[Forthcoming in *The Philosophical Quarterly* – this is an uncorrected draft, please don't quote from this version]

What can we know about unanswerable questions?

Thomas Raleigh, University of Luxembourg

ABSTRACT: I present two arguments that aim to establish logical limits on what we can know. More specifically, I argue for two results concerning what we can know about questions that we cannot answer. I also discuss a line of thought, found in the writings of Pierce and of Rescher, in support of the idea that we cannot identify specific scientific questions that will never be answered.

"Not ignorance, but ignorance of ignorance, is the death of knowledge."
(attributed to Lord Alfred North Whitehead)

1. INTRODUCTION

In this paper I will explore one particular way in which we can try to establish limits to what we can know. More specifically, I will try to establish a limit on what we can know about the *questions* we can or cannot answer. But before we begin, it may be worth just briefly mentioning a few possible ways in which one could try to trace the limits of what can be known that I will *not* discuss in this paper[1]:

(i) By appeal to facts about the finite, contingent nature of human brains.
Maybe neuroanatomical or psychological facts about how human minds are physically realised will allow us to set upper limits on: the complexity of propositions that can be learned or memorized or even entertained. Or perhaps precise limits on what the human brain can do could be framed in terms of information processing or computation. Of course this would then raise the question of whether such biological limits could be overcome via technology – neuro-enhancement or computer assistance, etc.

(ii) By appeal to general limits on provability and computability.
Famous limitative results, like those of Gödel, Church and Turing have established that there are limits on what can be proved within a formal axiomatic language and limits on what can be computed by any algorithmic (effective, mechanical) method. Thus there are undecidable problems – such as the halting problem – and there are non-computable functions and non-computable numbers – such as the 'busy beaver function' or 'Chaitin's constant'. Discussion of whether/how these results bear on the questions of what can be *known*, would raise questions about how formal proofs and algorithmic computations are related to knowledge – e.g. Searle's Chinese Room argument[2], or Lucas and Penrose's notorious argument that human mathematicians must somehow be performing non-computable functions that are not describable as formal proof systems[3]. One might also want to consider speculative

[1] This list is certainly not meant to be exhaustive – there could well be other ways by which we might try to establish facts about our own ignorance.
[2] See Searle (1980)
[3] See e.g. Lucas (1961), Penrose (1989).

suggestions that hyper-computational systems (i.e. which could compute a 'non-computable' function) could be physically realized – e.g. by quantum mechanical systems[4].

(iii) By appeal to algorithmic complexity.
As well as appealing to limits on what is in principle computable, one might also appeal to whether a problem can only be solved in super-polynomial time – e.g. if a problem can only be solved by an algorithm that requires exponential resources, then one might take it to be *practically* or *physically* insoluble (though an approximate or sub-optimal solution could still be feasible). Of course the issues here turn on the relationship between the complexity classes P and NP, which is widely regarded as the most important open question in theoretical computer science.

(iv) By appeal to the mathematical phenomenon of (deterministic) *chaos*.
The evolution over time of non-linear dynamic systems can exhibit such sensitivity to the initial conditions that prediction of future states of the system is impossible – since the initial state can never be measured accurately enough.

(v) By appeal to scientific or natural limits on information or measurement.
For example, given that nothing can travel faster than light, no kind of signal or causal influence from an event that is outside of our backwards-light cone could ever be detected by us. So specific information about events outside of the light cone cannot, as a matter of natural law, be known – though perhaps general facts about such events could still be known. Likewise, Heisenberg's Uncertainty Principle might well be thought to set a kind of natural limit on what can be known.

These are surely all interesting topics, eminently worthy of philosophical attention, but they will not be the topic of the present paper. Rather, I want to outline a kind of *logical* limit on what we can know about the questions we can or cannot answer – a limit that is perhaps in the same spirit as the well-known Church-Fitch Theorem. The plan for the paper then is as follows: in section 2 I will very briefly discuss the Church-Fitch result and also briefly discuss what can be known about various forms of higher-order ignorance. In section 3 I discuss a reason for thinking that we can never know for sure of some specific, currently unsolved empirical question whether or not the answer will remain forever unknown. In section 4 I discuss an interesting contradiction that points us towards some limits on what we can know about such unanswerable questions. In section 5 I present two arguments establishing that there are these logical limits on what we can know. Finally, section 6 provides a very short summary and conclusion.

