Knowledge from Falsehood and Truth-Closeness

Sven Bernecker[1,2]

Abstract
The paper makes two points. First, any theory of knowledge must explain the difference between cases of knowledge from falsehood (provided there are any) and Gettier cases where the subject relies on reasoning from falsehood. Second, the closeness-to-the-truth approach to explaining the difference between knowledge-yielding and knowledge-suppressing falsehoods does not hold up to scrutiny.

Keywords Knowledge · True belief · Gettier · Closeness · False lemma · Peter Baumann · Risto Hilpinen

Prima facie, one can come to know something through inference from a false premise. On the standard interpretation, what looks to be inferential knowledge obtained from a falsehood is instead inferential knowledge derived from a known proposition. One acquires knowledge *despite* the presence of a falsehood. A number of epistemologists have recently argued that apart from *knowledge despite falsehood* there is genuine *knowledge from falsehood*, where this is understood to require that the falsehood plays an important role in the inference-based production of knowledge.

I do not argue for the possibility of knowledge from falsehood. Instead I assume that knowledge from falsehood is possible and explore a follow-up issue: what (if anything) distinguishes cases of knowledge from falsehood from Gettier cases where the subject relies on reasoning from falsehood? I dub this the *false lemma problem*. The paper makes two points. First, it is a criterion of adequacy for any theory of knowledge that it explains the difference between knowledge from falsehood and falsehood-involving Gettier cases. Second, the explanation of this difference in terms of semantic or epistemic truth-closeness does not hold up to scrutiny.

Section 1 distinguishes between knowledge from falsehood and knowledge despite falsehood. Section 2 presents the false lemma problem. Section 3 argues that

✉ Sven Bernecker
    s.bernecker@uci.edu

1   Department of Philosophy, University of Cologne, Cologne, Germany

2   Department of Philosophy, University of California, Irvine, CA 92697, USA

the problem cannot be explained away. Sections 4 and 5 critically discuss the semantic and the epistemic version of the closeness-to-the-truth approach to the false lemma problem. Section 6 contains some concluding remarks.

1 Knowledge from Falsehood Vs. Knowledge despite Falsehood

The received wisdom in epistemology has it that knowledge-yielding competent deductive inference must issue from known premises. This principle has been dubbed *counter-closure* (Luzzi, 2010). If knowledge is justified non-gettierized true belief, as many epistemologists hold, counter-closure entails that to know a proposition via deductive inference each premise must be true, justifiedly believed, and non-gettierized. This paper focuses on challenges to the truth-condition of counter-closure.

Warfield's (2005: 407-8) handout case and similar cases are meant to show that there is knowledge from falsehood.[1]

> HANDOUT. Counting with some care the number of people present at his talk, Ted reasons: 'There are 53 people at my talk; therefore, *my 100 handout copies are sufficient*'. Ted's premise is false. There are 52 people in attendance -- he double counted one person who changed seats during the count. Does he know that 100 handout copies are sufficient?

Cases like this one are in no way unusual. We frequently rely on poor counting, rough calculation, and inexact measuring. Even though we are to some extent aware of the fact that these methods are imprecise, we derive precise conclusions from them. To deny that Ted knows that 100 handout copies are sufficient would make knowledge a rare commodity. Thus, Warfield concludes "that relevant falsehoods sometimes play a central epistemizing role in inference."[2] In recent years, a number of epistemologists have defended the possibility of knowledge from falsehood.[3]

Critics of the possibility of knowledge from falsehood argue that what superficially looks to be inferential knowledge obtained from *falsehood* is really inferential knowledge derived from *true belief*. In *Handout*, there is a true proxy belief that is less precise than the false premise-belief that does the epistemic work. The idea is that Ted forms not just *one* but *two* beliefs when he counts the people present at his

---

[1] For the history of this topic before the publication of Warfield (2005), see De Almeida (2019).

[2] Warfield (2005: 412). I take 'epistemization' to mean 'justification.'

[3] See Arnold (2013), Balcerak Jackson and Balcerak Jackson (2013), Bernecker and Grundmann (2019), Coffman (2008), De Almeida (2017), Fitelson (2010, 2017), Goldberg (2001), Hawthorne (2004: 57), Hilpinen (1988: 163-4), Klein (2008), Luzzi (2010, 2014, 2019), Murphy (2013), Sorensen (2016), and Turri (2011): 8; 2012: 217; Turri, 2019). Elgin (2019) holds that approximations and idealizations can give rise to scientific knowledge (and understanding) even though approximations and idealizations are literally false. For idealization-based belief to qualify as knowledge from falsehood, we have to assume that idealizations are outright falsehoods (as opposed to approximate truths) and that scientists are not cognizant of the fact that the idealizations they use are literally false. Both assumptions have been challenged by Sorensen (2013).

talk: the *explicit* belief that there are 53 people in the audience and the *implicit* belief that there are *approximately* 53 people in the audience. He then employs the latter belief in combination with his background knowledge that $53 < 100$ to conclude that his 100 handouts are enough. On this reading of the case, the false belief that there 53 people in the audience is inferentially inert. The false belief may figure in the *causal* production of the conclusion-belief and it may *contribute* to the justificatory status of the conclusion-belief but its contribution does not play an *essential* role. Given the proxy-premise strategy, *Handout* is a case of *knowledge despite falsehood* (*KDF* for short), as opposed to a case of *knowledge from falsehood* (*KFF* for short).[4]

The possibility of KDF is uncontroversial.[5] Suppose I base my belief that p on a dozen good reasons, but only eleven of them are true. If the eleven reasons are strong enough to justify p, then I know that p even though one of my reasons is false.

The controversial thesis states that there are KFF cases that are not reducible to instances of KDF. As was mentioned above, I assume the possibility of KFF and focus on a follow-up question: what distinguishes KFF cases (provided there are any) from Gettier cases where the subject relies on a reasoning from falsehood?

