ABSTRACT
According to the knowledge view of evidence notoriously defended by Timothy Williamson (2000), for any subject, her evidence consists of all and only her propositional knowledge (E=K). Many have found (E=K) implausible. However, few have offered arguments against Williamson’s positive case for (E=K). In this paper, I propose an argument against Williamson’s positive case in favour of (E=K). Central to my argument is the possibility of the knowledge of necessary truths. I also draw some more general conclusions concerning theorizing about evidence.

Keywords: functions of evidence, E=K, evidential probability, probability raising.

Introduction

One of the revolutionary theses of Williamson’s Knowledge and Its Limits (2000) is that, for any subject, her evidence consists of all and only her propositional knowledge (E=K). E=K can be seen as flagship of the positive part of the knowledge-first program in epistemology: not only can knowledge not be analysed in more fundamental terms (the negative part), it can also be successfully used in characterising other epistemologically interesting notions, such as evidence, for instance. E=K is constituted of two theses: E→K and K→E. In defence of the latter Williamson has proposed that it is pre-theoretically intuitive and that possible arguments against it do not succeed. His main argument for E→K appeals to our ordinary concept of evidence, and to considerations about what kind of entities can fulfill the central functions of that concept. According to Williamson, the central functions of our ordinary concept of evidence are figuring in inferences to the best explanation, playing a role in probabilistic reasoning, and enabling one to rule out inconsistent hypotheses (inconsistent with it). He argues that only known propositions can play the central role of our ordinary concept of evidence.

A striking feature of the Williamsonian knowledge-first approach is its aspiration for theoretical fruitfulness and economy. This is most remarkably illustrated in Williamson’s treatment of evidence and evidential probability. Given the thesis that evidence is knowledge, and some minimal assumptions about mathematics of probability and constraints on evidential support, Williamson is in a position to propose a powerful model of evidential probability that rivals traditional subjective Bayesian approaches. Now, the better a model fares with respect to its explanatory capacities, the more it will be insensitive to particular descriptive details. This seems to affect E=K and knowledge-based accounts of evidential probability as well. Williamson is clear about this; and when discussing, for instance, his understanding of the probability function, he recognizes that while his assumptions about the mathematics of probabilities (e.g. the probability axi-
omn) entail that logically equivalent propositions will receive same probability on given evidence, this should be considered a price to be paid for the greater clarity and explanatory power. This sort of stance in preferring simplicity and explanatory power to capturing all the particularities of \textit{explananda} seems to be a general trait of Williamson’s epistemology.

On the other hand, however, Williamson is also attached to characterizing our ordinary concepts. This is true in particular in Williamson’s treatment of the concept of \textit{evidence} (and evidence property). His argument for \textit{E=K}, as we have already noticed, depends crucially on considerations about the ordinary usage of the concept of \textit{evidence}. These two aspects of Williamson’s approach stand in tension. Some aspects of this sort of tension in Williamson’s epistemology, namely between his aspiration for mathematical clarity and in particular his aspiration to maintain a version of Bayesianism in theorizing about evidence and his reliance on the ordinary usage of concepts and intuitive judgements, have been already noticed in the literature. For instance, Dunn (2014) sheds light on the tension between Williamson’s Bayesian commitments and his view that we can gain evidence through inductive inferences. In what follows I aim to bring light to another place where this sort of tension surfaces. The problem is not the tension itself, but rather its impact on Williamson’s positive case in favour of \textit{E=K}. In short, unless Williamson is ready to give up crucial bits of his formal approach, which, I believe, he will not do, his main argument in favour of \textit{E=K} cannot be taken to support \textit{E=K}. This is not to undermine \textit{E=K} or the knowledge-first approach in general. Rather, this is to claim that there is no easy way of providing arguments for a simple theoretical model that are based on the use of ordinary concept of \textit{evidence}.