2. CHURCH-FITCH AND HIGHER-ORDER IGNORANCE

One logical limit on what can be known is provided by the Church-Fitch theorem[5]. This very short and apparently simple, though much-discussed and disputed proof purports to show

[4] See e.g. Kieu (2003), Ziegler (2005).
[5] The result first appeared in print in Fitch (1963) – though Fitch himself stated in a footnote that he owed the idea to an anonymous referee. Archival investigations by Joe Salerno eventually revealed that this referee was none other than Alonzo Church – see Salerno (2009), Church (2009). The theorem is sufficiently surprising that it is sometimes called the 'Church-Fitch Paradox'. However, Williamson

that: if there is at least one unknown truth, then there must be an unknowable truth. (Another way of putting this result is that: if every truth is in principle knowable, then every truth is in fact known.) A huge amount has since been written about the Church-Fitch theorem and how the result might be avoided or defended, but it would be a distraction to discuss any of this secondary literature here[6]. I mention the Church-Fitch theorem just to set the scene, since the results I am going to present are very much in the same spirit – logical limits on knowledge established via *reductio*. All that matters for present purposes is that the theorem states that it is impossible to know *of some specific proposition*, p, both that it is true and that it is unknown:

$$\neg\Diamond K(p \,\&\, \neg Kp)$$

However, it *is* clearly possible to know that one does not know *whether or not* some specific proposition, p, is true:

$$\Diamond K(\neg Kp \,\&\, \neg K\neg p).$$

And so it is perfectly possible to know that *there exists some or other proposition* that is true and that is unknown:

$$\Diamond K[\exists p(p \,\&\, \neg Kp)]$$

Given that my ultimate aim will be to show that there is something we will never know about what we will never know, I want to briefly clarify and distinguish between some different varieties of *higher-order ignorance*. Now, as a matter of linguistic fact, it seems that the English-language term 'ignorance' does *not* simply mean the absence of knowledge. For when you have a lucky or irrational true belief that p, you fail to know that p, yet you don't count as ignorant of the fact that p[7]. Also: 'ignorance' seems to be a *gradable* notion, whereas attributions of propositional knowledge are not gradable[8]. But so just for terminological convenience, let's stipulate the following definitions:

Ignorance of the fact that $p =_{df} p \,\&\, \neg Kp$	[This is factive state, requiring p to be true]
Ignorance whether $p =_{df} \neg Kp \,\&\, \neg K\neg p$	[This is not a factive state]

Notice that ignorance of the fact that p *entails* ignorance whether p, since p's truth entails that one cannot know not-p. And as we have just seen, whilst it is perfectly possible to know that you are *ignorant whether* p, the Church-Fitch result states that it is impossible to know that you are *ignorant of the fact* that p. But so what about higher-order ignorance? I.e. ignorance of one's own ignorance? Are there also logical constraints here on what we can know about higher-order ignorance? If the quotation attributed to Lord Whitehead at the start of this paper is correct, then we should be especially worried about this kind of 2nd-order ignorance. But is this 2nd order ignorance something it is even possible for one to know?

(2000) makes the following comment: "The conclusion that there are unknowable truths is an affront to various philosophical theories, but not to common sense. If proponents (and opponents) of those theories long overlooked a simple counterexample, that is an embarrassment, not a paradox. (Williamson, 2000, 271)

6 See e.g. Hart & McGinn (1976) Williamson (1987, 2000), Tennant (1997, 2010), Brogaard & Salerno (2002, 2006), Edgington (1985, 2010), Jago (2010), Kelp & Prichard (2009).

7 See Peels (2010, 2012).

8 See Brogaard (2016).

Well, notice first that given the two different 1st-order forms of ignorance – ignorance of the fact that p vs. ignorance whether p – there will then be four possible forms of 2nd-order ignorance:

(1) Ignorance of the fact that [One is ignorant of the fact that p]

$(p \,\&\, \neg Kp) \,\&\, \neg K(p \,\&\, \neg Kp)$

(2) Ignorance whether [One is ignorant whether p]

$\neg K(\neg Kp \,\&\, \neg K\neg p) \,\&\, \neg K\neg(\neg Kp \,\&\, \neg K\neg p)$

(3) Ignorance of the fact that [One is ignorant whether p]

$(\neg Kp \,\&\, \neg K\neg p) \,\&\, \neg K(\neg Kp \,\&\, \neg K\neg p)$

(4) Ignorance whether [One is ignorant of the fact that p]

$\neg K(p \,\&\, \neg Kp) \,\&\, \neg K\neg(p \,\&\, \neg Kp)$

Notice also that since ignorance of the fact that p always entails ignorance whether p, type (1) 2nd-order ignorance will entail type (4), and likewise type (3) entails type (2).

