

# Counter Closure and Knowledge despite Falsehood<sup>1</sup>

---

*Reasoning that essentially involves false conclusions, intermediate or final, cannot give one knowledge.*  
Gilbert Harman

## Abstract

Certain puzzling cases have been discussed in the literature recently which appear to support the thought that knowledge can be obtained by way of deduction from a falsehood; moreover, these cases put pressure, *prima facie*, on the thesis of *counter closure* for knowledge. We argue that the cases do not involve knowledge from falsehood; despite appearances, the false beliefs in the cases in question are causally, and therefore epistemologically, incidental, and knowledge is achieved *despite* falsehood. We also show that the principle of counter closure, and the concomitant denial of knowledge from falsehood, is well motivated by considerations in epistemological theory—in particular, by the view that knowledge is first in the epistemological order of things.

## Keywords

Epistemic Closure, Counter-Closure, Knowledge from Falsehood, Knowledge First

## 1. Introduction

Closure principles for knowledge and justification have been much discussed in the epistemological literature.<sup>2</sup> But consider the following principle of *counter closure* for knowledge<sup>3</sup>:

(CCK) If (i) *S* knows *q*, and (ii) *S* believes *q* solely on the basis of a competent deduction from some premises including *p*, then (iii) *S* knows *p*.

---

<sup>1</sup> This paper is fully collaborative; authors are listed alphabetically. Thanks to Branden Fitelson for discussion and to two anonymous referees for this journal.

<sup>2</sup> See, for instance, (Dretske 1970; Nozick 1981; Williamson 2000; Cohen 2002; Pritchard 2005).

<sup>3</sup> Warfield (2005) articulates a (multi-premise) principle of this kind concerning what he terms ‘inferential knowledge’; and Federico Luzzi (2010) termed (what appears to be a single premise version of) a principle of this kind ‘counter closure’. Like closure principles, principles of counter closure are subject to various refinements and nuanced differentiations; the principle below is our preferred, multi-premise version.

Intuitively, this principle is quite plausible; and indeed, we believe it is true. However, a number of examples have recently been discussed in the literature which appear to put some pressure on this principle. Consider the following two, both due to Ted Warfield:

*Handout*

Counting with some care the number of people present at my talk, I reason: ‘There are 53 people at my talk; therefore my 100 handout copies are sufficient’. My premise is false. There are 52 people in attendance—I double counted one person who changed seats during the count. And yet I know my conclusion. (Warfield 2005: 407-8)

*Fancy Watch*

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. (Warfield 2005: 408)

Each of these cases seems to describe a counterexample to the principle of counter closure for knowledge articulated above (CCK): for in each case it seems that there is a subject (Warfield) who (i) knows a proposition (that his 100 handout copies are sufficient; that he is not late for his meeting), which (ii) he believes solely on the basis of a competent deduction from (some premises including) another proposition which he believes (that there are 53 people at his talk; that it is exactly 2:58pm), but where (iii) that latter proposition is false—and hence, by the factivity of knowledge, not known. Warfield and others have called these examples cases of *knowledge from falsehood* (KFF); we prefer to describe them as instances of *knowledge despite falsehood* (KDF).

In what follows we defend the principle of counter closure. In Section 2 we discuss the above cases, showing in some detail that they do not involve knowledge from falsehood. In Section 3 we lay out our view of these cases of knowledge despite falsehood explicitly in schematic form, before comparing it, in Section 4, to alternatives in the literature. Finally, in Section 5 we provide independent theoretical support for (CCK). The upshot is a serious and novel account of some puzzling cases in epistemology that is both theoretically well-motivated and intuitively plausible.

## 2. On the Alleged Counterexamples

If the principle of counter closure (CCK) is true, as we maintain it is, then there cannot be any counterexamples to it. What then of the so-called cases of ‘knowledge from falsehood’ described above? To begin with, we note that when faced with any

purported counterexample to counter closure we must argue either: (i) that the subject *S* of the case does not know the proposition *q*; (ii) that the subject's belief that *q* is not solely based on a competent deduction from some premises including *p*; or (iii) that the subject does know *p* after all. These options exhaust the logical possibilities: if none of these claims can be defended in a given case, then clearly counter closure fails. Now, as a matter of fact, we will argue that in each of the cases *Handout* and *Fancy Watch*, it is the claim (ii) that the subject's knowledge that *q* is solely based on a competent deduction from some premises including her belief that *p* that fails; but we point out the possibility of other responses since there might be related cases in which another option for responding is preferable—and indeed, we shall see one such case below (*Temperature*).

Consider, then, the case of *Handout*. According to Warfield, in this case the subject (Warfield himself) bases his knowledge that his 100 handouts are sufficient for each audience member to have a copy solely on the belief that there are 53 people in the audience. But this belief is false in the case as described; and so Warfield concludes that *Handout* constitutes a case of knowledge from falsehood. Clearly, if this is correct, it is a counterexample to counter closure.