\section*{Argument from the Knowledge of Necessary Truths}

Unsurprisingly, over the last fifteen years \textit{E=K} has received sustained attention. Many philosophers have found it implausible.\footnote{Here is a non exhaustive list of more or less radical critics of \textit{E=K}: Harman (2002), Joyce (2004), Silins (2005), Brueckner (2005), Hawthorne (2005), Dodd (2007), Whitcomb (2008), Neta (2008), Kelly (2008), Conee and Feldman (2008), Goldman (2009), Schiffer (2009), Comesana and Kantin (2010), Schroeder (2011), Rizzieri (2011), Littlejohn (2012), Logins (2013), Hughes (2013), Arnold (2013), Dougherty and Rysiew (2013), Dunn (2014), McGlynn (2014) and Mitova (2014).} However, few have attempted to undermine the positive case (allegedly) supporting Williamson’s thesis (recent exceptions include Hughes, 2014; Goldman, 2009). In this paper, I present an argument against Williamson’s positive case in favour of \textit{E=K}. Central to my argument is the possibility of the knowledge of necessary truths. The point of this discussion is only to show that the existence of known necessary truths raises more trouble for \textit{(E=K)} than one may have expected.

In order to consider how knowledge of necessary truths raises a problem for Williamson’s positive case for \textit{(E=K)}, let us first focus on the following argument:

\begin{enumerate}
\item \textit{(E=K)} For any subject \textit{S}, \textit{S}’s evidence is all and only the propositions that \textit{S} knows.
\item \textit{(Functionality of Evidence FOE)} For any subject \textit{S}, for any proposition \textit{e}. If \textit{e} is part of \textit{S}’s evidence then \textit{e} is evidence for some hypothesis \textit{h}.
\item \textit{(Probabilism)} Prior (unconditional) probability of necessary truths is 1 (\textit{P(p)=1}, where \textit{p} is a necessary truth) (Kolmogorov’s 2nd axiom).
\item \textit{(Known Necessary Truths KNT)} For some subject \textit{S}, and some necessary truth \textit{p}, \textit{S} knows \textit{p} (e.g. \textit{2+2=4}).
\item \textit{(p)} (e.g. \textit{2+2=4}) is part of \textit{S}’s evidence. (1, 4)
\item \textit{(6)} (e.g. \textit{2+2=4}) is evidence for a hypothesis \textit{h}. (2, 5)
\item \textit{(EV)} For any subject \textit{S}, for any proposition \textit{e}, and for any hypothesis \textit{h}, \textit{e} is evidence for \textit{h} for \textit{S} if and only if \textit{e} is part of \textit{S}’s evidence and the probability of \textit{h} given \textit{e} is higher than the probability of \textit{h} alone (i.e. \textit{P(h|e)>P(h)}), given that \textit{P(h)=0}).\footnote{See Williamson’s original formulation: “\textit{EV} is evidence for \textit{h} for \textit{S} if and only if \textit{S}’s evidence includes \textit{e} and \textit{P(h|e)} > \textit{P(h)}” (Williamson, 2000, p. 187).}
\item \textit{(8)} \textit{P(2+2=4)} < 1 \cite{6, 7}
\item \textit{(9)} \textit{P(2+2=4) = 1} \cite{3}
\end{enumerate}

\textit{(1)-(9)} lead to a contradiction; in order to avoid inconsistency one has to reject either (1), (2), (3), (4), or (7).