2 The False Lemma Problem

Classical Gettier cases involve reasoning from falsehood. Consider, for example, Lehrer's (1965) well-known Ford case, which is in the spirit of Gettier's (1963) original job/coin case.[6]

> FORD. Smith has reasons to believe that Nogot owns a Ford and that Nogot works in his office. From this Smith infers that *someone in his office owns a Ford* (*p*). As it turns out, Nogot does not own a Ford but *p* is true. Unsuspected by Smith, there is another person in the office, Havit, who owns a Ford.

There is near universal agreement that Smith does *not* know that someone in the office owns a Ford. Opinions differ, however, as to *why* he does not know. Some (including Gettier) claim that Smith does not know because he essentially relies on a false premise.[7] However, this explanation is inconclusive because there are classical Gettier cases and Gettier-like failed threat cases that manage without false

---

[4] Among the proponents of the proxy-premise reading of (alleged) cases of KFF are Ball and Blome-Tillmann (2014), Lee (2021), and Schnee (2015). Buford and Closs (2018) provide a dilemma for those wanting to reject the possibility of KFF. Borges (2017: 286-9) and Montminy (2014: 466) argue that Ted's true belief that 100 handout copies are enough depends neither evidentially nor causally on the false belief that there are 53 people in attendance. Audi (2013: 514-5) constructs a case of testimonial KDF while Goldberg (2001: 516) gives a case of testimonial KFF. Bernecker and Grundmann (2019) argue for the possibility of memory KFF and Balcerak Jackson and Balcerak Jackson (2013) argue for the possibility of knowledge from false suppositions.

[5] The possibility of KDF is acknowledged by Goldman (1967: 368), Lehrer (1965: 169-71), Saunders & Champawat (1964: 9), and Swain (1981: 149-50).

[6] Unlike Gettier's job/coin case, this case does not rest on a confusion of the referential and the attributive sense of the definite description 'the man who will get the job.' See Biro (2017).

[7] See Clark (1963), Gettier (1963: 122), and Harman (1973: 195).

premises.[8]Consider, for example, the well-known failed threat case due to Ginet and Goldman (1976: 772-3).

> BARN. Henry is in an area where, unbeknownst to him, the inhabitants have erected a large number of fake barns. Fake barns are single-sided façades painted to look like barns. There is, though, one real barn among all the fakes, and Henry happens to be looking at it. Henry's looking at the barn causes him to believe that *there is a barn*.

Barn and *Ford* have in common that the respective beliefs do not qualify as knowledge because their truth is a result of luck. In many close possible worlds, the respective beliefs are false. This is why Gettier cases and failed threat cases are often lumped together. But there is an important difference between them: Smith's belief that someone in the office owns a Ford rests on a false premise. Henry, on the other hand, does not seem to rely (not even tacitly) on a false premise in coming to believe that there is a barn. He simply forms a belief on the basis of what his perceptual faculties present to him.[9]

Even though there are classical Gettier cases and Gettier-like failed threat cases that manage without false lemmas, many such cases, including Gettier's original ones, do involve reasoning from falsehood. And if we assume that falsehoods can play an *essential* role in the reasoning process giving rise to knowledge, we are left with the question of when the essential reliance on falsehoods leads to knowledge and when not. What distinguishes KFF cases from falsehood involving Gettier cases (*FIG cases* for short)? What distinguishes knowledge-yielding from knowledge-suppressing falsehoods in reasoning? This is the *false lemma problem*.

3 Dissolutions to the False Lemma Problem

A critic might try to explain away the false lemma problem by collapsing the FIG/KFF distinction. Yet reducing KFF cases to FIG cases is not an option because we are working from the assumption that there are genuine KFF cases. Reducing FIG to KFF is also not an option because then any true belief, no matter how unreliably and irrationally formed, qualifies as knowledge. And this view is clearly too implausible to deserve serious consideration.

Another strategy for rendering the false lemma problem a pseudo-problem is to acknowledge the distinction between FIG cases and KFF cases but to maintain that it is either not possible or not necessary to know the basis of the distinction.

---

[8] For classical Gettier cases without false premises, see Almeder (1973), Feldman (1974), Lehrer (1965: 170), and Swain (1972: 429-30). In light of these cases, the no-false-premise account of knowledge has been transformed into the no-essential-false-assumption account. See Feldman (2003: 37), Harman (1973: 46-50, 120-4), Lehrer (1974: 219-20), Levin (2006), and Lycan (2006: 156-7).

[9] Goldman (1976: 771), Levin (2006: 390), and Lycan (2006: 157-8) claim that Henry's barn-belief is non-inferential. Harman (1980: 176) disagrees. He claims that Henry assumes that it is unlikely that his belief is false. Since this assumption is false, he violates the no-essential-false-assumption principle and thus does not know.

A possible reason for thinking that the difference between the FIG/KFF distinction is unknowable is that the terms 'FIG' and 'KFF' have borderline cases. But even if 'FIG' and 'KFF' were vague terms, it would not follow that the basis of the FIG/KFF distinction is unknowable. To see this, consider the paradigmatically vague predicate 'heap.' No amount of conceptual analysis or empirical inquiry can settle whether adding one grain of sand atop a single layer produces a heap, and removing the last grain above the bottom layer destroys the heap. But while it is intrinsically uncertain how many grains of sand make a heap, we know which properties distinguish a heap from a non-heap, namely the number of grains of sand and how they are positioned. Which properties are the difference-makers is known; what is not known is the quantity or degree of these properties needed to make a difference. If the FIG/KFF distinction were like the heap/non-heap distinction, we should be able to identify the difference-makers.