But it should already seem mildly suspicious that, in the explicit description of the case, *only* the belief that there are 53 people in the room is mentioned as supporting the (justified) belief (indeed, knowledge) that Warfield's 100 handouts are sufficient; for surely the fact that 100 is more than 53 is *also* relevant.<sup>4</sup> That is, in order to infer that his 100 handouts are sufficient, from the premise that there are 53 people in the audience, Warfield must be in some way sensitive to the fact that 53 is less than 100. We mention this point not to suggest that in *Handout* Warfield does not know his conclusion (by virtue of failing to base his belief that he has sufficient handouts on this further premise), but rather to suggest that in the case of *Handout* there are almost certainly some *further*, albeit *tacit*, or *subconscious* beliefs present, which enable him to undertake the inference. And once this is recognized, we think it is fairly clear what is going on in the case at hand: the false belief that there are 53 people in the room is *epiphenomenal* as far as the causation and the justification of the belief that the 100

---

<sup>4</sup> Warfield seems to acknowledge that there is more going on in his alleged cases of knowledge from falsehood than meets the eye in his explicit descriptions: before describing the various examples, he says that he plans to 'focus on cases in which one has exactly one 'inferential argument' for one's conclusion and one's inferential argument consists of a single material premise and a suppressed conditional linking the premise to the conclusion via simple modus ponens' (Warfield 2005: 406).

handouts are sufficient is concerned; instead, this latter belief is causally grounded in, and informationally sensitive to, certain other, tacit or subconscious beliefs.<sup>5</sup>

More specifically, we think that our intuition that the subject of *Handout* knows his conclusion is tracking certain features of the case which are not spelled out explicitly, but which are nonetheless projected by readers as present. In particular, we think that what is really going on in *Handout* is something like the following: Warfield knows (perhaps tacitly) that the result of his count is ‘53’. He then forms two further beliefs: one is the explicit belief that there are 53 people in the room; the other is the tacit belief that there are approximately 53 people in the room. He then employs this latter belief, together with his knowledge that 53 is far less than 100, to conclude that his 100 handouts are sufficient. This last belief could not easily have been false, and indeed, constitutes knowledge;<sup>6</sup> but it is based upon other beliefs which also constitute knowledge, and so is no counterexample to counter closure.

It should be clear from this account of *Handout* what we will want to say in the case of *Fancy Watch*; but let us spell it out for the sake of clarity and explicitness. We think that, although the subject of the case (Warfield again) is described as inferring something true from something false, our intuition that the conclusion is known is sensitive to our understanding of how cases like this typically work, rather than to the explicit description given.<sup>7</sup> Thus, when Warfield looks at his fancy watch, he first forms the (perhaps tacit) belief that his watch reads ‘2:58’. From this perceptual belief, he infers two further things: an explicit belief, based in part on the additional

---

<sup>5</sup> As Lycan (2006) has noted, there is an unduly neglected research program in epistemology which, following Harman (1973), begins with the assumption that there can be no knowledge from falsehood, and turns this into ‘a methodology for investigating the structure of epistemic justification’ (Lycan 2006: 155); moreover, this methodology will reveal surprising facts about ‘the nature of inference itself’ (Lycan 2006: 155) and even uncover certain tacit beliefs which might otherwise go unrecognized. As may now be evident, our paper can be seen as contributing to this under-explored program. (Thanks to an anonymous referee for drawing this to our attention.)

<sup>6</sup> We do not mean to suggest that having a safe belief is sufficient for knowledge. The passage is merely meant to illustrate the plausibility of our assumption that the belief in question is in fact knowledge.

<sup>7</sup> Of course, Warfield is free to stipulate his examples in whichever way he likes; but it is a familiar point that one must be careful in describing an imagined, and genuinely possible, scenario since Kripke (1980) pointed out that a situation in which the heavenly body visible in the evening and called ‘Hesperus’ might be distinct from the one visible in the morning and called ‘Phosphorus’ without this being a situation in which Hesperus was not Phosphorus (such a situation being impossible). Our present point is the related one that our intuitions may be sensitive to the real features of such a possible situation, rather than its description. Indeed, this is what we believe to be going on in the case at hand: our intuitions, we believe, are sensitive not to Warfield’s stipulations and explicit descriptions of his cases, but rather to the way in which we respond to, interpret, and fill in the details of Warfield’s descriptions. Thus, the objection that we are disregarding Warfield’s explicit stipulations misses the point made in the main text.

belief that his watch is perfectly accurate, that it is exactly 2:58pm; and a tacit, or unconscious belief that it is almost 3pm. From this latter belief, together with the beliefs that his meeting is at 7pm, and that it takes much less than four hours to get to the meeting from his current location, Warfield infers that he is not late for his meeting. This latter belief constitutes knowledge since neither it, nor anything relevantly like it, is falsely believed in nearby possible worlds; but it is not in any way—whether evidentially, or causally—based on the belief that it is exactly 2:58pm.<sup>8</sup> This latter explicit belief, though false, is entirely epiphenomenal to the etiology of the knowledge that Warfield is not late for his meeting. Once again, we have a case of knowledge *despite*, rather than *from*, falsehood.

Our claim that Warfield's belief is in fact based on his knowledge that there are approximately 53 people in the room is strongly supported by other considerations. Assume that Warfield, in *Handout\**, is informed by someone who just counted the people in the room fifteen times and each time came to the same correct result (52), thus informing Warfield that the number of people in the room is certainly not (exactly) 53. Would Warfield get worried about whether or not he has enough handouts and retract his belief? Surely not. The explanation of why he doesn't, however, is that his belief that he has enough handouts is in fact based on his belief that there are *approximately* 53 people in the room. It is this belief that both causes and sustains Warfield's knowledge; it is therefore this belief upon which Warfield's knowledge is based.<sup>9</sup> Similar considerations hold for *Fancy Watch*.