A crucial step in the argument is the inference from (6) and (7) to (8). The inference is valid. (\textit{EV}) entails that \textit{P(e)} cannot be 1 (where \textit{e} is evidence for \textit{h} for \textit{S}). As Williamson puts it: “For if \textit{P(h|e)} > \textit{P(h)}, then \textit{P(e)} is neither 0 (otherwise \textit{P(h|e)} is ill defined) nor 1 (otherwise \textit{P(h|e)} = \textit{P(h)})” (Williamson (2000, p. 187)). The following shows why \textit{P(h|e)=P(h)} when \textit{P(e)=1}. Start with the definition of the conditional probability: \textit{P(h|e)=P(e|h)/P(e)}. Suppose that \textit{P(e)=1}. \textit{P(h|e)=P(e|h)/1=P(e|h)}. Now, \textit{P(h)=P(h&e)} + \textit{P(h&not-e)}. If \textit{P(e)=1}, then \textit{P(h&not-e)=0}. We have supposed that \textit{P(e)=1}. Hence, \textit{P(h)=P(h&e)}. Remember that if \textit{P(e)=1}, then \textit{P(h&e)=P(e)}. Therefore, \textit{P(h|e)=P(h)}. Hence, \textit{P(e)} has to be less than 1 if (\textit{EV}) holds. Now; (6) tells us that \textit{P(e)} in our case is \textit{P(2+2=4)}. Hence, we have to infer that \textit{P(2+2=4)} is less than 1.\footnote{Thanks to Julien Dutant and an anonymous referee here.}

Williamson (2000) is committed, on pain of inconsistency, to the rejection of (2), for he is explicitly committed to (1), (3), (4), and (7).\footnote{One could also argue that a proponent of \textit{E=K} may rather revise the \textit{EV} principle (premise 7) in order to avoid the inconsistency. For one thing, Williamson himself seems to be open to potential revisions of \textit{EV}: “At least as a first approximation, we can model the first
cost: one loses the important mathematical power of probabilism. By rejecting (4), one relinquishes a highly plausible view and by the same token concedes a lot to the sceptic, whereas by giving up (7), one forfeits simplicity and fruitfulness in theorising about evidence and evidential support in terms of probability. Of course, endorsing the claim that something can be part of one’s evidence set and yet not be evidence for any hypothesis is something of an oddity. However, on balance, it might appear to be the “lesser evil” in the present dialectical situation. Let me stress again that Williamson is committed to (1), (3), (4), and (7), and even if he were to revise his commitments the prospects for a reasonable rejection of one of these looks very bleak. We have to reject (2) given the extremely high plausibility of (3), (4), and (7) if we also want to maintain E=K.6

Furthermore, one might think that the rejection of (2) can also be motivated on independent grounds, for it is possible to distinguish between the following two concepts: evidence-for-a-hypothesis (evidence-for-h) and subject’s-body-of-evidence (S’s-evidence).7 Moreover, it seems that some passages from Williamson (2000) are hints towards, if not explicit commitments on this distinction.8 When we say things like, “the fact that photographs exist of ice on Mars is evidence for the hypothesis that there is water on Mars”, we use evidence in the evidence-for-h sense. Whereas when we advance that NASA has an impressive body of evidence, we are using evidence in the sense of S’s-evidence.9 Once the distinction is accepted, the proponent of (E=K) can claim that, while necessary truths can be S’s-evidence, they can never be evidence-for-h. This enables a proponent of (E=K) to motivate the rejection of (FOE), while maintaining that it is possible to know necessary truths, and that evidence for a hypothesis is that which raises the probability of a hypothesis (EV). Such a move is not completely far-fetched, since we rarely (if ever) talk of necessary truths as evidence-for-a-hypothesis.10 What...
is more, theoretically speaking, the gain in simplicity and in the ability to preserve probabilism, knowledge of necessary truths, and (EV) might, after all, be worth accepting the claim that known necessary truths are not evidence for any hypothesis, despite their being evidence possessed by some subjects.

How the two concept solution undermines the central function argument for E=K

However, giving up (2), i.e. (FOE), is not as benign to (E=K) as one might think. That is, giving up (2) entails an important, until now unnoticed, dialectical cost for an (E=K) theorist. Namely, the rejection of (FOE) undermines a major argument in favour of (E=K). More specifically, if (FOE) does not hold, one cannot use the argument from the central functions of our ordinary concept of evidence in favour of E=K.