One can also try to resolve the false lemma problem by declaring the explanation of the FIG/KFF difference superfluous. The idea is that one only needs to be able to distinguish FIG cases from KFF cases. It is not necessary that one should also be able to explain the difference and say on what basis one makes the distinction. As long as one can tell FIG cases from KFF cases in the same way, say, a chicken-sexer can tell male from female chicks, we do not need insight into the basis of one's discriminatory ability.[10]

The FIG/KFF distinction is conceptual in nature while the distinction between a male and a female chick is biological. Telling male from female chicks has a practical (economic) use that is largely independent of the ability to explain the principle for telling which is which. In contrast, the FIG/KFF distinction has merely theoretical or explanatory use. There is a clear explanatory advantage to be able to say not only *that* a given case is an instance of FIG or KFF, respectively, but also *why* it is. The explanatory advantage consists in a deeper understanding of the nature of knowledge. We know not only whether something is an instance of knowledge but also what makes it an instance of knowledge. The latter kind of knowledge gives us predictive powers and it allow us to intentionally grow our body of knowledge.

A theory of knowledge that allows us to sort FIG from KFF but that does not explain the principle behind the sorting meets the constraint of extensional adequacy but not that of explanatory adequacy. A successful theory of knowledge, however, is both extensionally and explanationally adequate. In other words, a successful theory of knowledge not only identifies all and only cases of knowledge but it also sheds light on the nature of knowledge. As theorists of knowledge we want to understand how we know the things we think we know (Stroud, 2002: 106). And we cannot satisfy ourselves on that score unless we can see ourselves as having good reasons for sorting FIG cases from KFF cases. It is a demand of the epistemological project itself that we seek to explain the principle behind the sorting of cases.

---

[10] Armstrong's (1963: 431-2) chicken-sexer example is used by Foley (1987: 168-170) to show that reliability is insufficient for knowledge. See also BonJour's (1980) clairvoyance case and Sosa's (2010: 88-9) blindsight case. According to Sosa (2009), the chicken sexer has animal knowledge that something is, say, a KFF case, but he lacks reflective knowledge about the FIG/KFF distinction.

4 Semantic Truth-Closeness

There are three kinds of explanations of the difference between knowledge-yielding and knowledge-suppressing falsehoods. *Defeater-based accounts* maintain that reasoning from a false premise yields knowledge if either the negation of the false premise is not a defeater for the true conclusion or the true conclusion is indefeasibly justified by a truth that is entailed by the false premise.[11]*Safety accounts* have it that reasoning from a false premise generates knowledge if the inferential path from the false premise to the true conclusion is modally stable – if it could not have easily given rise to a false conclusion.[12]Finally, *closeness-to-the-truth accounts* hold that reasoning from a false premise yields knowledge provided the false premise is close to the truth.

The closeness-to-the-truth approach to the false lemma problem comes in two flavors. There is a semantic version suggested (but not endorsed) by Hilpinen (1988) and an epistemic version put forth by Baumann (2020). I will start with the semantic closeness-to-the-truth account before tackling the epistemic version.

> Hilpinen (1988: 163) gives the following example of knowledge based on a false premise.
> THERMOMETER. A mother takes her son's temperature and reads the thermometer as 40 degrees Celsius. She comes to know that her son has fever because a temperature of greater than 37 degrees suffices for a fever. But the thermometer is not completely accurate (few ordinary thermometers are) and the son's temperature is actually 39.2 degrees.

The suggestion is that the mother knows that her son has a fever even though she reasons from a false premise. The example is said to illustrate that "perhaps relatively vague knowledge-claims can sometimes be justified by 'sharp' (or informative) but false beliefs which are reasonably close to the truth" (Hilpinen, 1988: 164). Hilpinen does not take himself to make a general point about KFF, but only to describe an instance of it that serves as a counterexample to the *sensitivity* condition of knowledge.[13]A belief that *p* is sensitive if and only if *p* were false and if one were to use the same method as in the actual world to arrive at a belief as to whether *p*, then one would not believe that *p* (Dretske, 1971; Nozick, 1981). The thermometer case challenges the necessity of sensitivity because there is a close possible world where the son does not have a fever -- his temperature is 36.9 -- even though the mother believes that he does have a fever because she relies on the inaccurate thermometer which reads '37.7'. Hilpinen concludes that.

---

[11] See De Almeida (2017), Feit and Cullison (2011), and Klein (2008). For a critical discussion of defeater-based accounts, see Bernecker (forthcoming).

[12] See Grundmann (2020: 5179), Luzzi (2019: 30, 70-1), and Warfield (2005: 414).

[13] The closest Hilpinen (2017: 148) comes to making a general comment about the difference between FIG and KFF is when he writes that "[i]n Gettier-type examples the inquirer's false belief falls short of epistemic perfection in a way that cannot be ignored."

a person can know things not only on the basis of (valid) inference from what he or she knows, but in some cases even on the basis of inference from what is *not* known (or even true) provided that the latter (evidential) propositions are sufficiently close to the truth (Hilpinen, 1988: 164).

The idea that propositions can be closer or further apart is highly intuitive. For instance, 'Max has a temperature of 40 degrees' and 'Max has a temperature of 37.7 degrees' seem to be closer than 'Maria has a temperature of 40 degrees' and 'Jill has a temperature of 36 degrees.' Many philosophers, including Hilpinen (1976), follow Lewis (1973) and Stalnaker (1968) in analyzing closeness of truth-conditions in terms of possible worlds. Proposition *p* is closer to the truth than proposition *q* if *p* is true in worlds that are closer to the actual one than the worlds in which *q* is true.[14]

The proposal under consideration states that a false lemma is knowledge-yielding if it is true in *close* possible worlds, and knowledge-suppressing otherwise. The difference between a FIG case such as *Ford* and a KFF case such as *Handout* is that the world where Nogot owns a Ford is further from actuality than the world where 52 people attend Ted's talk. Things have to go quite a bit differently than they actually do in the former case, but not that differently in the latter case.

