Reconsidering the Alleged Cases of Knowledge from Falsehood[1]

Kok Yong Lee, National Chung Cheng University

Abstract

A number of philosophers have recently proposed several alleged cases of "knowledge from falsehood," i.e., cases of inferential knowledge epistemised by an inference with a false crucial premise. This paper examines such cases and argues against interpreting them as cases of knowledge from falsehood. Specifically, I argue that the inferences in play in such cases are in no position to epistemise their conclusions.

1. Introduction

Recent literature has seen an increasing interest in *knowledge from falsehood*, i.e., knowledge epistemised solely by an inference that contains a false belief as one of its *crucial premises*, where a premise is crucial to an inference just in case the inference does not go through without it. Advocates have constructed several alleged cases of knowledge from falsehood, in which a certain false belief appears to play a crucial role in epistemising the target belief.[2] Opponents typically argue that, in such cases, the false belief in play is not a (relevant) premise of the epistemising inference.[3] To take a different route, I will construct an argument against the knowledge-from-falsehood interpretation of such cases by *granting* that the false belief is indeed a crucial premise of the inference in play.

The following will be structured as follows. Section 2 introduces an alleged case of knowledge from falsehood. Section 3 examines this case and shows that it involves a valid inference with a false crucial premise. I then argue that such inferences, by themselves, are in no position to epistemise its conclusion. Section 4 discusses another alleged case of knowledge from falsehood, which involves a strong inductive inference with a false crucial premise. I will argue that such inferences, by themselves, are in no position to epistemise its conclusion, too. Either way,

---

[2] Warfield (2005) and Klein (2008) are two early advocates of knowledge from falsehood. Also see Hiller (2013). Arnold (2013) has used the idea of knowledge from falsehood to argue against Williamson's (2000) account of evidence, which identifies evidence with knowledge. For a reply, see Littlejohn (2013).
[3] See Ball and Blome-Tillmann (2014), Montminy (2014), and Schnee (2015). For replies, see Buford and Cloos (2018) and Luzzi (2019).

the alleged cases of knowledge from falsehood are really not cases of knowledge from falsehood. Section 5 addresses and replies to two possible objections to my arguments.

2. An Alleged Case of Knowledge from Falsehood

Ted Warfield is one of the earliest theorists to argue for knowledge from falsehood. In the article "Knowledge from Falsehood," Warfield proposes five alleged cases of knowledge from falsehood, which have become the focus of the subsequent discussion. The following is one of Warfield's cases:

> *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.[4]

Intuitively, the subject knows that her 100 handout copies are sufficient for people at her talk. Moreover, on the face of it, the subject seems to acquire her knowledge *via* an inference that contains a false premise, i.e., the belief that there are 53 people at the subject's talk. Warfield further argues that this false belief is a crucial premise of the inference (in his words, a 'relevant premise').

Let us call the thesis that, in cases such as *Handout*, a certain false belief is a crucial premise of the inference in play 'the false-crucial-premise hypothesis' ('the FCP hypothesis' in short). Here, 'the inference in play' refers to the inference that supposedly epistemises the target belief. Clearly, the FCP hypothesis must hold in all cases of knowledge from falsehood.

Critics of knowledge from falsehood have typically focused on challenging the FCP hypothesis head-on. Some claim that the inference in play does not contain the false belief as one of its premises[5], while others argue that there are two distinct inferences involved, and the one that actually epistemises the target belief does not contain the false belief as one of its premises.[6] Both proposals have been criticised.[7] Without delving into the debate, I suggest looking at the issue by assuming the FCP hypothesis. Doing so allows us to highlight a serious problem in interpreting cases such as *Handout* as cases of knowledge from falsehood, or so I will argue.

3. Valid Deductive Inferences with a False Crucial Premise

To begin with, consider *Handout*.[8] Assuming the FCP hypothesis, the inference in play—that is, the inference that epistemises the conclusion that the subject's 100 handout copies are sufficient for the people at her talk—contains, as one of its crucial premises, the false belief that there are 53 people at the talk. As an illustration, we can reconstruct the inference in play in *Handout* as follows:

> P1   There are 53 people at my talk.