As we have already noted, the argument from the central functions of our ordinary concept of evidence (the central function argument) has been put forward by Williamson as the main positive case in favour of (E=K) (2000, p. 193-207). The argument is a defence of the claim that only known propositions can serve the central functions of our ordinary concept of evidence, and that, since only entities that can serve central functions of our ordinary concept of evidence can be evidence, one has to accept that only known propositions can be evidence. The central functions of our ordinary concept of evidence are supposed to be: to figure in inferences to the best explanation; to play a role in probabilistic reasoning/confirmation; and to enable one to rule out inconsistent hypotheses. Williamson’s central function argument proceeds by two steps (not logically dependent, but merely dialectical), by first defending the view that only propositional items can serve the central functions of evidence, and then by defending the view that nothing less than known propositions could serve these functions.

Now, the problem is that the central function argument is supposed to support the claim about S’s evidence: that one’s evidence is all and only one’s knowledge. In order for it to be successful, it has to show that only known propositions can serve the central functions of our ordinary concept of S’s-evidence. However, the functions that Williamson presents as the central functions of “our ordinary concept of evidence” are functions of the concept of evidence-for-h; these are not functions of the concept of S’s-evidence. Consider, for instance, the function of playing a role in probabilistic reasoning. The playing of a role in probabilistic reasoning is a central function of the concept of evidence-for-h, i.e. the concept of evidence-for-a-hypothesis. It seems implausible that this is a central role of the concept of S’s-evidence. Williamson considers probabilistic comparisons to be a paradigmatic form of probabilistic reasoning. Hence, he claims: “[o]ne way of using those probabilities is to regard h as more probable than h* given e (P(h|e) > P(h*|e)) if and only if h makes e more probable than h* does (P(e|h) > P(e|h*))” (2000, p. 196). Probabilistic comparison is central to our use of the concept of evidence-for-h. However, it is far from obvious that probabilistic comparisons are central to our use of the concept of S’s-evidence. Indeed, it seems that the contrary holds. If there is any meaningful way of distinguishing between concepts of evidence-for-h and S’s-evidence, then playing a role in probabilistic reasoning (e.g. in comparing hypotheses) is obviously one thing that should make a difference between the two concepts. Hence, the rejection of (FOE) and introduction of the distinction between evidence-for-h and S’s-evidence have the consequence that, at best, the main, allegedly positive, argument in favour of (E=K) can only be an argument in favour of the claim (Evidence-for-h=K), which is not what the argument is intended to support. In other words, if one rejects (FOE) and accepts the distinction between evidence-for-h and S’s-evidence, then one of the main arguments in favour of the view that one’s evidence is constituted by all and only one’s knowledge is undermined.

Furthermore, notice that, some passages in Williamson (2000) might be read as saying that he is committed to the view that only items that can serve the central functions of evidence can be part of one’s evidence. He seems to press this point in a reply to the objection against (E=K) according to which some non-propositional items can be one’s evidence. See, for instance: “Although evidence may well have central functions additional to those considered above, genuine evidence would make a difference to the serving of the functions considered above, whatever else it made a difference to” (Williamson, 2000, p. 197).

In short, if “genuine evidence” in the quoted passage means “genuine S’s-evidence”, then Williamson is committed to (FOE). However, if this is so, then there is a serious tension here, given his commitment to its negation (cf. his commitment to (E=K), (EV), probabilism, and (KNT)). Given such a tension, it seems that the most charitable reading of “genuine evidence” in the above quote should be one that takes it to mean “genuine evidence-for-h”. However, the trouble with this is that, if we accept this meaning, the argument from the central function of evidence does not support the view that S’s-evidence is knowledge (E=K), but merely that evidence-for-h is knowledge. In other words, the argument from the central functions of our ordinary concept of evidence misses its target and fails to provide support for the specific claim that one’s evidence is all and only one’s propositional knowledge.