A notorious problem of the possible-worlds framework is that there is no unique and non-ad hoc closeness metric for possible words (Baumann, 2009, 2016: 60–3; Zalabardo, 2012: 111–14). Lewis, who championed this framework, holds that the worlds most similar to actuality are the ones that have the same natural laws and initial conditions as the actual world. Yet he notes that precise similarity of fact across significant spatiotemporal regions can outweigh nomological similarity, provided it can be achieved as the cost of only small, localized miracles (Lewis, 1979: 466–7). In the end, he concedes that the closeness of worlds is context-dependent.[15] A world where, say, I am a brain in a vat is distant from the point of view of epistemology, but close from the point of view of physics (and metaphysics).

The Hilpinen-inspired distinction between FIG/KFF can handle *Handout* and *Ford*, but it has problems with other cases. Two drive this point home, consider two closeness measures for possible worlds – a subjective and an objective one. On the *subjective* metric, closeness is measured in terms of psychological distinguishability for the epistemic subject – the closer the world the less distinguishable from

---

[14] Hilpinen (1976) defends a possible-world approach to truthlikeness (verisimilitude) – a key concept in Popper's anti-inductivist philosophy of science. Science aims at producing true theories. But many past scientific theories have turned out to be false. It is therefore very likely that currently accepted theories will also turn out to be false. How can it be rational to pursue an unattainable goal? How can there be scientific progress under these circumstances? A possible answer is to scale down the goal of science. Science aims at developing theories, which approximate more closely to the truth. Popper (1963) defines verisimilitude in terms of logical consequence: theory A has greater verisimilitude than theory B if theory A has more true consequences and fewer false ones than theory B. For the discussion of the notion of verisimilitude in philosophy of science, see Oddie (2014).

[15] Lewis (1973: 50-2, 66-7, 91-5). See also Heller (1999: 505-7). According to Lewis' contextualism (Lewis, 1996: 557), Smith in *Ford* does not know that someone in the office own a Ford because he violates the rule of actuality and the rule of resemblance. The rule of actuality says that whatever is actual cannot be properly ignored in our search for knowledge. The rule of resemblance says that, of two very similar possibilities that saliently resemble each other, both should be either considered or rejected.

actuality. Putnam's (1973) Twin Earth, for example, is a close world in this respect. On the *objective* metric, it is the relative frequency of the actual occurrence of an event that determines whether a world is close.

First, consider a variation of the Ford case where, unbeknownst to Nogot, a distant relative recently bequeathed him a Ford. The transfer of ownership is delayed by a few hours. There is a possible world where paperwork has already been processed by the time Smith forms the belief that Nogot owns a Ford. The lemma is true. This world is *subjectively* close because, from the point of view of Smith and Nogot, it is psychologically indistinguishable from actuality. Given the Hilpinen-inspired proposal and a subjective closeness metric, Smith *knows* that someone in the office owns a Ford. However, we would still classify this as a Gettier case because the belief is true due to luck. The fact that the false lemma is true in subjectively close worlds is beside the point.

Next, consider a variation of the handout case. Suppose it is nothing short of a miracle that as many as 52 people show up for Ted's talk (the topic is esoteric, he is an uninspiring speaker, the talk is at an unpopular time of day, etc.). The world where an additional person shows up to the talk is *objectively* far from actuality. Intuitively, however, Ted still *knows* that his 100 handout copies are sufficient on the basis of having counted 53 people in the room. He knows because the margin of error is negligibly small.

The upshot is that regardless of whether we choose a subjective or an objective closeness metric, the semantic closeness-to-the-truth account of the FIG/KFF distinction is extensionally inadequate. Thus it seems that the FIG/KFF distinction is independent of the semantic truth-closeness of the false lemma.

As a last resort, the proponent of semantic truth-closeness could define closeness in terms of *numerical structure*. Given that the KFF cases discussed so far operate with consecutive integers (52/53 in *Handout*) and neighboring rational numbers (37/39.2/40 in *Thermometer*), it stands to reason to claim that the hallmark of KFF is that the false lemma is numerically close to the truth. In FIG cases, by contrast, the false lemma does not involve numerical structure and thus cannot be said to be close to the truth. The premise that Nogot owns a Ford is equidistant to the true proposition that someone in the office owns a Ford and to the false proposition that no one in the office owns a Ford.

The numerical elaboration of the semantic notion of closeness-to-the-truth fails for two reasons. First, there are KFF cases that manage without numerical structure. Consider the following case due to Turri (2012: 217).[16]

> DRESS. John awaits Monica's arrival. He wonders whether she will wear a scarlet dress. He hears a step on the staircase and swings around to see Monica enter the room. 'What a dazzling indigo dress!' he thinks and concludes,

---

[16] Non-numerical cases of KFF can also be found in Hiller (2013:10) and Klein (2008: 36). In Warfield's handout case, the reasoning is deductive; in non-numerical KFF cases, the reasoning is typically inductive.

'*Monica's dress is not scarlet.*' And he is right: her dress is not scarlet. But it is not indigo either. It is ultramarine.

This example does not involve numbers or countable nouns but, admittedly, it involves a sequence, namely the color chart, which allows us to 'measure' the relative distance of colors.

Second, not only are there KFF cases that manage without numerical structure, but also there are FIG cases that do involve such structure. Consider a variation on Gettier's (1963) job/coin case.

