Strengthening the case for knowledge from falsehood

Branden Fitelson

1. Background: one of Warfield's examples of knowledge from falsehood

Recently, several authors have offered examples of inferential knowledge, which is (at least prima facie) based on a falsehood. In this note, I will focus my attention on the following example, which is presented and discussed by Warfield (2005: 408).

> I have a 7pm meeting and extreme confidence in the accuracy of my fancy watch. Having lost track of the time and wanting to arrive on time for the meeting, I look carefully at my watch. I reason: 'It is exactly 2:58pm; therefore I am not late for my 7pm meeting'. Again I know my conclusion, but as it happens it's exactly 2:56pm, not 2:58pm.

For Warfield, an example counts as a case of *knowledge from falsehood* (KFF) just in case both of the following conditions are satisfied by the example (2005: 408).

> (1) the example involves inferential knowledge of a conclusion q.

> (2) the example involves a false relevant premise p.

Warfield considers various 'resistance' strategies, which attempt to undermine the claim that examples such as these are *bona fide* KFF-cases. He argues that none of these 'resistance' strategies will succeed. I will not enter into that debate here.[1] Rather, for the purposes of this note, I will just *assume* that Warfield's assessment is correct. My aim here is only to argue that *if* such examples *are bona fide* cases of KFF, *then* they can be *strengthened* in some epistemologically interesting ways.

2. A stronger conception of KFF, and an example thereof

Let us *suppose* that Warfield is correct — that his example (above) *does* satisfy both (1) and (2). I am interested in *stronger* conceptions of KFF. To wit, consider:

> (3) If the subject's belief p had *not* been *false*, then the example would *not* have constituted a case of inferential knowledge.

To require (3) is to require that the example be a case in which the inferential knowledge (that q, on the basis of p) *is counterfactually dependent on the falsity of the relevant premise p*. Adding (3) to Warfield's criteria (1) and (2) yields a *stronger* conception of KFF, which I will call KFF*. By modifying Warfield's example [which, intuitively, does *not* satisfy (3)], we can construct examples of KFF*. Here's one:

I have a 7pm meeting and extreme confidence in the accuracy of both my fancy watch and the Campanile clock. Having lost track of the time and wanting to arrive on time for the meeting, I look out of my office window (from which the Campanile clock is almost always visible). As luck would have it (owing, say, to the fluke occurrence of a delivery truck passing by my window), the Campanile clock is obscured from view at that instant (which *is* exactly 2:56pm). So, instead, one minute later, I look carefully at my watch, which (because my watch happens to be running one minute slow) reads exactly 2:56pm. I reason: 'It is exactly 2:56pm (p)

---

[1] I am currently preparing a longer study (Fitelson 2010), which will examine the broader debate in detail. There, I will further strengthen Warfield's case, by providing independent reasons to doubt the cogency of the sorts of 'resistance' strategies that are typically discussed in this recent literature.

therefore (*q*) I am not late for my 7pm meeting'. Thus (supposing Warfield is right), I have inferential knowledge that *q*, based on a relevant premise *p*, which is a falsehood. Now for the twist. If my belief that *p* had been *true,* then (we can plausibly suppose) it would have been based on my reading (at exactly 2:56pm) of the Campanile clock, which would have read exactly 2:56. Unbeknownst to me, however, the Campanile clock has been (and would have been) *stuck at 2:56* for some time.

This variation on Warfield's example seems not only to be an example of KFF (supposing Warfield is right about his original case), but also an example of KFF*. That is, we now seem to have a case of inferential knowledge based on a false relevant premise, *where the knowledge is counterfactually dependent on the falsity of said premise*.

3. Pushing the limits of KFF: refuting Coffman's conjecture

Examples of KFF* illustrate that the *falsity* of the basis belief *p* can play a crucial role in *explaining why* a KFF case involves inferential *knowledge*. But, as it stands, much of this explanation will *also* trade on *other explanans*. Specifically, we *also* have:

> (4) If the subject's belief that *p* had *not* been *approximately true*, then the example would *not* have constituted a case of inferential knowledge.

This might *seem* to reveal *a* sense in which the falsity of the subject's belief that *p* is *in*essential. Indeed, Coffman (2008: 190–1) conjectures that in *all* cases of KFF

> …we can identify a *true proposition p′* with the following two features:
>
> • the subject is (at least) disposed to believe *p′*,
>
> *and*
>
> • if the subject's inferential belief (that *q*) had been based on a belief in *p′*, the inferential belief would (still) have constituted knowledge.

Coffman's discussion suggests the following *alternative premise* (for our cases).

> (*p′*) It is *approximately* 2:56pm (*e.g.*, it is 2:56pm ± 2 minutes).

As it stands, our current example of KFF* does not seem to be a counterexample to Coffman's conjecture. For all I have said about our example, this (true) alternative premise *p′* may well have the two features Coffman desires. In order to turn our example into a case which (also) refutes Coffman's conjecture, we just need to modify it so as to ensure that if I had based my belief that *q* on a belief in *p′*, this would have traced back to my looking *at the Campanile* (*not* my looking *at my fancy watch*).[2] So, while the truth of (4) may[3] point (*via* Coffman's *alternative premise p′*) to *some* sense in which the *falsity* of the subject's belief that *p* is *in*essential to the acquisition of inferential *knowledge* in *the original examples of*

---

[2] One way to do this would be to modify our example above by adding to it the stipulation that I am confident that my fancy watch is *exactly* accurate, whereas I believe that the Campanile clock is only accurate to within (say) two minutes. And, as a result, I am disposed to come to believe 'it is *approximately t*' when I look at the Campanile clock and I see that it reads exactly *t*; whereas, I am disposed to come to believe 'it is *exactly t*' when when I look at my fancy watch and I see that it reads exactly *t*.

[3] Warfield gives various reasons to doubt that any Coffman-style strategy for resisting KFF can be made to work. My point here is that *even if* Coffman could effectively respond to Warfield's arguments, this wouldn't save his conjecture (as variations on our KFF*-example directly refute it). In the longer study I am preparing (Fitelson 2010), I will discuss the broader debate concerning 'knowledge from non-knowledge', and its ramifications for contemporary epistemology.

KFF, it seems that the same cannot be said regarding our strengthened examples of KFF*. In other words, it seems that *both* the *falsity and* the *approximate truth* of the subject's belief that *p* are explanatorily relevant to the presence of inferential *knowledge* in our examples of KFF*. If this is right, then we have managed to strengthen Warfield's example of knowledge from falsehood in some epistemologically interesting ways.[4]

University of California–Berkeley
Berkeley, CA 94720–2390
branden@fitelson.org

References

Coffman, E.J. 2008. Warrant without truth? *Synthese* 162: 173–94.
Fitelson, B. 2010. Knowledge from non-knowledge. Unpublished manuscript.
Warfield, T. 2005. Knowledge from falsehood, *Philosophical Perspectives* 19: 405–16.

---

[4] I would like to thank Jonathan Vogel and Ted Warfield for useful comments and discussion.