Conclusion and more general comments

Now, of course, this result does not undermine E=K. Indeed, in face of the argument from the knowledge of necessary truths a proponent of E=K may well choose to avoid the inconsistency by endorsing the two evidence concepts
solution and maintain E=K on the basis of its simplicity and theoretical fruitfulness. One may think of E=K, probabilism, and EV as constituting a simple theoretical model that help us to advance in epistemology and illuminate complex issues and concepts. There is a good chance that E=K is the most simple and elegant account of evidence possession on the present-day philosophical market.

However, I hope that the above discussion has shown that it is unlikely that one can theoretically motivate such a model by appeal to arguments that substantially rely on the main functions of our ordinary concept of evidence. I have shown that Williamson’s argument that appeals to the ordinary use of evidence in favour of E=K doesn’t support E=K, given his theoretical model. The moral that we might take from this result is that there is a tension in arguing about evidence: either we can have a simple model of evidence (and evidential probability) based on E=K (and mathematics of probability) that has a great simplicity and explanatory power, or we can have an argument for E=K that is based on the use of an ordinary concept of evidence. It seems that we cannot have both. In this context then, not only does Williamson’s central function argument fail, but any similar argument will fail. The best and the only thing we can do to theoretically motivate E=K if we also want the power and the simplicity of probabilism and EV (without giving a lot away to sceptics) is to argue for E=K on the basis of merely methodological considerations: it is the simplest and the most powerful model of evidence. This, however, is rather heterodox way of arguing for a view about evidence within contemporary mainstream epistemology. Most debates in mainstream epistemology, as far as I can see, crucially rely on (ordinary) intuitions about cases, not methodological arguments. Interestingly, this observation may help us to understand why Williamson’s E=K has attracted such an impressive number of critics without ever, as far as I can see, being shown to be inconsistent.

Now, what should we think of E=K? Should we reject it on the basis of the numerous, apparently counter-intuitive consequences it has? At this point, I would like to suggest that the answer to this question will ultimately and surprisingly depend on considerations about the methodology of epistemology. What kind of arguments do we want to be decisive in epistemology? Do we want to allow particularities of cases to undermine powerful theoretical models or should we value simplicity and explanatory power above all? These are essential questions that have to be addressed before we can hope to make further substantial progress in theorizing about evidence and other epistemologically interesting concepts.

Acknowledgements

Many thanks to Matthew Benton, Julien Dutant, Santiago Echeverri, Davide Fassio, Daniel Rubio, Tim Williamson, and audiences at Philosophisches Kolloquium (resp. prof. Barbara Vetter), Department of philosophy, Humboldt University of Berlin and European Epistemology Network Meeting 2016 for discussions and comments on earlier versions of this paper. The research work that lead to this article was supported by the Swiss National Science Foundation grant number 161761.

References


In a reply to a similar yet distinct worry (perhaps an even more urgent one), concerning the evidential probability of logically equivalent propositions (e.g. roughly, that the axioms of probability “entail that logically equivalent propositions have the same probability on given evidence”; Williamson, 2000, p. 212), Williamson explains that this is the price to be paid for the simplicity and power that is provided by the use of mathematics in theorizing about evidential probability: “We are using a notion of probability which (like the notion of incompatibility) is insensitive to differences between logically equivalent propositions. We therefore gain mathematical power and simplicity at the loss of some descriptive detail (for example, in the epistemology of mathematics): a familiar bargain” (Williamson, 2000, p. 212).
https://doi.org/10.1007/s11098-013-0184-9

https://doi.org/10.1111/j.1468-0149.2004.0356c.x

https://doi.org/10.1111/j.1747-9991.2008.00160.x

https://doi.org/10.1017/CBO9781139060097


https://doi.org/10.1057/9781137026460

https://doi.org/10.1093/bjps/axn003

https://doi.org/10.1007/s11098-009-9488-1

https://doi.org/10.1093/acprof:oso/9780199287512.003.0012


https://doi.org/10.1111/j.1520-8583.2005.00066.x

https://doi.org/10.1007/s11098-014-0339-3

https://doi.org/10.1007/s11098-006-9024-5