---

[4] Warfield (2005: 407-8).
[5] See Ball and Blome-Tillmann (2014).
[6] See Montminy (2014).
[7] See Buford and Cloos (2018) and Luzzi (2019).
[8] The argument below applies to Warfield's other cases.

P2   If P1, then my 100 handout copies are sufficient for the people at my talk.

C1   Therefore, my 100 handout copies are sufficient for the people at my talk.

Let us briefly comment on P1-C1. First, this reconstruction respects the FCP hypothesis; it is clear that the inference does not go through without the false belief P1. Second, P1-C1 also captures the inference that we intuitively envisage the subject carrying out in *Handout*. For one thing, it seems natural to read the subject's explicit inference in *Handout* as an enthymeme of P1-C1 or something equivalent. For another, it is plausible that both the premises and the conclusion of P1-C1 are exploited by the subject when conducting the inference. P1 and C1 are explicitly stated in *Handout*. Hence, it is natural to envisage that, in *Handout*, the subject actively infers C1 from P1. Moreover, the subject must also uphold P2, or something along this line, in order for her to competently deduce C1 from P1. Just imagine how baffling it will be if we are told that the subject does not believe P2 or has explicitly denied P2. By contrast, a subject, who infers C1 from P1 while explicitly denying P2, appears to be confused and thus cannot be considered as knowing C1 at all.

The FCP hypothesis, together with the fact that the subject's inference in *Handout* is deductively valid, gives rise to the following analysis:

A1   In *Handout*, the inference in play is a valid deductive inference, which contains a false belief as one of its crucial premises.

Granting A1, however, a serious problem arises for those who want to interpret *Handout* as a case of knowledge from falsehood.

The key point is that a valid inference with a false crucial premise is not *truth-conducive* in the sense that "[its] existence makes the [conclusion] more likely to be true, either objectively or from the point of view of the subject."[9] As a matter of fact, with one of its crucial premises being false, a valid inference is not even *truth-relevant* in the sense that it "affect[s] how likely it is that the [conclusion] is true, either by raising or by lowering that likelihood."[10] Just as we cannot tell whether the conclusion of an invalid argument is true or false *ceteris paribus*, we are equally incapable to tell whether the conclusion of a valid argument with a false crucial premise is true or false *ceteris paribus*.

However, on the ordinary notion of epistemising inference, it is extremely plausible that an inference is in a position to epistemise its conclusion only if it is truth-relevant, i.e., only if it either increases or decreases the likelihood of its conclusion (note: one might want to argue that the alleged cases of knowledge from falsehood showed precisely that this idea was wrong. I will come back to this point in Section 5). If an inference is not relevant to its conclusion being true, how can we come to know its conclusion solely on the basis of this inference? Put in another way, since a valid argument with a false crucial premise is no more relevant to the truth of its conclusion than an invalid argument is, the former is on an epistemic par with the latter. And since invalid arguments do not epistemise their conclusions, so do valid arguments with a false crucial premise (note: we are ignoring invalid arguments that are nonetheless inductively strong. We will address inductive arguments the next section).

Of course, the conclusion of an unsound inference may still be known if the false premise involved is redundant and/or if the conclusion is self-evident or substantiated by some other means.

---

[9] Stanley (2005: 1).

[10] DeRose (2009: 24).

But these are beside the point, as we are concerned with whether or not a deductive inference with a false *crucial* premise *by itself* is in a position to epistemise its conclusion.[11]

In brief, a valid inference with a false crucial premise, by itself, cannot epistemise its conclusion. To elaborate with an example, consider the following inference:

P3   Phlogiston exists.

P4   If P3, then there are some falsehoods in modern chemistry.

C2   Therefore, there are some falsehoods in modern chemistry.