So now we have these four different kinds of 2nd-order ignorance on the table, we can ask: is it possible to *know* that we are higher-order ignorant in one or more of these four senses? Well, three of the four kinds of higher-order ignorance are fairly simple cases. Both type (1) and type (3) are ignorance of the fact that one is ignorant in some way and so the Church-Fitch theorem would suffice to show that these two kinds of higher-order ignorance cannot be known. Whereas with type (4), it is immediately apparent that such higher-order ignorance *is* knowable. For whenever you are ignorant whether p, then for all you know it could be that $(p \,\&\, \neg Kp)$ but equally for all you know it could be that $\neg(p \,\&\, \neg Kp)$, since for all you know p could be false. So whenever you are ignorant whether p, you are also ignorant whether you are ignorant of the fact that p. And so assuming we can sometimes know that we are ignorant whether p, we can also sometimes know that we are type (4) higher-order ignorant. However, with type (2) higher-order ignorance, things get a bit more complicated. Kit Fine (2018) has recently shown that whether the type (2) kind of higher-order of ignorance can be known depends on which system of modal logic we choose. If we use a system, such as S4 or stronger, with the axiom $\Box p \rightarrow \Box\Box p$[9], then our model of knowledge will have the 'positive introspection' axiom $Kp \rightarrow KKp$. Given this rule that knowing entails knowing you know, Fine demonstrates that higher-order ignorance of type (2) *cannot* be known. And so if you are 2nd order ignorant in this sense, then you must also be 3rd order, 4th order, etc., ignorant. However, if we use a weaker modal logic, such as T, which does not include the $Kp \rightarrow KKp$ rule as a theorem, then Fine shows that higher order ignorance of type (2) *can* be known. In summary then: of the four kinds of 2nd-order ignorance, two of them are unknowable (types 1 and 3) assuming one endorses the Church-Fitch result, one of them clearly is knowable (type 4) and for one kind (type 2) it depends on our choice of modal logic.

[9] So the accessibility relation between worlds is reflexive and transitive.

3. WHAT CAN WE KNOW ABOUT QUESTIONS WE WILL NEVER ANSWER?

I want to focus now on establishing a kind of logical limit on what we can know that is couched in terms of the *questions* that we can or cannot know the answer to. I will be discussing whether it is possible for us to identify specific questions that we know we will not ever be able to answer. To be slightly more precise, I will be concerned exclusively with questions about *empirical or scientific matters* – so not purely logical or mathematical questions, nor moral or evaluative questions – which can be posed in yes-or-no format, such that we know that either a proposition or its negation is the correct answer but we cannot know which. Is it possible for us to identify such 'scientific insolubilia'? Of course it is surely extremely plausible that there are *some or other* empirical questions that will never in fact end up being resolved – i.e. that no actual, finite inquirer will ever know the answer to. After all, there are presumably just too many possible questions concerning arbitrary matters of empirical fact for us to know the answers to them all. E.g. the number of molecules in my left hand thumbnail, the number of hairs in Donald Trump's left nostril etc. So it is very plausibly possible for us to know that there exists some or other scientific question that will never in fact be resolved. But what we are interested in is whether we can know *of a specific empirical question*, which is currently unsolved, that it will remain forever unsolved? Compare: as we saw with discussion of the Church-Fitch result in the previous section, there is no problem knowing that there exists *some or other* fact that I am ignorant of. But it is impossible to know of a *specific* fact both that it is a fact (i.e. true) and that I am ignorant of it.

A number of prominent thinkers have counselled that we should never be too sure that some specific scientific question will remain forever unanswered. For example, Charles Darwin warned that:

> "…it is those who know little, and not those who know much, who so positively assert that this or that problem will never be solved by science." (Darwin *The Descent of Man*, 1871)

Likewise C. S. Pierce wrote:

> The history of science affords illustrations enough of the folly of saying that this, that, or the other can never be found out. Auguste Comte said that it was clearly impossible for man ever to learn anything of the chemical constitution of the fixed stars, but before his book had reached its readers the discovery which he had announced as impossible had been made. Legendre said of a certain proposition in the theory of numbers that, while it appeared to be true, it was most likely beyond the powers of the human mind to prove it; yet the next writer on the subject gave six independent demonstrations of the theorem" (Pierce, *Science & Immortality,* 1887)