> COIN. Smith and Jones have applied for the same job. Smith has counted *ten* coins in Jones' pocket and has overheard the president of the company state that Jones would in the end be selected. Smith therefore believes that Jones gets the job and Jones has nine, ten or eleven coins in his pocket. From this he infers that *the man who gets the job has nine, ten or eleven coins in his pocket*. The belief is true but the reason it is true is that Smith gets the job and that Smith has *nine* coins in his pocket.

Clearly Smith does *not* know that the man who will get the job has nine, ten or eleven coins in his pocket. The fact that there is only a one-coin difference between the change in Smith's pocket and the change in Jones' pocket does not change the fact that it is a Gettier case. In sum, the semantic approaches to truth-closeness fail to distinguish FIG and KFF.

5 Epistemic Truth-Closeness

Semantic truth-closeness, as we have seen, consists in the similarity of truth conditions. Epistemic truth-closeness, on the other hand, is a property of belief-forming methods, not of propositions. A method is close to the truth if it actually yields a false belief but could have easily generated a true belief of the same kind. In this section, I will analyze Baumann's (2020) notion of epistemic truth-closeness (step 1) and critically discuss the FIG/KFF distinction based on this notion (step 2).

Step one. On the widespread reliabilist conception of knowledge, a true belief counts as knowledge if its mode of acquisition rules out all serious or 'relevant' alternatives in which the belief would be false. The safety theory is a popular version of reliabilism.[17] Baumann uses the safety theory to explicate the notion of epistemic truth-closeness.

In its bare-bone form, safety states that someone knows that *p* only if, in all close possible worlds where they believe that *p*, *p* is true. As it stands, the condition is too

---

[17] The safety theory was introduced by Sosa (1996: 276-7) and Williamson (1994: 230-4). S's belief that *p* formed via method *M* is safe if and only if in all close possible worlds in which S believes *p* via *M*, S's belief is true. A method that yields true beliefs in all close possible worlds is a reliable method, in a perfectly ordinary sense of 'reliable method' (Williamson, 2009: 306–8). Moreover, given that Williamson (2000: 101, 126) maintains that knowledge is safe belief, and that knowledge requires reliability, it follows that safe belief is reliable belief.

stringent. Typically, safety is relativized to belief-forming methods. A method is safe if, in all close possible worlds where S believes that *p based on the same method as in the actual world*, *p* is true. As it stands, the condition is deemed to be too weak. The standard way to strengthen method safety is to demand that a safe method yields true beliefs not only in the target proposition *p* but also in *propositions relevantly similar to p*.[18] The qualified safety condition reads:

> (Method Safety). A method used in the acquisition of a true belief that *p* is safe just in case use of the method in all (or most) close possible worlds leads to true beliefs of the same kind as the belief that *p* (Baumann, 2020: 1770).

Just as a method can safely or unsafely generate *true* beliefs, it can safely or unsafely generate *false* beliefs. A method generating a false belief that *p* is *inversely safe* if, in close possible worlds it leads to false beliefs of the same kind as the belief that *p*. An inversely safe method could not have easily generated a true belief. The opposite of inverse method safety is epistemic truth-closeness. A method giving rise to a false belief that *p* is close to the truth if it could have easily given rise to true beliefs in propositions relevantly similar to *p*. Baumann articulates the notion of epistemic truth-closeness as follows:

> (Safe Method-Closeness) A method used in the acquisition of a false belief that *p* is safely close in case use of the method in all (or most) close possible worlds leads to true belief of the same kind as the belief that *p* (Baumann, 2020: 1773).

Two clarificatory remarks. First, the formulation of safe method-closeness pre-supposes an unique closeness metric of possible worlds – the very metric which we saw cannot be had. Yet this problem can be circumvented by spelling out the reliabilistic idea not in terms of close possible worlds but in terms of probabilities. In fact, Baumann (2009, 2016: 58-63) advocates a probabilistic version of reliabilism whereupon a method *M* generating a true belief *T* is reliable if and only if the probability Pr (*T*/*M* happens) $> s$ (for $s > .5$), and Pr (*T*/*M* happens) $> Pr$ (*T*/*M* does not happen). Probabilistic reliabilism does not face the problem of the missing closeness metric for possible worlds.

Second, a token belief-forming method is reliable only relative to its being a token of a certain type. Since any token is a token of many types, which can differ in degree of reliability, the question arises: which method type determines the reliability of beliefs formed by its tokens? This is known as the *generality problem* (Conee & Feldman, 1998) and it affects modal and probabilistic versions of reliabilism alike. Baumann (2008: 29n4; Baumann, 2009: 83) argues that the generality problem cannot be solved. Yet this does not speak against the notion of safe method-closeness because the generality problem is not specific to reliabilism but afflicts every major epistemology, including internalism (cf. Bishop, 2010; Comesaña, 2006; Conee, 2013).

---

[18] For a critique of global method safety, see Bernecker (2020).

Step two. Building on the notion of safe method-closeness, Baumann develops an explanation of the FIG/KFF distinction. Before proposing his preferred explanation, Baumann rejects a simplistic proposal.

The simplistic proposal has it that a false lemma is knowledge-yielding if and only if the method giving rise to it is close to the truth and that it is knowledge-suppressing if and only if the method is far from the truth. This proposal will not do, for the condition of safe method-closeness is met not only in KFF cases but also in FIG cases. To see this, consider Gettier's (1963) Ford/Barcelona case: Smith falsely believes that Jones owns a Ford and infers from this the true disjunction that either Jones owns a Ford or Brown is in Barcelona but he has no reason to believe the second, true disjunct. Even though *Ford/Barcelona* is clearly a FIG case, the false first disjunct meets the condition of safe method-closeness. For

> if [Smith] had not even been close to the truth of the first disjunct (for instance, if he had just made a wild guess about Jones' Ford ownership) then we would not consider him to be justified and thus also not gettiered. We only do count Smith as justified and in a Gettier case because he was close to the truth. More precisely: Smith is in a Gettier case because he is method-close to the truth (Baumann, 2020: 1773-4).