Of course, it might be insisted that the cases at issue *are* ones in which the false, explicit belief is causally efficacious in producing the belief in the conclusion; and that while we may have a neat account of some relevantly similar cases in which a false belief accompanies some knowledge in an epiphenomenal manner, this does nothing to address the alleged counterexamples to counter closure that Warfield has articulated. But we think this response is misguided. Obviously there are cases in which subjects really do (causally) base certain beliefs inferentially on explicitly held false beliefs: but we think such cases will not involve knowledge of the conclusions; and, moreover, if they are described appropriately, we will not be tempted to think of

---

<sup>8</sup> See fn. 6.

<sup>9</sup> The phrase 'based on' tracks causal sustenance, which may come apart from causal pedigree (though not in this case).

those conclusions as known (that is, we will not have the intuition that they are known). Consider, for example, the following case:

*Temperature*

Bill lives in Pittsburgh, but has to go on a business trip to Toronto in March. He has heard that Canada has long, cold winters, and wonders whether he will need to bring his winter coat. He searches on Google for ‘temperature Toronto March’ and finds that the average low is -5 degrees. But then he remembers that Canada uses degrees Celsius and wonders how much that is in Fahrenheit. He decides to calculate. Unfortunately, Bill misremembers the conversion formula: although he should add 32 (to get 27), then divide by 9 (to get 3), and finally multiply by 5 (to get 15), he thinks that he is supposed to divide by 5 and multiply by 9. At the same time, however, Bill makes typing errors on his calculator: he keys in 5 (rather than -5) then adds 3 and adds 2, rather than adding 32; he divides the result (10) by 5 (to get 2), then multiplies by 9 (to get 18). As a result he believes that the average low in Toronto in March is 18 degrees Fahrenheit, and he concludes he will need his winter coat.

Intuitively, in *Temperature* Bill does not know that he will need his winter coat: if he hadn’t made the typing errors, he would have believed that it would be 49 degrees Fahrenheit in Toronto in March, and would have concluded that he did not need his winter coat; and since he might easily not have made the typing errors, he might easily have had a false belief regarding the issue of whether he ought to bring his coat. Moreover, in our opinion, the reason why Bill’s belief that he will need his winter coat does not constitute knowledge is that in *Temperature*, Bill’s explicit numerical beliefs play a crucial role in causing his belief that he won’t need his winter coat: accordingly, his false belief that it will be 18 degrees Fahrenheit in Toronto in March is not merely epiphenomenal in the causal derivation of his belief that he won’t need his coat; and as a result, this latter belief does not constitute knowledge.<sup>10</sup>

---

<sup>10</sup> Note also that, in Warfield’s case, his belief that he has enough handouts (isn’t late for his meeting) is safe: there are no nearby worlds in which Warfield believes falsely that he has enough handouts. Of course, on our account that is so because his belief is based on his knowledge that there are approximately 53 people in the room (it’s almost 3pm). In *Temperature*, however, Bill’s belief that he won’t need his winter coat isn’t safe and doesn’t constitute knowledge. Moreover, note that it is not possible to turn Warfield’s cases into cases in which it is obvious that Warfield doesn’t know the propositions in question, as is the case in *Temperature*. That is so because we cannot easily stipulate away the fact that Warfield, in his examples, believes that there are approximately 53 people in the room (it’s almost 3pm), while keeping in place the stipulation that he believes that there are exactly 53 people in the room (it’s 2:58pm). Such a stipulation would make Warfield’s cases exceedingly strange and psychologically implausible: our intuitions would, with such stipulations in place, be no longer distinct and reliable. Finally, it is noteworthy that it is the very fact that we cannot coherently stipulate away that Warfield believes that there are approximately 53 people in the room (it’s almost 3pm) which makes the intuition that Warfield knows that he has enough handouts (isn’t late for his meeting) so strong: we fill in the implicit detail that Warfield holds the mentioned implicit belief and uses it as a

If *Handout* and *Fancy Watch* were taken by readers to be genuinely like *Temperature* in involving a falsehood in the causal etiology of the belief in the conclusion, they too would be intuitively judged to be cases in which the subject lacked knowledge. It is only because readers are sensitive to how the relevant beliefs are in fact formed and sustained that they have the intuition that the beliefs in the conclusions are known; but given the actual causal origin of these beliefs in typical cases of this kind, the cases do not constitute counterexamples to the principle of counter closure (CCK).

### 3. The View

We take our discussion of the above cases to be fairly simple and straightforward. Nevertheless, it will be useful to have an explicit, schematic characterization of our view available. Simply put, we endorse the following account of the kinds of cases of knowledge despite falsehood discussed above:

*Knowledge Despite Falsehood (KDF)*

In apparent cases of KDF, there are two true propositions  $t_1$  and  $t_2$  such that:

- (1)  $t_1$  evidentially supports both  $p$  and  $t_2$  for  $S$ ;
- (2)  $t_2$  is entailed by  $p$ ;
- (3)  $S$  knows both  $t_1$  and  $t_2$ ;
- (4)  $S$ 's belief that  $q$  is properly based on her knowledge that  $t_2$ .