And in his recent book *Human Compatible*, Stuart Russell recounts the story of how Lord Ernest Rutherford, perhaps the leading nuclear physicist of his day, publically stated in September 1933 that: 'Anyone who looks for a source of power in the transformation of atoms is talking moonshine.' But so then the Hungarian physicist Leo Szilard read this statement from Rutherford in the newspaper the next day, went for a walk to think about it and promptly invented the idea of neutron-induced chain reaction. Russell comments:

> “The problem of liberating nuclear energy went from impossible to essentially solved in less than twenty-four hours. Szilard filed a secret patent for a nuclear reactor the following year. The first patent for a nuclear weapon was issued in France in 1939.’ (Russell, 2019, p8)

The philosopher who has probably written most extensively on this topic is Nicholas Rescher. In a number of different works, Rescher argues against the idea that we could know that a specific scientific question will remain forever unsolved:

> “It is in principle infeasible for us to tell now how future science will answer present questions, or even what questions will figure on the question agenda of the future, let alone what answers they will engender. In this regard, as in others, it lies in the inevitable realities of our cognitive condition that the detailed nature of our ignorance is – for us at least – hidden away in an impenetrable fog of obscurity.” (Rescher, 2009a, 44)

> “The quest for scientific insolubilia is a delusion; no one can say in advance just what questions natural science can and cannot answer. Identifiable insolubilia have no place in an adequate theory of scientific inquiry.” (Rescher 2009b, 51)

To be clear: we are concerned with whether or not we might be able – at a particular time, in a particular state of scientific knowledge – to *identify* a specific instance of a forever unanswerable scientific question. This leaves entirely open that there may well *be* such unanswerable questions. Another important clarification that Rescher makes concerns the scope of ‘future science’. One obvious way that we might come to know that we humans will never answer some specific scientific question is if we learn that the human race will become extinct very shortly. So if we know that a meteor is going to destroy all life on Earth next week then we could know that various specific scientific questions will never be answered by humans. But so the interesting issue isn’t really about this sort of contingent, parochial limit on specifically human science or human knowers, but about *any* finite minded subject in our physical universe. So even if we humans go extinct, or the whole Solar system or whole galaxy becomes devoid of life etc., still there may well be other alien species with their own scientists. In what follows I will continue to say ‘we know’, ‘we will know’, etc. But by ‘we’ I mean any actual, physical knower, past, present or future, of any species in our universe. (So no omniscient gods, no merely possible knowers in other possible worlds, and no idealized rational agents.)

But so why should we believe – what we might call the ‘Pierce-Rescher Thesis’ – that we cannot ever know for sure of some specific scientific question that it will never be answered? We have some striking examples from the history of science where people have claimed that a question is unsolveable only to end up looking foolish when the question is solved soon after. But such examples are hardly decisive by themselves. Is there an actual argument why we should accept that it is futile to try to identify an unanswerable scientific question? Well, I think that Rescher supplies us with at least a fairly plausible reason. He emphasizes how new discoveries and theoretical revolutions in science create totally radical changes in our concepts and ontologies. We cannot possibly predict what future science will or will not be able to answer given that future scientific theories will almost certainly involve utterly new, radically different scientific concepts and ontologies that have not occurred to anyone yet.

And of course we cannot begin to imagine what these radically new, future scientific conceptions of the world will be like, since any act of imagination or prediction will be performed within the constraints of our current conceptual framework.

> "Clever though he unquestionably was Aristotle could not have pondered the issues of quantum electrodynamics. The scientific questions of the future are, at least in part, bound to be conceptually inaccessible to the inquirers of the present. The question of just how the cognitive agenda of some future date will be constituted is clearly irresolvable for us now. Not only can we not anticipate future discoveries now; we cannot even pre-discern the questions that will arise as time moves on and cognitive progress with it." (Rescher, 2005, 96)

This seems to me to provide at least *some* good reason to think that we can never know for sure of some specific, currently unsolved empirical question whether or not it will forever remain unsolved – for we just cannot really begin imagine what the science of the far future will and will not be able to answer.

4. AN INTERESTING CONTRADICTION

However, Rescher also goes on to draw the following conclusion:

> 'It is clear on this basis that the question "Are there nondecidable scientific questions that scientific inquiry will never resolve, even were it to continue *ad infinitum*" – the insolubilia question, as we may call it – is one that cannot ever possibly be settled in a decisive way. After all, how could we possibly establish that a question Q about some issue of fact will continue to be *raisable and unanswerable* in every future state of science, seeing that we cannot now circumscribe the changes that science might undergo in the future? **And since this is so, we have it that this question is, quite interestingly, self-instantiating: it is a question regarding an aspect of reality (of which, of course, science is a part) that scientific inquiry will never, at any specific state of the art, be in a position to settle decisively**." (Rescher, 2005, 96 – emphasis added)

Rescher here appears to have fallen into a quite interesting self-contradiction. For he is claiming both that we cannot now identify a specific scientific question that will remain forever unsolved and also to be providing an example of just such a question that science will never be able to settle. As we will soon see, avoiding this kind of contradiction establishes some limits on what we can know, limits that are, I suggest, somewhat similar to the Church-Fitch result.

