{PERC})$$

In other words, assuming Bayesian conditionalization holds, our confidence in the underminer after learning that  $E$  is lower than the prior confidence we had in the falsity of the second premise. It is important to bear this result in mind, because sometimes the discussion in the literature proceeds as if learning that  $E$  were intrinsically bad for anti-skeptical purposes – but this is clearly not the case. Putting 3) and 4) together, we have

$$5) P(\neg U) > P(\neg U/E) > P(\text{PERC}/E) > P(\text{PERC})$$

Notice, however, that equation 5) is obtained just because  $\neg U$  is bound to be true when  $E$  is false. In other words,  $E$  lowers  $\neg U$  because, the way it has been defined,  $U$  implies  $E$ , so  $P(E/U) = P(\neg U/\neg E) = 1$ . But this way of building the case is contentious.

To see why, let me step back for a moment and consider other arguments that are also said to follow the I-II-III pattern. Take Roger White's example in (2006):

WI: It appears to me that this is a hand.

WII: This is a hand.

WIII: This is not a fake-hand.

Soon we learn that “This is not a fake hand”, according to White's analysis, should be understood as “either this is a real hand *or I do not have hand-like experiences at all*”. But this is not an interesting statement we would like to plug into the I-II-III pattern, so the intended argument is to some extent circumvented. Intuitively, “This is not a fake hand” implies that I *do* have the experience as of a hand, *and* that it is a real hand. We also have the intuition that “this is a fake hand” equally implies that I have the experience as of a hand, though this time it is not a *real* hand. But then WIII and the underminer are not strictly speaking contradictory statements. Here are some alternatives:

WIII.a) What looks like a hand is actually a real hand (as opposed to a fake hand)

WIII.b) What looks like a hand is not a real hand, but a fake hand;

in which case both  $P(\text{WI}/\text{WIIIa}) = P(\text{WI}/\text{WIIIb}) = 1$ . Then we face a dilemma: either we respect this intuition, but then the underminer does not strictly speaking contradict the intended conclusion of the argument, or the underminer is the negation of line III, but line III is not exactly what we are interested in, so the morals we draw from here are quite limited.

Let's go back to Tim's paper. The problem here is not as transparent as in White's case ( $\neg U$  is not explicitly of the form  $(PERC \vee \neg E)$ ), but it is still there. Why is it assumed that  $\neg U$  does not imply  $E$ ? Presumably, you could be a real human being (as opposed to a brain in a vat) and still not have the particular experience captured by statement  $E$ . So far so good. But then, *the same rationale should apply if you are envatted*. So, again, either line III and the underminer are both taken to imply  $E$ , and hence the two statements are not strictly construed as contradicting each other, or *none of them* is taken to strictly imply line 1. Otherwise we are introducing an asymmetry that is not well motivated. The whole argument against dogmatism seems to rely on this illegitimate asymmetry.

To put it more explicitly, consider the following alternative ways to disambiguate the relevant statements:

(U.a): I am having perceptual experience as of hands, and by having this experience I am being deceived by a mad scientist;

(U.b): I am having perceptual experience as of hands, and by having this experience I am not being deceived by a mad scientist;

as opposed to:

(U.a') = (U.a): I am having perceptual experience as of hands, and by having this experience I am being deceived by a mad scientist;

(U.b'): Either I am not having a perceptual experience as of hands, or I am not being deceived by a mad scientist.

Clearly, (U.b') is close to what the paper takes to be line III, but is not intuitively what we want to assert as receiving support from premise II. Consider now the corresponding version of equation 5), if (U.a) and (U.b) are taken to be the right pair:

$$5') P(U.b) > P(U.b/E) > P(PERC/E) > P(PERC)$$

As opposed to 5), equation 5') cannot be asserted:  $P(U.b/E)$ , and not  $P(U.b)$ , will now be the largest of the four quantities, while  $P(PERC)$  will still be the smallest.

Would we then be any closer to refuting the skeptic? Of course not: the formalism does not tell us whether  $P(U.b/E)$  is higher or lower than  $P(U.a/E)$ . This is, I think, as it should be, if we want to model the idea that skeptical scenarios are indistinguishable from non-skeptical ones. Still, it is good to know that gathering new evidence cannot make us worse off.

### 3. Which liberalism?

The observations from the previous section should not be taken to constitute an argument *in favor of dogmatism*, in Pryor's sense. Neither should they be taken to mean that Bayesianism actually *supports* Pryor's model – they only tell us that certain ways of attacking dogmatism are better to be avoided. In any case, a full defense of Pryor's dogmatism has never been Tim's explicit intention. This is, I think, the correct attitude, given that Pryor's general standpoint does not seem to be compatible with Bayesianism, for reasons that are completely independent from the Cohen-White argument. The source of the problem is *not* the supposed violation of certain Bayesian theorems,<sup>2</sup> but a deeper reluctance to assimilate “dogmatic evidential support” to other kinds of evidential support; “dogmatic evidential support” is *sui generis*, as Pryor himself acknowledges (cf. his 2007).