To see how this account handles the above cases, note that the variables in *Handout* and *Fancy Watch* take the following values:

*Handout:*

- $t_1$ : The result of my (Warfield's) count was '53'.  
 $t_2$ : There are 53 people at my talk *give or take a few*.  
 $p$ : There are 53 people in the room.  
 $q$ : I have enough handouts.

*Fancy Watch:*

- $t_1$ : My (Warfield's) watch reads '2:58'.  
 $t_2$ : It's approximately 3pm.  
 $p$ : It's 2:58.  
 $q$ : I'm not late for my meeting.

In both cases,  $t_1$  is clearly known by the subject: it also evidentially supports or justifies both  $t_2$  and  $p$  to at least some degree; and in both cases  $t_2$  is entailed by  $p$ . Finally, note that the truth  $t_2$  is a rough-and-ready approximation of the falsehood  $p$ ,

---

base for his knowledge that he has enough handouts (isn't late for his meeting), as we have argued in the main text.

which is, in addition, known in the examples. Moreover, since Warfield's belief that  $q$  is properly based in his knowledge that  $t_2$ —namely, by competent deduction—it follows that the examples are not cases of knowledge from falsehood but rather cases of *knowledge despite falsehood*.

Now, we will not attempt to justify the claim that  $t_2$  is known tacitly in this paper by means of the provision of an analysis of tacit knowledge: we take it as given that there is such a phenomenon. Of course, it might be furthermore considered controversial that a subject can have tacit knowledge that plays an active role in his or her reasoning, as is required by our account of the above cases. However, it should be noted that the derivation itself is most plausibly construed to be automatic and nonconscious or tacit, too; indeed, as cognitive psychologists teach us, much of our actual cognitive processing is of precisely this kind.<sup>11</sup> In any case, we take it to be obvious from our considerations in Section 2 that the beliefs at issue are in fact based on the beliefs in the relevant propositions  $t_2$ , as required by our condition (4), and as the counterfactual considerations there have shown.

Finally, note that (KDF) avoids a potential problem to which Warfield (2005) has drawn attention.<sup>12</sup> He claims that showing that there is a true proposition  $t$  in the vicinity that functions as a basis for  $S$ 's knowledge that  $q$  will not suffice, if the mentioned truth  $t$  is only believed because it follows from  $p$ ; for in that case, the falsehood  $p$  would still play an essential role in justifying the belief that  $q$ , and we would still be dealing with a case of knowledge from falsehood. But clearly, this is not a problem for (KDF). To see this note that in the above examples, the entailed truth  $t_2$  is believed because  $t_1$  is believed, not because  $p$  is believed—that was the upshot of Section 2. Moreover, note that  $p$  is believed because  $t_1$  is believed, but it is not the case that  $t_2$  is believed because  $p$  is; rather, there is an evidential shortcut, directly from  $t_1$  to  $t_2$ , by-passing the falsehood  $p$  entirely. And, finally, since  $t_2$  is known and  $q$  competently deduced from  $t_2$ —together with auxiliary premises that are also known—it is perfectly suited to function as a basis for the knowledge that  $q$ .

(KDF) articulates in theory the approach we adopted in practice in response to the examples described above. In the next section, we compare it to the alternatives in the literature.

---

<sup>11</sup> See, for instance, Evans and Stanovich (2013).

<sup>12</sup> See also Klein (2008) for pertinent discussion.



## 4. Comparison with Other Accounts

### 4.1 Warfield

Warfield (2005) discusses a number of possible accounts of KDF (as opposed to KFF), only two of which he considers even remotely plausible (and all of which he rejects). Here is the first of these:

#### *Warfield's First Proposal*

In apparent cases of knowledge from falsehood, there is a true proposition  $t$  such that:

- (1)  $t$  is entailed by the involved falsehood  $p$ ;
- (2)  $t$  is (propositionally) justified for  $S$ ;
- (3)  $S$  (at least) dispositionally believes  $t$ .<sup>13</sup>

Warfield has it that this first proposal fails for the same reasons as what we shall call his second proposal. Let us therefore consider that proposal next:

#### *Warfield's Second Proposal*

In apparent cases of knowledge from falsehood, there is a true proposition  $t$  such that:

- (1)  $t$  is evidentially supported by the evidence for the involved falsehood  $p$ ;
- (2)  $t$  is (propositionally) justified for  $S$ ;
- (3)  $S$  (at least) dispositionally believes  $t$ .<sup>14</sup>

Obviously, our account of KDF is similar to Warfield's proposals in some respects: taking Warfield's  $t$  to be our truth  $t_2$  throughout, we accept the first condition (1) of each of his proposals; and since knowledge requires both justification and belief, our condition (3) plausibly commits us to both of Warfield's (2) and (3). Nevertheless, we are not even remotely attracted to these accounts of KDF, for they are both silent on why we should take the relevant proposition  $t$  to be the basis of the subject's belief that  $q$ : it is not clear from Warfield's proposal why we should take the examples at issue to be cases of knowledge *from*  $t$ . What is most obviously missing is anything corresponding to our condition (4) above.