The inference P3-C2 is valid, and P3, which is false, is a crucial premise for deducing C2. It is clear that the inference P3-C2, by itself, does not substantiate C2. Consequently, the inference P3-C2 is not in a position to epistemise C2—there is no way one can come to know C2 solely on the basis of P3-C2, not even if C2 is true, and one is justified in believing both P3 and P4.

Now, as noted, when *Handout* is interpreted in accordance with A1, the subject's inference can be reconstructed as P1-C1, which, though valid, contains the false belief P1 as one of its crucial premises. Hence, just as the inference P3-C2, by itself, fails to epistemise C2, the inference P1-C1, by itself, also fails to epistemise C1. In other words, if *Handout* is to be interpreted along the line of A1, the case is not a case of knowledge from falsehood, as it is not a case of knowledge in the first place. Since A1 is a corollary of the FCP hypothesis, this shows that, given the FCP hypothesis, it is problematic to interpret cases such as *Handout* as cases of knowledge from falsehood.

Let us make three comments. First, the argument above does not imply that the subject in cases such as *Handout* lacks knowledge of the target proposition. In fact, the arguments against the FCP hypothesis proposed in this paper (including the aforementioned argument and the one articulated in Section 4) are compatible with the idea that the subject in the alleged cases of knowledge from falsehood possesses inferential knowledge of the target proposition. My thesis is just that the FCP hypothesis gives rise to serious difficulties for interpreting such cases as cases of knowledge from falsehood. And it is compatible with the thesis that the subject in such cases knows the target belief.

Second, admittedly, if the arguments in this paper are successful, the inferences in play in the alleged cases of knowledge from falsehood must contain no false crucial premise, provided that the subject in such cases does know the target proposition *via* this inference. Naturally, this leads to a series of questions: what is the structure of the epistemising inference in play? How should we identify the epistemising inferences? Given that the inference in play in such cases *appears* to contain a false crucial premise, we will have to explain this phenomenon away. But how? These are certainly important questions that must be answered by any complete account. But since my goal here is just to highlight a serious problem of the FCP hypothesis, I will not tackle the question of how (if at all) the subject acquires knowledge in the alleged cases of knowledge

---

[11] A crucial premise may still be redundant. Consider the argument from (p & q) and (p → r) to r. Although both premises are crucial, the first premise, i.e. (p & q), is in a sense redundant, since what is needed, in order to draw the conclusion, is not the whole premise but a part of it, i.e. p. Put differently, what is really "crucial" to this argument is not the first premise as a whole, but a part of it. In this paper, I assume that cases of knowledge from falsehood are such that the false crucial premise in play is not redundant in the sense that it contains a true part, which is solely responsible for epistemising the target knowledge.

from falsehood. Nor will I propose a positive account of the purportedly epistemising inference in such cases.

Third, although I have argued that the inference in play in *Handout* can be reconstructed as P1-C1, I do not mean that there is only one possible way to reconstruct the inference. Nor do I contend that P1-C1 is *the* correct reconstruction. The following, for instance, appears to be an equally plausible reconstruction of the inference in play:

P5   There are 53 people at my talk.

P6   100 handout copies are more than enough for 53 people.

P7   If P5 and P6, then my 100 handout copies are sufficient for the people at my talk.

C3   Therefore, my 100 handout copies are sufficient for the people at my talk.

Fortunately, we are not forced to choose between P1-C1 and P5-C3. Nor, for that matter, are we forced to choose between P1-C1 and any other plausible reconstruction. The bottom line is that *Handout* is most naturally and charitably read as involving a deductively valid inference, or at any rate, an inference whose conclusion is inferred *via* a deductive step. If so, no matter how we reconstruct the inference in play, the argument above will still hold.

4. Strong Inductive Inferences with a False Crucial Premise

Some might argue that the inference involved in *Handout* should be reconstructed as an *inductive* (non-deductive) inference rather than a deductive one. The aforementioned argument did not hold, the critics claimed, because the inference in play had been formulated incorrectly.

