

Chadwick Prize Winner: Christian Michel

Why Truth Trackers Do Not Need to Reject Epistemic Closure

In this paper I shall argue that Fred Dretske's rejection of the principle of epistemic closure under known implication is both ineffective and unnecessary. It is ineffective, because Dretske cannot argue convincingly to reject closure independently from his knowledge account. Unnecessary, because his theory of knowledge does not need to reject closure - contrary to the commonly held belief that the closure principle and truth tracking accounts are incompatible.

1. Closure, the argument from ignorance and truth tracking

The principle of epistemic closure under known implication holds roughly that if S knows P and also knows that P entails Q then S knows Q, or slightly more precisely: then S has all that he needs to know Q. Here I follow the formulation of Dretske (2005).¹

This seemingly innocent and highly intuitive principle is, however, heatedly debated in epistemology, for good reasons. A lot seems to be at stake. The discussion about the validity of the closure principle has close links to the discussion of the "skeptical argument from ignorance". This argument states, in a nutshell, that we cannot know ordinary things like "I have two hands", if we don't know that we are not manipulated brains in a vat (BIVs) or otherwise cleverly and systematically deceived creatures. Formally, it runs like this:

- (1) We don't know that we are non-BIVs
- (2) Assuming that we have hands, then we are non-BIVs (because we are not floating handless brains in a vat)
- (3) Closure principle: If we know P and we know that from P follows Q, then we know Q. Or equivalent: If we don't know Q and from P follows Q, then we don't know P

Conclusion:

Setting P = we have hands and Q = we are non-BIV in (3), we conclude with (1) and (2) that we don't know we have hands.

The question now is, how to avoid the skeptical conclusion, which seems counterintuitive and unacceptable. There are many strategies to respond to the skeptic,² but the focus in what follows is on the defence through truth tracking accounts of knowledge, like that proposed by Dretske.

Let's abbreviate in what follows a proposition about ordinary things (e.g. "I have two hands") by "OT" and a skeptical hypothesis of the type "I am a BIV" by "SH".

It is commonly accepted that the rejection of closure is a necessary consequence of "truth tracking" accounts of knowledge. Dretske's knowledge theory based on "conclusive reasons" goes, roughly, like this: S knows P if S has conclusive reasons R for believing P. Conclusive reasons R are such that if P were false R would not hold. For OT we may have conclusive reasons but for non-SH we have no conclusive reasons, as SH is - by definition - constructed such that our belief is insensitive to truth. Hence, a truth tracking account implies that we know OT but we don't know non-SH.

(1) is then true and the conclusion from the argument from ignorance needs to be false, because truth tracking allows for knowing ordinary things. To render the argument from ignorance invalid and save us from the skeptic conclusion Dretske rejects premise (3), the closure principle.

¹ The principle of closure needs some qualifications and strictly cannot stand in this simple form (see e.g. Hawthorne, 2005). However for the purpose of this paper I consider the simplified version as sufficient. There are many trivial counterexamples based e.g. on the possibility that S may not believe Q and therefore does not know Q, or that usually one does not know all known implications. I assume that the closure principle can be qualified as to eliminate such trivial counterexamples.

² See deRose (1995) for a discussion of various strategies and his own contextualist account.

There is another possibility to render the argument from ignorance invalid, namely to reject (2). In the different formulations of the argument from ignorance premise (2) is usually tacitly assumed as being true [see e.g. deRose (1995)] and the possibility of its falsehood is not considered. But if we accept that the truth of a proposition is a *necessary condition* for knowledge, then we could develop a strategy to show that (2) is false and the truth tracker need not anymore to reject closure in order to avoid the skeptical conclusion.

2. Dretske's independent defense of the rejection of closure

The denial of closure is a very controversial position because closure seems so intuitive and productive for expanding our knowledge. The fact that the rejection of closure follows from Dretske's account of knowledge could be used as an argument against him. Therefore, to find arguments for the rejection of closure *independently* from his theory of knowledge would support truth tracking theories. Indeed, Dretske intends to provide different such independent arguments (2005). I will focus on the main argument, the one related to transmission of evidential warrant from OTs to SHs and show that it is ineffective.

Dretske argues (2005, p14) that "our ways of discovering P are not necessarily the ways to discover what we know to be implied by P". Dretske uses the notion of "protoknowledge" to make his point: there are things that need to be true for what we perceived to be true. That there are material objects need to be true for there are cookies in the jar to be true. Dretske continues to argue that some reasons for believing P do not transmit to things Q known to be implied by P. At this point Dretske introduces the notions of *heavyweight* (HW) and *lightweight* (LW) propositions in order to make the non transmissibility plausible. A heavyweight proposition is of the SH-sort: "I am a brain in the vat", "There is no external world", "The past is not real". Indeed it is plausible that the reasons that make me believe a LW are not effective to make me believe a HW. Nevertheless, Dretske is convinced that "There are things, heavyweight implications, we needn't know, even though we know our knowledge depends on them" (2005, p17).