Be that as it may, one might nevertheless suspect that Warfield's response to his proposals creates problems for our approach, too. What is Warfield's objection? He writes:

---

<sup>13</sup> 'The suggestion is that one has knowledge despite the presence of an involved falsehood if there is a justified and (at least) dispositionally believed truth entailed by the falsehood' (Warfield 2005: 411). It seems clear that the justification in question is propositional, and that the subject is taken to possess it; but we will not make anything of this in what follows.

<sup>14</sup> 'One has knowledge despite the presence of an involved falsehood if there is a justified and (at least) dispositionally believed truth evidentially supported by the evidence for the involved falsehood' (Warfield 2005: 411-2).

Both proposals fail catastrophically in Gettier cases. I seem to see a dog in the yard [...]. I form the belief that there is a dog in the yard and then reason [...] to the conclusion that there is at least one animal in the yard. My belief is false (there is no [real] dog, only [a] toy) and my conclusion though true, because of [the presence of a] squirrel behind the brush, is not known. The [proposals] get this clear ‘no knowledge’ case wrong. They both imply that I know that there’s at least one animal in the yard. After all, there is a justified and dispositionally believed truth that is *both* evidentially supported [...] and entailed by my false belief that there is a dog in the yard: the truth is ‘there is a dog or squirrel in the yard’. The [proposals] rule that this belief epistemizes my conclusion. I therefore am judged to have knowledge [that there is an animal in the yard] in this case which goes against the standard and correct Gettier verdict. (Warfield 2005: 412)

Yet it is easy to see how our account avoids Warfield’s objection: conditions (3) and (4) of our account are not satisfied in Warfield’s Gettier case. To see this note that the proposition spelling trouble for Warfield’s proposal—namely, the proposition that there is a dog or a squirrel in the yard—is not *known*; nor does the subject (Warfield) *base* his belief on it. The case accordingly poses no trouble for us.

#### 4.2 Coffman

It will also be illuminating to compare our view to EJ Coffman’s. Coffman (2008) accepts that there are genuine cases of knowledge from falsehood: that is, unlike us, he doesn’t argue against this view of Warfield’s; rather, he attempts to uncover necessary conditions for cases of knowledge from falsehood. In particular, he proposes the following:

##### *Coffman’s Conjecture*<sup>15</sup>

In all cases of knowledge from falsehood, ‘we can identify a *true proposition* which has the following two features:

- (1) The subject is (at least) disposed to believe the proposition, *and*
- (2) If the subject’s inferential belief were based on a belief in that proposition, the inferential belief would still constitute knowledge (other things equal).<sup>16</sup>

As should be obvious to the reader, we agree with Coffman that his two conditions are satisfied in (e.g.) Warfield’s cases; though, of course, we do not agree with his assumption (which, to be fair, he doesn’t intend to argue for in his paper) that the cases at issue are genuine cases of knowledge from falsehood.

#### 4.3 Fitelson

Branden Fitelson (2010) purports to refute *Coffman’s Conjecture*. While Fitelson doesn’t offer a view of KFF or KDF himself, he presents the following interesting case:

---

<sup>15</sup> The label ‘Coffman’s Conjecture’ is due to Fitelson (2010).

<sup>16</sup> (Coffman 2008: 190-1).

### *Campanile*

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*) therefore (*q*) I am not late for my 7pm meeting’. Thus [...], 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. (Fitelson 2010: 667, emphasis in original)

Fitelson claims that this example is stronger than Warfield’s original *Fancy Watch* case, for in *Campanile* his knowledge that he isn’t late for his meeting is, he suggests, ‘counterfactually dependent on the falsity of the premise’ that it is exactly 2:56pm when he looks at his watch: if Fitelson’s belief that it is exactly 2:56pm had been true, then he would have been Gettiered, and so wouldn’t have known that proposition.

It is not this case which Fitelson takes to pose a problem for Coffman, but it is worth remarking that our account can handle *Campanile* elegantly and with ease. In fact, our treatment of this case doesn’t differ from that of the others we have discussed. To see this note that Fitelson knows (perhaps tacitly) that his fancy watch reads ‘2:56pm’. He then forms two further beliefs: one is the explicit belief that it is exactly 2:56pm; the other is the tacit belief (and indeed knowledge) that it is approximately 2:56pm. He then employs this latter belief, together with his knowledge that his meeting is at 7pm, and that it takes much less than four hours to get to the meeting from his current location, to conclude that he isn’t late for his 7pm meeting. This latter belief constitutes knowledge since it is based on knowledge and neither it, nor anything relevantly like it, is falsely believed in nearby possible worlds; but it is not based on the belief that it is exactly 2:56pm. This latter explicit belief, though false, is entirely epiphenomenal to the etiology of the knowledge that Fitelson is not late for his meeting. Once again, we have a case of knowledge *despite*, rather than *from*, falsehood and nothing about Fitelson’s new example presents a problem for the independently motivated account proposed above.