It is worth noting that in *Handout*—in fact, all alleged cases of knowledge from falsehood proposed by Warfield—the inference that the subject explicitly carries out does appear to be a deductive inference. This is why inferences such as P1-C1 are very plausible reconstructions in the first place. But admittedly, an inductive step can also be added into the reconstruction, as the false belief in play is arguably acquired through an inductive inference. So, to incorporate this feature, one might reconstruct the subject's inference in *Handout* as follows:

P8   I have counted that there are 53 people at my talk.

P9   There are 53 people at my talk.

P10   If P9, then my 100 handout copies are sufficient for the people at my talk.

C4   Therefore, my 100 handout copies are sufficient for the people at my talk.

Suppose that P9 is inferred from P8 by induction. The inference P8-C4 thus has an inductive flavor in it. But even so, the aforementioned argument is not sabotaged, for in P8-C4, C4 is still *deductively* inferred in part from the false belief P9.

A related and more promising suggestion is to deny that cases of knowledge from falsehood can only arise from deductive inferences. On this line of suggestion, cases of knowledge from falsehood involving inductive inferences could also be constructed. If so, our critics continued, the aforementioned argument would have no force against such cases of knowledge from falsehood.

The critics are right that a seeming case of knowledge from falsehood may just involve an inductive inference. While the cases originally proposed by Warfield do characterise the subject

as explicitly performing a deductive inference, Peter Klein has proposed a number of alleged cases of knowledge from falsehood in which the inferences involved are inductive. Consider, for example, the following:

> *Appointment.* On the basis of my apparent memory, I believe that my secretary told me on Friday that I have an appointment on Monday with a student. From that belief, I infer that I do have an appointment on Monday. Suppose, further, that I do have an appointment on Monday, and that my secretary told me so. But she told me that on Thursday, not on Friday.[12]

Intuitively, the subject knows that she has an appointment on Monday, and this belief appears to be inferred from the false belief that the secretary told the subject on Friday that she has an appointment on Monday. More important for the present discussion, the inference involved is most naturally construed as an inference from testimony, which is an inductive inference. For the sake of argument, let us assume that the inference in play is inductively strong. The following is a plausible reconstruction of the subject's inference:

> P11   My secretary told me on Friday that I have an appointment on Monday.

> C5   Therefore, I have an appointment on Monday.

Let us make two remarks. First, the reconstruction respects the FCP hypothesis; the inference will not go through without the false belief P11. Second, the inference P11-C5 appears to be in line with the explicit inference that the subject carries out in *Appointment*. Of course, as noted, a given inference may have several equally plausible reconstructions. So, perhaps one wants to add the following premise into the reconstruction and has P11 to be inferred from this new premise:

> P12   I remember that my secretary told me on Friday that I have an appointment on Monday.

Reconstructing the inference in this way will not affect the points I want to make below. For the bottom line is that any plausible reconstruction of the subject's inference in *Appointment* should contain the step of inductively inferring C5 from P11 or similar premises, and it is this inferential step with which I am concerned.

The FCP hypothesis in combination with the assumption that the subject's inference in *Appointment* is inductively strong gives rise to the following analysis:

> A2   In *Appointment*, the inference in play is a strong inductive inference, which contains a false belief as one of its crucial premises.

But A2 leads to a serious problem for those who want to interpret *Appointment* as a case of knowledge from falsehood.

The problem is that a strong inductive inference with a false crucial premise, like its deductive counterpart, is not relevant to the truth of its conclusion. A strong inductive inference with a false crucial premise neither increases nor decreases the likelihood of its conclusion. Likewise, with one of its crucial premises being false, a strong inductive inference is no more effective in substantiating its conclusion than a weak inductive inference is.

---

[12] Klein (2008: 36).