So consider the following question – or we could call it a meta-question as it is a question about questions:

- $\mathbf{Q_{INS}}$: is there a specific yes/no question about a matter of empirical fact that we will never solve?

By ‘solve’ I just mean *know* the correct answer to. (So it is not enough to guess the correct answer or form a lucky true belief in the answer.) And by ‘specific’ I mean: can we *identify* a specific question – as opposed to knowing that there will be *some or other* question that is never answered. So to solve Q_{INS} we would either need to know *de re* of some specific empirical question that it will never be solved [YES answer], or know that there is no such question – i.e. know that every empirical question will eventually be solved [NO answer].

The sorts of considerations Rescher mentions, about the unimaginable nature of science in the far future, might well be thought to support the idea that we will never be able to solve Q_{INS} After all, these considerations make it implausible that at any specific point in time we will be able to identify a specific question that we know will remain unsolved throughout all future times and all future scientific theory-change – so it seems implausible that we can ever know that the answer to Q_{INS} is YES. But equally it seems implausible that we can know that the answer to Q_{INS} is NO, as it seems even more implausible that we will ever be able to establish that every empirical question *will* eventually be answered.

But it also seems very plausible that Q_{INS} is itself a specific yes/no question about a matter of empirical fact. For after all, the course of human scientific knowledge, or indeed the course of empirical knowledge amongst any species anywhere and anytime in the physical universe, is itself just a matter of contingent empirical fact. The answer is either: ‘YES – *this* particular empirical question will never be answered by anyone.’ Or ‘No – there is no empirical question that will not eventually be answered.’ So Q_{INS} is itself amongst the set of empirical questions. (This is surely intended by Rescher given his claim that the insolubilia question is ‘self-instantiating’ and his reminder that scientific inquiry is itself a part of reality that is thus open to scientific inquiry.)

So now suppose that we could *know* that Q_{INS} is itself insoluble – i.e. we could identify it as a specific example of an insolubilium – then we would know that the answer to Q_{INS} is “Yes, there is a specific empirical question that we will never solve”. Thus Q_{INS} would be solved. Contradiction. Thus we must reject the idea that we can know that Q_{INS} is insoluble and conclude instead that we cannot know that Q_{INS} is insoluble.

This line of thought seems to have an affinity with the Church-Fitch reasoning in the following respect[10]: the proposition (p & ¬Kp) can be true but it cannot be known since if, *per impossibile*, the proposition were known then it would be false. (Knowing the left-hand conjunct ensures that the right-hand conjunct is false.) Likewise, the following proposition could be true: Q_{INS} is insoluble. But again, if *per impossibile* this proposition were known then it would be false – since to know this proposition would be to know a solution to Q_{INS}.

In the next section I will regiment this line of thought into an argument with numbered, premises and also present another, closely related argument.

10 I am grateful to an anonymous referee for this journal for pressing me to say more about the comparison with the Church-Fitch theorem.

5. TWO ARGUMENTS

The line of thought presented in the previous section can be put into numbered premises as follows:

(1) Q_{INS} is itself a genuine, empirical, yes/no question.	[PREMISE]
(2) We know that Q_{INS} will never be solved	[SUPPOSE for *reductio*]
(3) Q_{INS} will never be solved	[from 2, factivity of K]
(4) We know of a specific, yes/no empirical question that it will never be solved	[from 1, 2]
(5) We know the answer to Q_{INS}	[from 4, meaning of Q_{INS}]
(6) We will never know that Q_{INS} will never be solved	[from 3, 5 by *reductio*]

Let me now make a few comments about this argument.