The upshot is that safe method-closeness, by itself, is insufficient to differentiate between FIG and KFF.

Close method-safety, Baumann argues, is necessary but not sufficient for KFF. What sets KFF apart from FIG is that there is a less precise method available to the subject such that had they used this method they would have still arrived at the true conclusion. Baumann illustrates this point with the help of a variation on Warfield's handout case.

> [Ted] could have made rough but unrisky estimates (rather than counted) and reasoned, for instance, in the following way: 'There are at most 15 people over there, at most 20 over there, and, finally, certainly not more than 30 people over there; this is still way below 100; hence, I have a sufficient number of handouts.' In contrast to this, there is no such less precise method available to the subject in Gettier cases involving false premises (Baumann, 2020: 1774).

The hallmark of KFF, according to Baumann, is that the truth of the conclusion is doubly secured: the method giving rise to the false lemma is safely close *and* a less precise method is available to the subject for generating the true conclusion. On the flip side,

> [a] justified true belief based upon and dependent on a false premise is a *gettiered* belief only if the subject's belief in the false premise is method-close to the truth and there was no less precise method available to the subject (Baumann, 2020: 1774; my emphasis).

I have two questions for clarification and an objection.

Question one. The method of 'rough but unrisky estimates,' although less precise than counting, is equally accurate as counting in that it also meets the condition of safe method-closeness. Presumably, the reason Baumann focuses on less precise rather than more precise methods is that a more precise substitute for a safe method-close method tends to be more accurate. Yet it is surprising if a less precise method is equally accurate. But what about *equally precise* methods? Why should it not be indicative of KFF that the subject could have easily employed an equally precise safely close method to generate the true conclusion? It seems that Baumann could have extended the criterion of the FIG/KFF distinction to equally precise methods.

Question two. Baumann explains the difference between FIG and KFF in terms of the *availability* of a less precise method but he does not specify what it means for a method to be available to a subject. Are we talking about a subject's psychological or practical abilities? Are judgments about the availability of methods made from the perspective of the epistemic subject or from the perspective of an external observer? Moreover, given that vignettes are notoriously underdescribed, how can we ever be sure whether a less precise method is available to the subject?

Objection. According to Baumann, what sets KFF apart from FIG is that there is a less precise method available to the subject such that had they used this method they would have still arrived at the true conclusion. As it stands, the proposed criterion for identifying KFF is too weak. To drive this point home, I present two cases of non-knowledge that meet Baumann's criterion for KFF.

First, consider a variation of the Ford case. Instead of basing the belief that someone in the office owns a Ford on Nogot's testimony and existential generalization, Smith could have used perception and probabilistic reasoning. Suppose Smith can see from his office window that there are some Ford models in the company's parking lot and suppose he knows Ford's market share in vehicle sales. This method is equally or less precise as the method he actually employs. Moreover, the methods have a very similar accuracy. After all, the Ford models in the company's parking lot may not belong to the employees and the percentage of Ford ownership in the workforce may not correspond to the national average. Thus we have a FIG case where the false belief is actually formed via a safe method-close method and where there is an equally or less precise method available to the subject such that had they used it they would have still arrived at the true conclusion.

Next, consider a variation of the Gettier-like Barn case.[19] In the original case, Henry detects with the naked eye that there is a barn. In the modified case, Henry looks at regions in the country with a satellite. Each time he uses the satellite he has to select the viewing area. Since he wants to look at some farms, he picks what he thinks is an agricultural region in the Midwestern United States, region *R*, and then zooms in on *R* really close so that he sees only a small portion of *R* that has one barn-like structure. Based on the satellite's visual image Henry forms the belief that there is a barn in region *R*. The belief is true. However, if Henry had zoomed out just a little so that he had taken in slightly more of region *R*, he would have looked at a lot of fake barns. This method is less precise because zooming out provides less

---

visual detail of the structures in view. But use of this method, which is clearly available to Henry, would have still led him to a true belief. For region *R* contains one real barn.

In light of these counterexamples, Baumann might be tempted to strengthen the account of KFF by adding a further condition. Intuitively, counting and unrisky estimates (in the two versions of *Handout*) are more similar to one another than either testimony cum existential generalization and perception cum probabilistic reasoning (in the two versions of *Ford*) or unaided perception and satellite imagery (in the two version of *Barn*). Thus, Baumann might argue that what sets KFF apart from FIG is that (i) the method giving rise to the false lemma is safely close, (ii) a less precise method is available to the subject for generating the true conclusion, and (iii) *the less precise method is similar to the actually employed method.*

The crux of this response to the objection at hand is that, just as there is no unique closeness metric for possible worlds, there is no unique similarity metric for belief-forming methods. Similarity assessments are context-dependent and, in Baumann's (2009: 81) own words, "everything is similar in some respect to everything else." It depends on the context whether the (equally or) less precise method the subject employs in a close possible world is (sufficiently) similar to the actually employed method. And if the satisfaction of condition (iii) of KFF is context-dependent, so is the distinction between FIG and KFF. Yet this conclusion is hard to believe, for studies in experimental philosophy suggest (after initially claiming the opposite) that the Gettier intuition is stable across a number of contextual factors such as age, culture, gender, and language (Kim & Yuan, 2015; Machery et al., 2017; Nagel et al., 2013).