Still, I am not completely sure whether Tim is not implicitly endorsing something very close to Pryor's framework. Let me identify two different possible goals one might have in mind at the time of reflecting on the transmission of warrant along the I-II-III pattern. We might want to:

- a. argue that sometimes the lack of prior justification for line III (or, perhaps, the lack of prior justification to reject the underminers for line II), does not prevent us from having transmission of warrant from I to II.
- b. argue that, as sometimes we do not need prior justification for line III in order to have transmission of warrant from I to II, we can further acquire *warrant for III* through II.

The second goal is obviously more ambitious; strictly speaking, in order to argue in favor of liberalism, goal (a) suffices. To see this, let me recall that one can be liberal and still accept that transmission of justification from II to III may well fail. In other words, there are models that save liberalism, and for which the putative problem that constitutes the backbone of this paper (i.e., the putative violation of Bayesian considerations) does not even arise. By way of illustration, take Akeel Bilgrami's understanding of the I-II-III pattern (cf. for instance Bilgrami (2004)). As long as we do

---

<sup>2</sup> By “Bayesian theorems” I do not mean here “*probability* theorems”; the probability calculus, by itself, is not concerned with diachronic considerations.

not have reasons to doubt line III – Bilgrami contends – we are entitled to carry out the argument from I to II; in particular, we do not need to have positive reasons to believe that the relevant underminers are false (other than the fact that their negation is entailed by our prior endorsement of line III). Unlike Pryor, however, Bilgrami takes the explanation for this phenomenon to be that type-III propositions are typically part of our background (we believe them already, before facing a particular argument), and propositions we already believe in *are not the type of thing that can be justified or unjustified*. The crucial idea here is that the concept of justification only applies to epistemic changes, and not to antecedent elements. Thus, in most characteristic instances of the I-II-III pattern no premise can give us additional reasons to assert III, so there is no transmission of justification from II to III. Still, it would not be correct to say that III is “justified by default,” as other authors would have it; justification questions only arise when we want to modify our background. So, for example, we *would* need justification to assert that line III is false (technically speaking, that would amount to engaging in *belief revision*), but not to keep on maintaining it.

Were we to adopt this or similar lines of reasoning, we would fulfill goal (a) and offer a defense of liberalism that does not require examining the C-W strategy at all. This seems to suggest that the paper is implicitly committed to (b) rather than (a), and that Tim ultimately seeks to show that justification transmits to III, as Pryor would have it.

#### **4. Against Conservatism**

In the light of the previous section, let me assume that the goal of the paper is actually (b). Let’s also grant, for the sake of the argument, that, in spite of the observations from section 2, there are philosophically interesting instances of the I-II-III pattern to which the C-W argument applies. Then it is right to point out that, in order for probabilistic considerations to bear any relevance on issues of propositional justification, we need some auxiliary principles to hold. Let me focus on the rejection of the Auxiliary Thesis about Independent Justification (from now on, ATIJ), as well as on the related direct attack on conservatism.

The thought is that we could well have  $P(X/E) \leq P(Y)$ , “where E is part of the evidential base of X, yet where one does not need independent justification to believe Y in order to have justification to believe that X” (p. 16). I am absolutely sympathetic to this idea, but I would root it in a more general theory. In general, I am convinced that the probability of  $p$  (for – almost – any  $p$ ) is not *all* we have to consider in order to come to believe that  $p$ . Depending on the details, it may be perfectly acceptable to have  $P(p) < P(q)$  and yet rationally come to believe  $p$  but not  $q$  – say, because  $p$  has higher explanatory value, or because it better cohere with the rest of our beliefs, among many other possibilities.

Tim does not go through this path. He does not point to a comprehensive story that will allow us to reject ATIJ in a principled way. Rather, he offers some examples with the hope of showing that ATIJ leads to intuitively unpalatable results. Analogously, he will later argue against conservatism by offering examples in which common sense intuitions and conservatism yield different verdicts. I have some qualms about this line of reasoning, because it is not clear to me whether putative common sense intuitions, by themselves, can override the diagnosis that would follow from accepting conservatism.

Consider the example that attempts to deactivate ATIJ (“Etch-A-Sketch with Conditional Pills”). He first seeks to show that

$$6) P(p/E) \leq P(E \rightarrow p)$$

where ‘E’ is the evidence and ‘p’ the proposition that the picture has disappeared. Notice that if  $E \rightarrow p$  is taken to be a material conditional, equation 6) is trivially true.<sup>3</sup> On the other hand, if  $E \rightarrow p$  is not a material conditional, it is not so clear what to say; it would be desirable to find some clarifications on how the arrow is to be understood. (For example, it’s not the case, for every possible type of conditionals, that the truth of the consequent always entails the conditional as a whole, as required by the proof that appears on p. 17). Incidentally, I am not completely sure of which concept of probability is at stake in the paper. Tim mentions “epistemic” probabilities all along

---

<sup>3</sup> Proof (where  $P(E) > 0$ ):

- 1)  $P(p/E) \leq P(\neg E \vee p)$
- 2)  $P(p \& E) / P(E) \leq P(\neg E) + P(p) - P(p \& \neg E)$
- 3)  $P(p \& E) + P(p \& \neg E) / P(E) \leq P(\neg E) + P(p)$
- 4)  $P(p) / P(E) \leq P(\neg E) + P(p)$
- 5)  $0 \leq P(\neg E)$ .