One might worry whether Q_{INS} can count as an *empirical* question if as a matter of logic it is unsolvable[11]. To allay this worry: firstly, notice that I have not suggested that Q_{INS} is insoluble *as a matter of logic*. In the previous sections I suggested that the Rescher's line of thought about the unimaginability of future scientific concepts supports the idea that we will never know a positive 'Yes' answer to Q_{INS}. On the other hand, I suggested, it seems very implausible that we will ever know a positive 'No' answer to Q_{INS}, since this would require knowing that every empirical question will eventually be answered. Together these two suggestions give us at least some reason to think that Q_{INS} is insoluble. However, this is not just a matter of logic – it depends on considerations about the imaginability of future science, etc. What I *do* claim to establish as a matter of logic is that: *we will never know* that Q_{INS} is insoluble. Secondly, as a more general point: contingent, empirical claims can be unknowable as a matter of logic. And so there can be empirical questions with correct answers that are unknowable as a matter of logic. Once more, a comparison with the Church-Fitch result may be helpful here. Consider the following proposition, which is one instance of the general schema (p & $\neg Kp$): 'It is raining and nobody knows it is raining'. This is presumably an empirical, contingent claim about the state of the actual, physical world. But as a matter of logic it cannot be known. And so the question: 'Is it the case that: it is raining and nobody knows it is raining?' should presumably count as an empirical question even if the correct answer turns out to be a proposition that cannot, as a matter of logic, be known.

Another possible worry is that the move to line (4), based on (1) and (2), involves a kind of intensional substitution that is not always legitimate. The inference is effectively of the following form: We know X is F, X is a Y, so we know that there is a Y which is F. And of course, from the fact that Lois knows Superman can fly and the fact that Superman is Clark Kent, it does not follow that Lois knows Clark Kent can fly. However, in the present context the move from (1) and (2) to (4) seems harmless insofar as we can legitimately take ourselves to know (1) – that Q_{INS} is a genuine empirical question. And this brings me to another important clarification: we should be clear that premise (1) is not just a trivial assumption, for it is not just trivial that Q_{INS} *is a genuine question*. After all, and in light of Rescher's warnings about the inscrutability of future science, it can clearly happen that as we gain more scientific knowledge and change our best theories and ontologies etc. we come to realise that a question is *not well defined*, i.e. contains a false presupposition or is somehow confused.

[11] Once more, I am very grateful to a referee for this journal for pressing me to clarify this point.

The classic example of a question with a false presupposition is the lawyer asking an innocent defendant: 'When did you stop beating your wife?' But we could also think of questions involving out-dated, obsolete theoretical terms: 'How dense is Phlogiston?', 'What is the velocity of the Earth through the lumineferous aether?"[12] And so we should at least allow for the *possibility* that in the future, after some great theoretical and conceptual revolution, we come to realise that Q_{INS} is somehow failing to be a genuine question after all. (In which case premise 1 would be false.) So then it could be that lines 2 and 3 of the argument are true, since Q_{INS} turns out to have NO genuine answer, and yet line 4 could be false since we have not yet identified a specific genuine empirical question that will never be solved. Which is all just to emphasize that this argument relies on the assumption that Q_{INS} is a genuine question (not a pseudo-question). To repeat: this is not just a trivial assumption. But neither, I think, is it implausible. It is not like the question involves reference to some highly theoretical scientific term or some part of our physical ontology that future science might well reject – e.g. a question about string theory or leptons or gravity waves etc. It is just a question about questions and it seems at least pretty unlikely we will eliminate questions from our theoretical vocabulary.

Let's turn now to a second, closely related line of thought. Having shown that we cannot know that Q_{INS} will never be answered, we might instead try to draw a weaker moral from the unpredictable, unimaginable nature of future scientific progress: that we at least cannot know that Q_{INS} *will* eventually be solved. For to do this we would have to somehow establish that either we *will* eventually identify a specific insoluble question or that we will know that all empirical questions will eventually be answered! But here too there is a lurking danger of self-contradiction that needs to be avoided. For consider the following question, which I will call QQ_{INS}:

- **QQ_{INS}:** Is Q_{INS} soluble? (I.e. will we eventually know the answer to Q_{INS}?)

[Notice: this is now a meta-meta-question – a question about a question about questions.]

The answer to QQ_{INS} must either be: YES, Q_{INS} is soluble (we will eventually answer it), or, NO, Q_{INS} is not soluble (we will never know the answer to Q_{INS}). And so to *know* the answer to QQ_{INS} we would either have to know that Q_{INS} is soluble or know that Q_{INS} is not soluble. Now, the conclusion of the previous argument was that we cannot know that Q_{INS} is insoluble. And so, assuming that the previous argument is indeed a valid deduction from known premises, we can know this conclusion:
I.e. We know (we cannot know that Q_{INS} is insoluble)
So now let's suppose (for *reductio*):
We know (we cannot know that Q_{INS} is soluble)
If we both know that we cannot know Q_{INS} is insoluble and we know that we cannot know that Q_{INS} *is* soluble, then we would know that we cannot know either of the two possible answers to QQ_{INS} So we would know that QQ_{INS} is an example of an insoluble question. Which would mean that we would after all know the answer to Q_{INS} – and so we would after all know that Q_{INS} is soluble. Contradiction! So we have to reject the assumption that we can