Now, as noted, an inference is unable to epistemise its conclusion, if it is not truth-relevant to its conclusion. It follows that a strong inductive inference with a false crucial premise, by itself, cannot epistemise its conclusion. Of course, the conclusion of an inductive inference containing a false premise may still be known if the false premise involved is redundant and/or if the conclusion is self-evident or substantiated by some other means. But these are beside the point, as we are concerned with whether or not an inductive inference with a false *crucial* premise *by itself* is in a position to epistemise its conclusion.

In brief, an inductively strong premise with a false crucial premise, by itself, cannot epistemise its conclusion. To illustrate, consider, as an example, the following inference:

P13   My professor told me that phlogiston exists.

C6   Therefore, there are some falsehoods in modern chemistry.

Suppose that the inference from P13 to C6 is inductively strong (e.g., my professor is very knowledgeable and honest) and that P13, a crucial premise of P13-C6, is indeed false (e.g., I have hallucinated that my professor told me so). It is clear that the inference P13-C6, by itself, does not substantiate C6. As a result, the inference P13-C6 cannot epistemise C6—there is no way one can come to know C6 simply *via* P13-C6, not even if C6 is true and one is justified in believing P13.

Now, as noted, when *Appointment* is interpreted along the line of A2, the inference in play can be reconstructed as P11-C5, which, though inductively strong, contains the false belief P11 as one of its crucial premises. Hence, just as the inference P13-C6, by itself, fails to epistemise C6, the inference P11-C5, by itself, also fails to epistemise C5. But if so, then *Appointment* is not a case of knowledge from falsehood, as it is not a case of knowledge in the first place. Since A2 follows from the FCP hypothesis, this shows that, given this hypothesis, it is implausible to interpret cases such as *Appointment* as cases of knowledge from falsehood.

5. Possible Objections and Replies

The aforementioned arguments depend on two theses:

(a)   An inference is in a position to epistemise its conclusion only if it is truth-relevant.

(b)   Neither a valid deductive inference with a false crucial premise nor a strong inductive inference with a false crucial premise are truth-relevant.

Some might turn the tables on me and claim that the point of cases such as *Handout* and *Appointment* was precisely that (a) and/or (b) is false, namely, either non-truth-relevant inferences might still epistemise its conclusion, and/or even with a false crucial premise, a valid deductive inference or a strong inductive inference can still be truth-relevant. In fact, some might even claim that the proposal of knowledge from falsehood should be more charitably interpreted as a denial of (a) and/or (b). If so, the critics continued, I had begged the question against the proponents of knowledge from falsehood by accusing them of mistakenly attributing knowledge based on (a) and/or (b).

To reply, let us first note that (a) and (b) are supported by certain initially plausible considerations. As regards (a), the requirement that inferential knowledge must be based on a certain truth-relevant inference is derived from, or is a special case of, the more general requirement that justification, or warrant, required by knowledge must consist (at least in part) in

truth-relevant factors—such as the strength of one's evidence or reasons, the degree of reliability of the belief-forming process, etc.

The truth-relevant requirement of knowledge is overwhelmingly plausible and constitutes an essential part of the ordinary notion of knowledge. A brief survey of the literature should show that this view has been very widely accepted. Moreover, the assumption that knowledge requires truth-relevant factors is equally endorsed by non-orthodox views such as *pragmatic encroachment*.[13] Proponents of pragmatic encroachment deny that truth-relevant factors are the only factors that play a significant epistemic role, but they do not deny that truth-relevant factors play an essential role in turning true beliefs into knowledge. In addition, the idea that truth-relevant factors play an essential role in determining knowledge fits perfectly well with the ordinary notion of knowledge. I find it incomprehensible to think about the notion of knowledge without the requirement of the target belief being based (at least in part) on truth-relevant factors, in as much the same way as I find it incomprehensible to think about the notion of knowledge without the requirement of the target belief being true.