To this I respond firstly that I *cannot* have *any reason at all* to believe HW propositions, as my belief of them is insensitive to truth. Secondly, the reasons for believing P and Q should not be the same, because then I could know P and Q independently for the same reason. But the point of closure is precisely that I can extend my knowledge from some proposition P via *other reasons*, e.g. via a logical principle, to a different proposition Q.

To appeal, via the distinction of HW and LW propositions, to a property of the *content* of the proposition P and Q, in order to explain the failure of closure seems not satisfactory. Others use similar strategies, like Colin McGinn (2002) with his account of "discriminatory knowledge". He states that different types of propositions have different discriminatory requirements and closure fails for the problematic propositions for which I cannot tell whether they are true or not (those correspond roughly to Dretske's HW propositions).

Part of the intuition we have about the closure principle is that it seems to be structured like a universal formal law (though more complicated than just a logical law, like modus ponens, because it includes the knowledge operator) that depends only on the truth value of the propositions P and Q, not on a *specific content* of those two propositions. The principle is based on a formula that contains variables P, Q standing for propositions, logical operators and the operator "know". It is the crux of our intuition that the *content* of the proposition P, Q shall not matter as long as P, Q and "P entails Q" are true. The knowledge operator, of course, is not a logical operator, but a vague one that may include psychological states, like believing, that could behave strangely and provoke, under certain circumstances, closure to fail. However, it seems we intuit that a rational person *should (and normally would)* believe Q, because the fact that P implies Q renders Q true if P is true and the logical implication is a strong reason to belief Q. You have then all you need to know Q. Even without committing to any theory of knowledge and leaving knowledge a quite vague notion, we strongly intuit that closure must hold.

3. A thought experiment

It is important to stress that the argument from ignorance only works if proposition “P entails Q” is true, provided we endorse the commonly accepted claim that we can know only true propositions. I will argue that a LW proposition cannot entail a HW proposition. I will show that to say P implies Q where P is a LW and Q a HW proposition is vague in a certain sense and requires a qualification. Then I show that if “P entails Q” is meant in a one sense it does not suppose any problem with regard to the skeptical argument. If it is meant in the other possible sense it is intuitively false. Our perception of a dilemma through the argument from ignorance arises from conflating both senses.

Let’s analyse a proposition where a LW proposition entails some HW proposition. The following entailment seems intuitively true and is typically used in the closure formulas in the argument from ignorance:

(I) “If I have two hands then I am not a brain in a vat “

Why is this intuitive ? Most plausibly because we think like this:

(II) “I am a BIV” entails “I have no hands” (and not two hands !)

To illustrate why and where exactly the entailment (I) requires some qualification let us carry out a thought experiment. Imagine you are a BIV and you only have the hallucination of having two hands, so you do not have two hands (“in reality”). Only an observer, let’s say the scientist controlling you being a BIV, standing next to the vat and observing you being a brain in the vat can tell whether you are a BIV or not. Now the vat-observer has a more objective and comprehensive perspective than you as a BIV. He lives in an epistemic world, let’s call it “world 2” that has a higher level of objectivity. “World 1”, corresponding to your world as a BIV, is embedded in world 2 such that your beliefs that are insensitive to truth for you, are sensitive to truth for the vat-observer (e.g. you don’t know that you are a BIV, but the observer knows that you are a BIV). He is observing you, and how you are being deceived in the vat. He could say something like this to you:

“Yes, you have dreamed hands - world 1-hands - which are “real” hands according to your criteria and perspective - but they are not “ultimately real” or “absolutely real”. But you don’t have “real” hands as I have - world 2-hands. I have full transparency on your states and more information than you. So I understand what you mean and you are right in a certain sense, from your lower perspective.

It is possible that there is even a world 3 in which I am myself floating in a tank as a BIV ! Who knows!”

With this example as a background, the implicit meaning of (I) is the *intuitively false* implication:

(III) “If I have two world 1-hands then I am not a world 2- BIV.”

Indexing the world we refer to, makes the fallacy obvious. The following proposition using a different indexation is obviously true:

(IV) “If I have two world 1-hands then I am not a world 1-BIV.”

As (IV) talks only about world 1, this type of BIV scenario is not problematic, because world 1 subjects with world 1-hands could recognize their situation (world 1 consists of all truths to which the beliefs of world 1 subjects are sensitive). Therefore, premise (1) in the argument from ignorance is false and hence the skeptic conclusion does not obtain. What we really mean, or we are driven towards unconsciously, when we reflect on the argument from ignorance is (III), i.e. the claim to be a world 2-BIV from a world 1-perspective, because this implies that premise (1) is true (i.e. we don’t know we are BIVs – and we cannot even in principle know because of the way the the BIV scenario is defined).