[12] This latter question – or rather pseudo-question – was what Michelson and Morley's famous experiment in 1887 was designed to answer. The experiment led to physicists eventually abandoning the whole idea of a medium through which light waves propagate.

know that we cannot know that Q_{INS} is soluble. Maybe it is *true* that we cannot know that Q_{INS} can be solved, but this is *not* something we can ever *know*.

Let's now also lay out this second line of thought as numbered, natural language premises:

(1) QQ_{INS} is a genuine yes/no question about a matter of empirical fact. [PREMISE]
(2) We know [we will never know that Q_{INS} will never be solved]. [PREMISE - previous proof]
(3) We know [we will never know that Q_{INS} will be solved]. [ASSUMPTION for *reductio*]
(4) We will never know that Q_{INS} will be solved. [From 3 by factivity of K]
(5) We know [we will never know that Q_{INS} will never be solved]
AND we know [we will never know that Q_{INS} will be solved]. [From 2 & 3]
(6) We know that QQ_{INS} will never be solved. [From 5]
(7) We know that a specific, yes/no empirical question will never be solved. [From 1 & 6]
(8) We know the answer to Q_{INS}. [From 7]
(9) We know that Q_{INS} will be solved. [From 8, plausible luminosity]
(10) We will never know [we will never know that Q_{INS} will be solved]. [From 3, 4, 9, by *reductio*]

Once more, as with the previous argument, the first premise makes the non-trivial assumption that QQ_{INS} is a genuine question. If it turned out to be a mere pseudo-question with no genuine answer, then line (6) could be true and yet line (7) false. Another thing that is worth commenting on is the move from line (8) to line (9), which relies on a kind of luminosity or KK principle. For line (8) just says that we know the answer to some question, whereas line (9) effectively says that *we know* we know the answer to that question. Now of course in general the KK principle is controversial and arguably false of actual, non-ideal human knowers. But in this specific case it is surely a pretty harmless assumption, given that we are considering the knowledge of all sentient beings in the universe ever, that if 'we' know the answer to Q_{INS} then we will also know that we know the answer – i.e. we will know that we have solved Q_{INS}.

Finally, I should make clear that although in the foregoing discussion I sometimes talked of what we *can* and *cannot* know, in presenting the pair of arguments in this section I have deliberately stated all of the premises in terms of what anyone *will* or *will not* ever know (anywhere, at anytime). This is partly because I want Q_{INS} and QQ_{INS} to be questions about matters of empirical fact and the modal notions of what is possible or impossible for us to know (at least in some senses of possibility) are perhaps less straightforwardly empirical. But more importantly, possibilities like this can *fail to aggregate.* It could be possible for me to be married to Jane and possible for me to be married to John and yet it is not possible for me to be both married to Jane and married to John. Likewise, perhaps it is possible for us to know X and possible for us to know Y and yet it is impossible for us to both know X and know Y. Indeed, this seems to be just what Heisenberg's famous Uncertainty Principle tells us. Beyond a certain degree of accuracy one can never know both the position and the velocity of a particle at the same time, though either quantity could individually be known to a greater degree of accuracy – for the more accurately one quantity is known, the less accurately the other can be known. These issues with non-aggregation are avoided if we stick to talking about actual knowledge by any actual subject anywhere and anytime in the actual universe.

6. CONCLUSION

In summary: we started with the plausible looking line of thought that we can never know for sure of some specific empirical question that it will never be answered since we cannot begin to imagine the future development of science, which will presumably involve radical conceptual revolutions and totally different ontologies. This line of thought might well incline us to think that the question Q_{INS} [is there a specific empirical question that will never be answered?] will never be solved. However, the idea that we can *know* that Q_{INS} will never be answered leads us into contradiction. So we cannot know that Q_{INS} will never be answered – a logical limit on what we can know. We then considered whether we might at least establish that we *cannot know* that Q_{INS} *will* eventually be solved. Again, the unpredictable, unimaginable shape of future science might seem to support this. But here also, the idea that we can *know* that we cannot know that Q_{INS} will eventually be solved also leads us into contradiction. So we cannot know that we cannot know that Q_{INS} will eventually be solved – another logical limit on our knowledge. These results are interesting and somewhat surprising, I hope, precisely because the Pierce-Rescher line of thought about the unpredictability of future science supports the idea that it is *true* that Q_{INS} will never be answered and so also supports the idea that we cannot know that Q_{INS} will be answered. But it turns out, on pain of contradiction, that we cannot *know* either of these things[13].