As regards (b), the idea that such inferences are not truth-relevant is derived from, or supported by, an essential feature of inferences. That is, in essence, the inferential relationship between a set of premises and a conclusion *by itself* has no bearing on the truth of the conclusion, no matter how good, logically speaking, the inferences are. Logically good inferences such as valid deductive inferences and strong inductive inferences are characteristically truth-preserving or leading to truth-preservation. But a truth-preserving, or approximately truth-preserving, inference *by itself* does not increase or decrease the likelihood of its conclusion. It is when paired with true premises, and only true premises, that a logically good inference substantiates or renders its conclusion more probable or less so. Put differently, as a belief-forming process, an inference is only *conditionally reliable*, and whether or not the belief-forming process is *reliable simpliciter* (i.e., whether or not all its output-beliefs, or a sufficient proportion of them, are true) depends on whether or not all of its input-beliefs are true.[14] In other words, a conditionally reliable belief-forming process may nonetheless be unreliable *simpliciter* if some of its input-beliefs are false. It is because of this essential feature of inferences that neither a valid deductive inference with a false crucial premise nor a strong inductive inference with a false crucial premise is truth-relevant.

All in all, since (a) and (b) are individually plausible and well supported by principled reasons, if the proposal of knowledge from falsehood is to be construed as clashing with (a) and/or (b), we should see the clash as a *reductio* against knowledge from falsehood.

6. Conclusion

At its core, the aforementioned argument consists of two steps. First, I have shown that, given that the false belief in question in cases such as *Handout* and *Appointment* is a crucial premise of the inference in play (i.e., the FCP hypothesis), the inference in play must be either a valid deductive inference with a false crucial premise or a strong inductive inference with a false crucial premise. Second, I have argued that these two kinds of inferences, by themselves, are not in a position to

---

[13] See Fantl and McGrath (2009).
[14] Goldman (1979).

epistemise their conclusions. Hence, assuming the FCP hypothesis, cases such as *Handout* and *Appointment* are not cases of knowledge from falsehood.

Department of Philosophy

National Chung Cheng University

No.168, Sec. 1, University Rd.,

Minhsiung, Chiayi 62102,

Taiwan (R.O.C.)

kokyonglee.mu@gmail.com

References

Arnold, A. (2013). "Some Evidence Is False". *Australasian Journal of Philosophy* 91: 165–72.

Ball, B., and M. Blome-Tillmann. (2014). "Counter Closure and Knowledge Despite Falsehood". *The Philosophical Quarterly* 64: 552–68.

Buford, C., and C. M. Cloos. (2018). "A Dilemma for the Knowledge Despite Falsehood Strategy". *Episteme* 15: 166–82.

DeRose, K. (2009). *The Case for Contextualism*. Oxford: Clarendon Press.

Fantl, J., and M. McGrath. (2009). *Knowledge in an Uncertain World*. Oxford: Oxford University Press.

Goldman, A. I. (1979). "What Is Justified Belief?". In G. S. Pappas (eds.), *Justification and Knowledge: New Studies in Epistemology*. Boston: D. Reidel Publishing Company, pp. 1-23.

Hiller, A. (2013). "Knowledge Essentially Based upon False Belief". *Logos and Episteme* 4: 7–19.

Klein, P. (2008). "Useful False Beliefs". In Q. Smith (eds.), *Epistemology: New Essays*. Oxford: Oxford University Press, pp. 25–61.

Littlejohn, C. (2013). "No Evidence Is False". *Acta Analytica* 28: 145–59.

Luzzi, F. (2019). *Knowledge from Non-Knowledge: Inference, Testimony and Memory*. New York: Cambridge University Press.

Montminy, M. (2014). "Knowledge Despite Falsehood". *Canadian Journal of Philosophy* 44: 463–75.

Schnee, I. (2015). "There Is No Knowledge from Falsehood". *Episteme: A Journal of Social Epistemology* 12: 53–74.

Stanley, J. (2005). *Knowledge and Practical Interests*. Oxford: Clarendon Press.

Warfield, T. (2005). "Knowledge from Falsehood". *Philosophical Perspectives* 19: 405–16.

Williamson, T. (2000). *Knowledge and Its Limits*. Oxford: Clarendon Press.