There is, however, a slight worry one might have. Fitelson suggests that we can modify *Campanile*

so as to ensure that if I had based my belief that [I won't be late for my meeting] on a belief in [the proposition that it is approximately 2:56pm], this would have traced back to my looking at the Campanile (not my looking at my fancy watch). (Fitelson 2010: 668)

Clearly, Fitelson presupposes here that in *Campanile* he doesn't in fact base his belief that he won't be late on his knowledge that it is approximately 2:56pm: this is a (pragmatic) presupposition of (an utterance of) the quoted counterfactual. As should be painstakingly clear by now, we reject that presupposition. However, Fitelson proposes that the counterfactual in the above-quoted passage (and its presupposition) is true and, moreover, remains true if we amend *Campanile* by stipulating that he believes his fancy watch to be exactly accurate, but the Campanile clock only to be 'accurate within (say) two minutes.' To illustrate this further he says in a footnote:

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 [...] I look at my fancy watch and I see that it reads exactly  $t$ . (Fitelson 2010: 668, fn. 2)

If we understand him correctly, what Fitelson has in mind here is that, in the amended case, he believes (or is disposed to believe) that it is approximately 2:56pm *only* when looking at the Campanile, but not when looking at his fancy watch. He takes this to show that Coffman's Conjecture is false, since (if Fitelson is right) there is no true  $t$  such that were one to base one's belief that  $q$  on it, that belief would constitute knowledge. And the case might be thought to pose problems for us too: in *Modified Campanile*, one might argue, Fitelson cannot have based his belief that he won't be late on his belief that it's approximately 2:56pm, for he doesn't even have that belief!

We find this line of argument unconvincing. To see why, note that it is by no means clear that Fitelson would, or even could, fail to believe that it is approximately 2:56pm in the sorts of case at issue. The stipulation that this is so in *Modified Campanile* is problematic, for Fitelson is also stipulated to believe, in the same case, that it is *exactly* 2:56pm. Given that the proposition that it is exactly 2:56pm rather obviously entails the proposition that it is approximately 2:56pm, it would, assuming minimal rationality, be rather surprising for Fitelson to believe the former but not the latter; and indeed, some have argued that this is impossible.<sup>17</sup>

---

<sup>17</sup> Note that the notion of implicit belief is frequently defined in terms of 'obvious entailment'. Thus, this claim comes out analytic on these accounts. In (Field 1978), for instance, which builds upon an

Now, one might be tempted to think that one can and often does believe that it's exactly 2:56pm without believing that it's approximately 2:56pm because of the felicity of utterances of the form of (1) or (2):

- (1) I don't believe it's *approximately* 3pm; I believe it's *exactly* 3pm.
- (2) It's not *approximately* 3pm; it's *exactly* 3pm.

However, utterances of this form are cases of meta-linguistic negation, such as utterances of (3)-(6):

- (3) I don't believe she's *smart*; I believe she's *brilliant*!
- (4) She isn't *smart*; she's *brilliant*!
- (5) I don't believe he's *tall*; I believe he's *huge*!
- (6) He isn't *tall*; he's *huge*!

Accordingly, we take it that Fitelson's stipulation to the effect that he doesn't believe that it's approximately 2:56pm is either incoherent or, at the very least, extremely strong. Thus, in standard cases of the kind suggested by the proposed modification to *Campanile* Fitelson believes and knows (at least tacitly) that it is approximately 2:56pm, and he bases his belief that he isn't late on that knowledge—despite Fitelson's claim that the belief that he isn't late is 'inferred' from his explicit belief that it's exactly 2:56pm.<sup>18</sup> By contrast, if there are possible cases in which a subject has the combination of attitudes that Fitelson stipulates,<sup>19</sup> that subject simply will not count as knowing his or her conclusion.<sup>20</sup>

Fitelson's *Modified Campanile* case therefore fails to pose problems for us or for Coffman. So, is there anything wrong with *Coffman's Conjecture*? Aside from the assumption, previously noted, that there are genuine cases of knowledge from falsehood, we can find no fault with it. It is, however, incomplete: while Coffman's conditions are indeed necessary for knowledge despite falsehood, they are not sufficient (of course, he never said they were). That this is so can be readily seen by

---

account in (Dennett 1978), we find the idea that a subject  $x$  implicitly believes  $p$  iff for  $x$   $p$  is an obvious entailment of a  $q$   $x$  explicitly believes and  $x$  does not believe  $p$  explicitly.

<sup>18</sup> Again, we take it to be true that, in *Campanile*, if Fitelson were not to believe that it is approximately 2:56pm, he wouldn't believe (and so wouldn't know) that he isn't late for his meeting. That is, in the closest worlds in which Fitelson does not believe that it is approximately 2:56pm he also doesn't believe that he isn't running late—independently of whether those worlds are ones in which he looks at the *Campanile* clock or his fancy watch.

<sup>19</sup> Perhaps, for instance, an artificial intelligence could have the beliefs in question explicitly encoded in some way.

<sup>20</sup> Note that if you think that the semantics of 'approximately' doesn't work in the way outlined above and that 'approximately,  $p$ ' entails  $\neg p$  (and, therefore, that  $p$  entails '¬approximately  $p$ '), then nothing stops us from rephrasing our view by picking out  $t$  differently—for instance, as the proposition that it is *in the region of 2:56pm, around and including 2:56pm, or approximately and including 2:56pm*, etc.

noting that Coffman's conditions are met in Warfield's Gettier case described above. As we have seen, our account fares better here, and can therefore be considered an improvement.