References

- Brogaard, B. (2016) "Ignorance and Incompetence: Linguistic Considerations" in Peels & Blauw (eds.) *The Epistemic Dimensions of Ignorance*, Cambridge University Press.
- Brogaard, B. & Salerno, J., (2002) "Clues to the Paradoxes of Knowability: Reply to Dummett and Tennant," *Analysis*, 62: 143–150
- Brogaard, B. & Salerno, J. (2006) "Knowability and a Modal Closure Principle," *American Philosophical Quarterly*, 43: 261–270
- Church, A. (2009) "Referee Reports on Fitch's 'A Definition of Value'," in Salerno (ed.) *New Essays on the Knowability Paradox*, Oxford: Oxford University Press.
- Darwin, C. (1871) *The Descent of Man, and Selection in Relation to Sex*, London: John Murray
- Edgington, D. (1985) "The Paradox of Knowability," *Mind*, 94: 557–568
- Edgington, D. (2010) "Possible Knowledge of Unknown Truth," *Synthese*, 173: 41–52
- Fine, K. (2018) 'Ignorance of Ignorance', *Synthese* 195 (9): 4031-4045
- Fitch, G. (1963) "A Logical Analysis of Some Value Concepts," *The Journal of Symbolic Logic*, 28: 135–142.
- Hart, W. D. & McGinn, C., (1976) "Knowledge and Necessity," *Journal of Philosophical Logic*, 5: 205–208
- Jago, M. (2010) "Closure on Knowability," *Analysis*, 70: 648–659
- Kelp, C. & Pritchard, D. (2009) "Two Deflationary Approaches to Fitch-Style Reasoning," in Salerno (ed.) *New Essays on the Knowability Paradox*, Oxford: Oxford University Press.

[13] Earlier versions of this paper were presented at events in Trondheim and in Luxembourg. I am grateful to the audiences on those occasions for their questions and comments. I am deeply indebted to Julien Dutant who generously provided written comments that greatly improved the paper. Needless to say, he is not responsible for any errors that remain nor should he be assumed to endorse any of the views advanced in the paper. Many thanks also to Dov Gabbay, Frank Hofmann, Aleks Knoks, Jonathan Knowles, Reka Markovich, Anders Nes and Filippo Vindrola. Finally, I am very grateful to the anonymous referees for this journal, whose reports also substantially improved this paper.

- Kieu, T. (2003). "Quantum Algorithm for the Hilbert's Tenth Problem". *International Journal of Theoretical Physics* 42 (7): 1461–1478.
- Lucas, J. (1961) "Minds, Machines and Gödel", *Philosophy* 36 (April–July): 112–127
- Peels, R. (2010) "What is Ignorance?" *Philosophia* 38 (1):57-67.
- Peels, R. (2012) "The New View of Ignorance Undefeated", *Philosophia* 40 (4):741-750
- Peirce, C. S. (1887) 'Science & Immortality', *The Christian Register* 66: 214
- Penrose, R. (1989) *The Emperor's New Mind*, Oxford University Press.
- Rescher, N. (2005) *Epistemic Logic*, University of Pittsburgh Press.
- Rescher , N. (2009a) *Ignorance*, University of Pittsburgh Press.
- Rescher, N (2009b) *Unknowability*, Lexington Books.
- Russell, S. (2019) *Human Compatible*, Viking Press Books.
- Salerno, J. (2009) ""Knowability Noir: 1945–1963," in Salerno (ed.) *New Essays on the Knowability Paradox*, Oxford: Oxford University Press.
- Searle, J. (1980) 'Minds, Brains and Programs', *Behavioral and Brain Sciences*, 3: 417–57
- Tennant, N. (1997), *The Taming of the True*, Oxford: Oxford University Press
- Tennant, N, (2010) "Williamson's Woes," *Synthese*, 173: 9–23
- Williamson, T. (1987), "On the Paradox of Knowability," *Mind*, 96: 256–61
- Williamson, T. (2000), *Knowledge and its Limits*, Oxford: Oxford University Press
- Ziegler, M. (2005). "Computational Power of Infinite Quantum Parallelism". *International Journal of Theoretical Physics* 44 (11): 2059–2071.