6 Conclusion

We saw that it is a criterion of adequacy for any theory of knowledge that it explains the difference between knowledge-yielding and knowledge-suppressing falsehoods in reasoning. According to the closeness-to-the truth approach, the difference has to do with the relative closeness of the falsehood to the truth. The semantic notion of truth-closeness faces the problem of the missing closeness metric for possible worlds. What is more, semantic truth-closeness, no matter how it is determined, is not what sets KFF apart from FIG. Baumann's epistemic account of truth-closeness fails because either it is either too weak to distinguish KFF from FIG or it faces the problem of the missing similarity metric for belief-forming methods. A positive account of the FIG/KFF distinction is the topic for another paper.[20]

References

Almeder, R. (1973). Defending Gettier's counterexamples. *Australasian Journal of Philosophy, 53*, 58–60.

Armstrong, D. M. (1963). Is introspective knowledge incorrigible? *Philosophical Review, 72*, 417–432.

Arnold, A. (2013). Some evidence is false. *Australasian Journal of Philosophy, 91*, 165–172.

Audi, R. (2013). Testimony as a Social Foundation of knowledge. *Philosophy and Phenomenological Research, 87*, 507–531.

Balcerak Jackson, M., & Balcerak Jackson, B. (2013). Reasoning as a source of justification. *Philosophical Studies, 164*, 113–126.

Ball, B., & Blome-Tillmann, M. (2014). Counter closure and knowledge despite falsehood. *The Philosophical Quarterly, 64*, 552–568.

Baumann, P. (2008). Is knowledge safe? *American Philosophical Quarterly, 45*, 19–30.

Baumann, P. (2009). Reliabilism – Modal, Probabilistic or Contextualist. *Grazer Philosophische Studien, 79*, 77–89.

Baumann, P. (2016). *Epistemic Contextualism: A Defense*. Oxford University Press.

Baumann, P. (2020). Close to the truth. *Philosophia, 48*, 1769–1775.

Bernecker, S. (2020). Against global method safety. *Synthese, 197*, 5101–5116.

Bernecker, S. (forthcoming). Knowledge, falsehood, and defeat. In R. Borges & I. Schnee (Eds.), *Illuminating errors: New essays on knowledge from non-knowledge*. Routledge.

Bernecker, S., & Grundmann, T. (2019). Knowledge from forgetting. *Philosophy and Phenomenological Research, 98*, 525–540.

Biro, J. (2017). Non-Pickwickian belief and 'the Gettier problem. *Logos and Episteme, 8*, 47–69.

Bishop, M. A. (2010). Why the generality problem is Everybody's problem. *Philosophical Studies, 151*, 285–298.

BonJour, L. (1980). Externalist theories of empirical knowledge. *Midwest Studies in Philosophy, 5*, 53–73.

Borges, R. (2017). Inferential knowledge and the Gettier conjecture. In R. Borges, C. De Almeida, & P. D. Klein (Eds.), *Explaining knowledge: New essays on the Gettier problem* (pp. 273–291). Oxford University Press.

Buford, C., & Cloos, M. (2018). A dilemma for the knowledge despite falsehood strategy. *Episteme, 15*, 166–182.

Clark, M. (1963). Knowledge and grounds. A comment on Mr. *Gettier's Paper. Analysis, 24*, 46–48.

Coffman, E. J. (2008). Warrant without truth? *Synthese, 162*, 173–194.

Comesaña, J. (2006). A well-founded solution to the generality problem. *Philosophical Studies, 129*, 27–47.

Conee, E. (2013). The specificity of the generality problem. *Philosophical Studies, 163*, 751–762.

Conee, E., & Feldman, R. (1998). The generality problem for Reliabilism. *Philosophical Studies, 89*, 1–29.

De Almeida, C. (2017). Knowledge, benign falsehoods, and the Gettier problem. In R. Borges, C. De Almeida, & P. D. Klein (Eds.), *Explaining knowledge: New essays on the Gettier problem* (pp. 292–311). Oxford University Press.

De Almeida, C. (2019). Review of F. Luzzi, *knowledge from non-knowledge* Cambridge University Press. *Notre Dame Philosophical Reviews* 2019.12.06.

Dretske, F. (1971). Conclusive reasons. *Australasian Journal of Philosophy, 49*, 1–22.

Elgin, C. Z. (2019). Epistemically useful falsehoods. In B. Fitelson, R. Borges, & C. Braden (Eds.), *Themes from Klein: Knowledge, Scepticism, and justification* (pp. 25–39). Springer.

Feit, N., & Cullison, A. (2011). When does falsehood preclude knowledge? *Pacific Philosophical Quarterly, 92*, 283–304.

Feldman, R. (1974). An alleged defect in Gettier examples. *Australasian Journal of Philosophy, 52*, 68–69.

Feldman, R. (2003). *Epistemology*. Prentice Hall.

Fitelson, B. (2010). Strengthening the case for knowledge from falsehood. *Analysis, 70*, 666–669.

Fitelson, B. (2017). Closure, counter-closure, and inferential knowledge. In R. Borges, C. De Almeida, & P. D. Klein (Eds.), *Explaining knowledge: New essays on the Gettier problem* (pp. 312–323). Oxford University Press.

Foley, R. (1987). *A theory of epistemic rationality*. Harvard University Press.

Gettier, E. (1963). Is justified true belief knowledge? *Analysis, 23*, 121–123.

Goldberg, S. C. (2001). Testimonially based knowledge from false testimony. *The Philosophical Quarterly, 51*, 512–526.

Goldman, A. I. (1967). A causal theory of knowing. *Journal of Philosophy, 64*, 357–372.

Goldman, A. I. (1976). Discrimination and perceptual knowledge. *Journal of Philosophy, 73*, 771–791.

Grundmann, T. (2020). Saving safety from counterexamples. *Synthese, 197*, 5161–5185.

Harman, G. (1973). *Thought*. Princeton University Press.

Harman, G. (1980). Reasoning and evidence one does not possess. *Midwest Studies in Philosophy, 5*, 163–182.

Hawthorne, J. (2004). *Knowledge and lotteries*. Oxford University Press.