#### 4.4 Klein

Peter Klein (2008) accepts the view that there are cases of knowledge from falsehood, and aims to give a *definition* of them—he calls them cases of '*useful falsehood*'. Here is Klein's (2008: 48) proposed definition:

*Klein's Definition*

*S*'s belief that *p* is a useful falsehood for acquiring knowledge that *q* [...] iff:

- (1) *p* is false,
- (2) *S*'s belief that *p* is doxastically justified,
- (3) *S*'s belief that *p* is essential in the causal production of the belief that *q*,
- (4) *p* propositionally justifies *q*,
- (5) *p* entails a true proposition, *t*,
- (6) *t* propositionally justifies *q*,
- (7) whatever doxastically justifies the belief that *p* for *S* also propositionally justifies *t* for *S*.

As is quickly clear on reflection, our view as explicated above is perfectly compatible with the idea that in Warfield's cases, and in other cases of apparent knowledge from falsehood, all of Klein's conditions are satisfied—apart from condition (3): as we argued in Section 2, the relevant false beliefs are not essential in the causal production and/or sustenance of the subject's knowledge. To see this, note that Klein understands the notion of essentiality at issue in (3) counterfactually: had *S* not come to believe the false *p*, then she wouldn't have come to conclude/infer that *q*.<sup>21</sup> Yet this conditional is false: in the cases at issue, there is an independent proposition, *t*<sub>2</sub>, in the vicinity of *p* that is known and the knowledge of which both initially caused and now causally sustains *S*'s belief that *q*; accordingly, even if *S* had not come to believe *p*, she would still have concluded that *q*. Thus, Klein's essentiality thesis is false. Indeed, this was made obvious already in Section 2 by the fact that if (as in *Handout\**) someone told Warfield that there are certainly not 53 people in the room, he would still believe *q*—that he has enough handouts. Similar considerations apply to the cases Klein

---

<sup>21</sup> See (Klein 2008: 41): 'A false belief plays an essential causal role in producing knowledge in each of the cases. That is, if the false belief were simply removed from the actual causal chain that resulted in knowledge, no causal chain resulting in the cognition would remain.'

discusses; but to spare the reader unnecessary repetition we will not go through the exercise of showing that this is so.<sup>22</sup>

Clearly, Klein assumes that the above notion he defines is not empty. But if our account of the mentioned cases is right, then the subject doesn't believe the relevant proposition solely on the basis of a competent deduction from a falsehood, and the belief in the falsehood is, accordingly, inessential in Klein's sense. The above definition does nothing to guarantee that there are any cases of KFF so understood; on the contrary, it remains plausible that there are none.

## 5. Theoretical Reflections

We take it that our explanation of the above examples (Sections 2 and 3) is intuitively plausible and superior to alternative accounts in the literature (Section 4). However, there are further, theoretical reasons to prefer our approach: in particular, the principle of counter closure (CCK) can be derived from independently motivated considerations in epistemological theory; and our preferred account of the alleged counterexamples therefore receives independent corroboration from such considerations. We begin with some points of conceptual and terminological clarification.

Assuming a broadly evidentialist account of epistemic justification, our beliefs (and other credences) ought to be based on our evidence; and they are justified only if they are so based. From an evidentialist point of view, this much strikes us as uncontroversial, and indeed, constitutes a simple articulation of the relationship between our concepts of evidence, justification, and the basing relation. Nevertheless, it allows us to define evidence functionally as *whatever it is that our beliefs can be based upon that can serve to justify them*. Now, a *knowledge first* approach to epistemology is one according to which it is knowledge which can serve to justify our beliefs and credences—that is, it is knowledge which can serve as the epistemological basis for our beliefs. Given the functional definition of evidence above, the knowledge firster can therefore say that knowledge is evidence ( $E = K$ )—that is, that the totality of a subject's evidence is the collection of all and only those propositions that he or she knows. Indeed, in the context of the above definition of evidence, the

---

<sup>22</sup> Klein (2008: 37) discusses four cases in total. Three of them are very similar to the cases discussed here, and our account applies; but in the final case—that of the *Ptolemaic Astronomer*—we (and our informants) do not share Klein's assessment that the subject has knowledge.

claim that knowledge is evidence in this sense is nothing other than the commitment to a knowledge first approach to epistemology.<sup>23</sup>

Next, we consider the nature of epistemic justification. It is common in epistemology to distinguish between propositional and doxastic justification. From an evidentialist perspective, propositional justification is, in effect, a logical notion: some propositions  $\Gamma$  justify a proposition  $p$  to degree  $d$  if and only if the probability of  $p$  given  $\Gamma$  is  $d$ ; as a special case, they justify  $p$  fully (conclusively, or to degree 1) when  $p$  is a logical consequence of  $\Gamma$ . But what is doxastic justification? It seems that (at least) three conditions must be met if some beliefs are to justify a credence: first, the beliefs in question must themselves be justified; second, the propositions believed ( $\Gamma$ ) must propositionally justify the object of the credence ( $p$ ) to the degree ( $d$ ) to which it is believed; and third, the credence must be caused by the beliefs in a manner that is sensitive to this second fact.

We take this evidentialist view of epistemic justification to be rather plausible and attractive, not least due to its remarkable simplicity. However, we do not intend to argue for these independently plausible claims here; we simply take them for granted. Our aim, recall, is to provide a novel argument in support of the counter closure principle for knowledge. But we are now in a position to do this: for counter closure can be seen to follow quite naturally from knowledge first and some auxiliary assumptions.