Incidentally, if the arrow here is the material conditional, line 1 on p.17 just *is* line 5, as  $P(\neg E \vee p/E) = P(p/E)$ .



section IV, which seems to point to something akin to Williamson's (2000) sense, but he also speaks in terms of agents having "rational confidence" in a proposition, in which case a personalist interpretation seems to be called for. It should also be noted that equation 6), as it stands, does not encode a diachronic statement. In general,  $P(p/E)$  does not necessarily state our confidence in  $p$  "upon learning that  $E$ ". We might well believe that strict Bayesian conditionalization should hold in the present context, but this is actually a *further* assumption we have to add to those already on the table.

In any case, we don't need to rely on conditionals of any sort for the structure of the example to work – and in the end, this might be problematic: the structure of the case enables us to prove too much. Take *any* logically independent  $Y$ ,  $X$  and  $E$  with priors in  $(0,1)$ , for which the agent antecedently has  $P(X/E) < P(Y)$ , and such that the agent is intuitively justified in believing that  $X$  once she learns that  $E$  is true. Then plug a "pill situation" in: assume that upon taking a pill, the agent won't be justified in believing that  $Y$ . Is the agent still justified to come to believe that  $X$ ?

More generally, the structure of the example seems to be like this:

- Intuitively, it seems that you have  $J$  (where  $J$  is any property you find desirable).
- But your theory says that in order to have  $J$  you need  $F$ .
- Something prevents you from having  $F$ .
- Do you actually have  $J$  under the circumstances?

All we can say is that in such a case there is a clash between what our theory predicts and our pre-theoretical intuitions, but it's not obvious that the right moral here is that the *theory* should go. The problem is compounded by the fact that what prevents us from having  $F$  in Tim's example is a science fiction trick: science fiction scenarios are seldom helpful to elicit clear intuitions. The theory could well yield that we take the scenario to be, say, metaphysically impossible. In short, I'm not sure how to arrive at the desired verdict without begging the question against conservatism; it seems to me that in order for the example to be successful we need to be already convinced that ATIJ can be violated.

Very similar considerations apply to the direct strategy against conservatism in section V. The conservative can very well bite the bullet and claim that the cases discussed are not counterexamples at all – or perhaps that they are not *relevant* enough. Ideally, for a putative counterexample to be significant we should show how to incorporate it to a comprehensive rival picture; alternatively, we could try to show how to use it as input to modify the original position somehow, in a sort of reflective

equilibrium mechanism. Just advancing the intuition that probabilities do not translate un-problematically into justification for coming to believe might be short of what we need; an independent story that tells us how to link probabilities and belief seems to be desirable.

## 5. Conclusions

To sum up:

- (1) The particular instance of the I-II-III pattern offered in the paper shows that there is no transmission of Bayesian confirmation from line II to line III (which might – but need not – be taken to indicate that there is no transmission of warrant from II to III), but this is due to an artifact of the example. Once we correct this, the C-W argument is rendered ineffective against most paradigmatic examples of the pattern.
- (2) There are interpretations of the I-II-III pattern that save liberalism, and for which the putative problem highlighted by the C-W argument does not even arise. In other words, we can be liberal and yet accept that transmission of warrant from II to III fails.
- (3) As for the specific proposal of divorcing probability from justification, it would be desirable to find a more robust general account of how probabilities and acceptance (or belief) are linked; without such a general account, we might worry that the morals drawn from the examples in sections IV and V beg the question against conservatism.

## References

- Bilgrami, A. (2004). "Skepticism and pragmatism." In D. McManus (Ed.), *Wittgenstein and skepticism*. London: Routledge, pp. 56–75.
- Cohen, S. (2005). "Why Basic Knowledge is Easy Knowledge". *Philosophical and Phenomenological Research* 70: 417-430.
- Pryor (2000). "The Skeptic and the Dogmatist". *Nous* 34: 517-549.
- Pryor (2007). "Uncertainty and Undermining". Manuscript.
- White, R. (2006). "Problems for Dogmatism". *Philosophical Studies* 131: 525-557.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.
- Willenken, T. (2009). "Against Epistemic Conservatism". Manuscript.
- Wright, C. (1985). "Facts and Certainty." Henriette Hertz Philosophical Lecture. *Proceedings of the British Academy*. Oxford: The British Academy, pp. 430-472.