Heller, M. (1999). The proper role for Contextualism in an anti-luck epistemology. *Philosophical Perspectives, 13*, 115–130.

Hiller, A. (2013). Knowledge essentially based upon false belief. *Logos and Episteme, 4*, 7–19.

Hilpinen, R. (1976). Approximate truth and Truthlikeness. In M. Przelecki, K. Szaniawski, & R. Wójcicki (Eds.), *Formal methods in the methodology of the empirical sciences* (pp. 19–42). Reidel.

Hilpinen, R. (1988). Knowledge and conditionals. *Philosophical Perspectives, 2*, 157–182.

Hilpinen, R. (2017). Sed ubi Socrates currit? On the Gettier problem before Gettier. The lesson of Gettier. In R. Borges, C. De Almeida, & P. D. Klein (Eds.), *Explaining knowledge: New essays on the Gettier problem* (pp. 135–151). Oxford University Press.

Kim, M., & Yuan, Y. (2015). No cross-cultural differences in the Gettier Car case intuition: A replication study of Weinberg et al. 2001. *Episteme, 12*, 355–361.

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

Lee, K. Y. (2021). Reconsidering the alleged cases of knowledge from falsehood. *Philosophical Investigations, 41*, 151–162.

Lehrer, K. (1965). Knowledge, truth and evidence. *Analysis, 25*, 168–175.

Lehrer, K. (1974). *Knowledge*. Oxford University Press.

Levin, M. (2006). Gettier cases without false lemmas? *Erkenntnis, 64*, 381–392.

Lewis, D. K. (1973). *Counterfactuals*. Blackwell.

Lewis, D. (1979). Counterfactual dependence and Time's arrow. *Nous, 13*, 455–476.

Lewis, D. K. (1996). Elusive knowledge. *Australasian Journal of Philosophy, 74*, 549–556.

Luzzi, F. (2010). Counter-closure. *Australasian Journal of Philosophy, 88*, 673–683.

Luzzi, F. (2014). What does knowledge-yielding deduction require from its premises? *Episteme, 11*, 261–275.

Luzzi, F. (2019). *Knowledge from non-knowledge: Inference, testimony and memory*. Cambridge University Press.

Lycan, W. G. (2006). On the Gettier problem problem. In S. Hetherington (Ed.), *Epistemology futures* (pp. 148–168). Oxford University Press.

Machery, E., Stich, S., Rose, D., Chatterjee, A., Karasawa, K., Struchiner, N., et al. (2017). Gettier across cultures. *Noûs, 51*, 645–664.

Montminy, M. (2014). Knowledge despite Falsehood. *Canadian Journal of Philosophy, 44*, 463–475.

Murphy, P. (2013). Another blow to knowledge from knowledge. *Logos and Episteme, 4*, 311–317.

Nagel, J., San Juan, V., & Mar, R. A. (2013). Lay denial of knowledge for justified true beliefs. *Cognition, 129*, 652–661.

Nozick, R. (1981). *Philosophical explanations*. Harvard University Press.

Oddie, G. (2014). Truthlikeness. In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/truthlikeness/

Popper, K. R. (1963). *Conjectures and refutations: The growth of scientific knowledge*. Routledge and Kegan Paul.

Putnam, H. (1973). Meaning and reference. *Journal of Philosophy, 70*, 699–711.

Saunders, J. T., & Champawat, N. (1964). Mr. Clark's definition of 'knowledge. *Analysis, 25*, 8–9.

Schnee, I. (2015). There is no knowledge from falsehood. *Episteme, 12*, 53–74.

Sorensen, R. (2013). Veridical Idealizations. In M. Frappier, L. Meynell, & J. R. Brown (Eds.), *Thought experiments in science, philosophy, and the arts* (pp. 30–52). Routledge.

Sorensen, R. (2016). Fugu for Logicians. *Philosophy and Phenomenological Research, 92*, 131–144.

Sosa, E. (1996). Postscript to 'proper functionalism and virtue epistemology'. In J. Kvanvig (Ed.), *Warrant in contemporary epistemology* (pp. 271–281). Rowman & Littlefield.

Sosa, E. (2009). Human knowledge, animal and reflective. In E. Sosa (Ed.), *Reflective knowledge: Apt belief and reflective knowledge* (Volume II ed., pp. 135–153). Oxford University Press.

Sosa, E. (2010). Replies to my Critics. *Crítica: Revista Hispanoamericana de Filosofía, 42*, 77–93.

Stalnaker, R. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory* (pp. 98–112). Blackwell.

Stroud, B. (2002). Understanding human knowledge in general. In B. Stroud (Ed.), *Understanding human knowledge: Philosophical essays* (pp. 99–121). Oxford University Press.

Swain, M. (1972). An alternative analysis of knowing. *Synthese, 23*, 423–442.

Swain, M. (1981). *Reasons and knowledge*. Cornell University Press.

Turri, J. (2011). Manifest failure: The Gettier problem solved. *Philosopher's Imprint, 11*, 1–11.

Turri, J. (2012). In Gettier's Wake. In S. Hetherington (Ed.), *Epistemology: The key thinkers* (pp. 214–229). Continuum.

Turri, J. (2019). Knowledge from falsehood: An experimental study. *Thought, 8*, 167–178.

Warfield, T. A. (2005). Knowledge from falsehood. *Philosophical Perspectives, 19*, 405–416.

Williamson, T. (1994). *Vagueness*. Routledge.

Williamson, T. (2000). *Knowledge and it limits*. Oxford University Press.

Williamson, R. (2009). Reply to Alvin Goldman. In P. Greenough & D. Pritchard (Eds.), *Williamson on knowledge* (pp. 305–312). Oxford University Press.

Zalabardo, J. L. (2012). *Scepticism and reliable belief*. Oxford University Press.