The argument goes as follows. Assume that  $S$  knows that  $q$ . Then  $S$ 's belief that  $q$  must be justified (since justification is necessary for knowledge<sup>24</sup>). Now assume further that  $S$ 's belief that  $q$  is solely based on her belief that  $p$  (by way of a competent deduction). Then  $S$ 's belief that  $p$  must serve to justify her belief that  $q$ . But then, by the principle of knowledge first, according to which it is knowledge which serves to justify beliefs and credences,  $S$ 's belief that  $p$  must constitute knowledge. So, if  $S$  knows  $q$ , and her belief that  $q$  is solely based on her belief that  $p$  (by way of a competent deduction), then  $S$  knows  $p$ ; but this is just the principle of counter closure itself, which therefore must be true.

---

<sup>23</sup> See (Williamson 2000) for a knowledge first approach to epistemology.

<sup>24</sup> In fact, this assumption is strictly stronger than required, for we apply it only in the case of *inferential* knowledge—a version of the assumption restricted to just such cases would accordingly serve our purposes equally well (thanks to an anonymous referee for drawing this point to our attention). Nonetheless, we accept the assumption in full generality, and due its simplicity, have appealed to this unrestricted thesis in the main text.



The principle of counter closure (CCK), far from being an *ad hoc* and dogmatic assumption (as advocates of KFF suggest), follows from some very well-motivated principles of epistemological theory. In the absence of compelling reasons to reject it, we submit that it is true.

## 6. Conclusion

In the literature, the existence of knowledge from falsehood (KFF) has been presented as an interesting result that is supported by particular examples. We have argued that the examples do not in fact support this view, but are instead quite naturally regarded as cases of knowledge despite falsehood (KDF). Moreover, we have argued that the principle of counter closure for knowledge (CCK) follows fairly straightforwardly from a knowledge first approach to epistemology. Finally, we have shown that the only ‘downside’ of a well-developed KDF position such as the one defended here is the recognition that subjects have tacit knowledge which is causally efficacious in their deductive reasoning. But this can hardly be considered a cost: it is now well-established that not all genuine psychological processes are open to our view. Accordingly, we concur with Gilbert Harman (1973: 120), when he says that ‘[r]easoning that essentially involves false conclusions, intermediate or final, cannot give one knowledge.’

Brian Ball

St Anne’s College, 56 Woodstock Road, Oxford, OX2 6HS, email: [brian.ball@philosophy.ox.ac.uk](mailto:brian.ball@philosophy.ox.ac.uk)  
New College of the Humanities, 19 Bedford Square, London, WC1B 3HH, email: [brian.ball@nchum.org](mailto:brian.ball@nchum.org)

Michael Blome-Tillmann

Faculty of Philosophy, University of Cambridge, Sidgwick Avenue, Cambridge, CB3 9DA, UK, e-mail: [mb938@cam.ac.uk](mailto:mb938@cam.ac.uk)  
Department of Philosophy, McGill University, 855 Sherbrooke St. W., Montreal, QC, H3A 2T7, CA, e-mail: [michael.blome@mcgill.ca](mailto:michael.blome@mcgill.ca)

## References

- Coffman, E. J. (2008). ‘Warrant without truth?’, *Synthese* 162(2): 173-94.  
Cohen, S. (2002). ‘Basic Knowledge and the Problem of Easy Knowledge’, *Philosophy and Phenomenological Research* 65(2): 309-29.  
Dennett, D. C. (1978). ‘Brain Writing and Mind Reading’, in: *Brainstorms: Philosophical Essays on Mind and Psychology*. D. C. Dennett. Montgometry, Vermont, Bradford Books: 39-50.  
Dretske, F. (1970). ‘Epistemic Operators’, *Journal of Philosophy* 67: 1007-23.

- Evans, J. ST. B. T. and Stanovich, K. E. (2013). 'Dual-Process Theories of Higher Cognition - Advancing the Debate', *Perspectives on Psychological Science* 8(3): 223-41.
- Field, H. H. (1978). 'Mental Representation', *Erkenntnis* 13: 9-61.
- Fitelson, B. (2010). 'Strengthening the case for knowledge from falsehood', *Analysis* 70(4): 666-9.
- Harman, G. (1973). *Thought*. Princeton NJ, Princeton University Press.
- Klein, P. D. (2008). 'Useful False Beliefs', *Epistemology: New Essays*. Q. Smith. Oxford, Oxford University Press: 25-62.
- Kripke, S. (1980). *Naming and Necessity*. Oxford, Blackwell.
- Luzzi, F. (2010). 'Counter-Closure', *Australasian Journal of Philosophy* 88(4): 673-83.
- Lycan, W. (2006). 'The Gettier Problem Problem', *Epistemology Futures*. S. Hetherington. Oxford, Clarendon Press: 148-68.
- Nozick, R. (1981). *Philosophical Explanations*. Oxford, OUP.
- Pritchard, D. (2005). *Epistemic Luck*. Oxford, Clarendon.
- Warfield, T. A. (2005). 'Knowledge from Falsehood', *Philosophical Perspectives* 19(1): 405-16.
- Williamson, T. (2000). *Knowledge and Its Limits*. Oxford, OUP.