The upshot is that the proposition (I) seems ambiguous under the above analysis and we need to consider “world levels”. The interpretation (IV) of (I) is unproblematic and trivial. Hence, the meaning (III) must be the relevant, but (III) is intuitively false.

4. How to preserve Dretske's motivation to reject closure without rejecting closure

The argument from transmission of warrant of Dretske can be interpreted, under the light of the vat-observer thought experiment, in a plausible way, though those reasons do not constitute at all reasons to reject closure.

The "right" intuition of Dretske is obviously that there are different types of propositions that when appearing in a closure inference generate problems. It seems indeed absurd to infer from a trivial fact like "I see cookies in the jar" such far-reaching implications like "There is an external world". HW propositions are simply propositions the beliefs of which are insensitive to truth. They are beyond our human epistemic reach. As Dretske correctly points out, there is nothing in the material world that can indicate to us that there is a material world.

Indeed, justification doesn't transfer from LW to HW propositions - not because closure does not hold, but because we should not plug the LW->HW entailments into the closure equation *in the first place* because they are false and hence cannot constitute a known entailment.

The final upshot is that we should not restrict the scope of the closure principle artificially, and in an ad-hoc fashion, only to LW propositions as Dretske and McGinn do, yet at the same time we can continue to defend a truth tracking account. As LW->HW entailments are false in its proper interpretation, the closure principle is not applicable in the first place to them, because such entailments cannot be known.

5. Some objections

The kind of stratification of worlds proposed may be considered an ad-hoc solution. However, I think that the thought experiment is plausible and intuitive. The "world levels" define perspectives of epistemic accessibility. Most thinkers assume, more or less consciously, models of stratification of reality or objectivity. Simply think of the epistemological concepts of "externalism" and "internalism". Externalism is the position that the justification to believe a proposition does not require conscious access by the subject to the reasons why the belief is justified, while internalism does require it. Those concepts imply also a certain stratification relative to epistemic access, so stratification is not such an exotic move in epistemology.

One could further argue that the stratification strategy does not work for more limited and down to earth examples, like the famous cleverly painted mules. If I am in a zoo and see zebras I know that there are zebras, though I don't know that they are not cleverly painted mules. The difference from the BIV example is that the scope of deception is much more limited. Dretske's strategy in this example is to argue that we can endorse, without being guilty of logical contradiction, the conjunction "I know there is a zebra" and "I don't know that it is a cleverly painted mule", because closure does not hold. According to the account proposed in this paper, we can endorse this conjunction, but not because closure fails, but because the implication (OT) "There is a zebra" -> (non-SH) "It is not a cleverly painted mule" is not true. The scenario of the cleverly painted mule is equivalent to the BIV scenario, though the deceptive scope of the former ranges only over the very narrow domain of knowledge of a part of the inhabitants of a zoo. With respect to the knowledge of the inhabitants of the zoo we can consider world 1 and world 2 levels. World 2 subjects are then not represented by vat-observers, but e.g. by zoological experts that can distinguish zebras and cleverly painted mules. And the judgment about whether the mule is cleverly painted or not is - per construction of this scenario - beyond the epistemic reach of the world 1 subject, the visitor, so she knows in world 1 that this is a zebra, but she does not know whether it is a mule in world 2.

Finally, one may argue that the proposed approach endorses skepticism. If indeed no ultimate view exists and if we are only epistemically competent in the limited world 1 then we are quite ignorant, at least ignorant with respect to other levels of reality. However, the account given here concedes to skepticism that we cannot know *all* or we cannot know anything in an absolute, ultimate sense (if this is conceding anything to skepticism) but the upshot is we can know *something*, namely world 1 things.

6. Conclusion

I have argued that entailments of the type “If I have two hands, then I am not a brain in a vat” are ambiguous and need to be interpreted such that they are either false or are true only in an unproblematic sense. This allows us to argue that truth tracking and the closure principle need not be incompatible to render the skeptical conclusion false.

Bibliography

- DeRose, K., 1995 ‘Solving the Skeptical Problem’, *The Philosophical Review*, Vol.104, No.1 (Jan., 1995), pp. 1-52.
- Dretske, F., 1981 ‘The pragmatic Dimension of Knowledge’ from *Philosophical Studies*, 40. Extracts in: *Reading Epistemology* by Sven Bernecker, 2006, Blackwell Publishing.
- Dretske, F., 2005 ‘The Case against Closure’ in *Contemporary Debates in Epistemology*, edited by M.Steup and E. Sosa., Blackwell Publishing.
- Hawthorne, J., 2005 ‘The Case for Closure’ in *Contemporary Debates in Epistemology*, edited by M.Steup and E. Sosa., Blackwell Publishing.
- McGinn, C., 2002 *Knowledge and Reality*, OUP.