Mechanistic reasoning and the problem of masking

Michael Wilde[1]

Abstract

At least historically, it was common for medical practitioners to believe causal hypotheses on the basis of standalone mechanistic reasoning. However, it is now widely acknowledged that standalone mechanistic reasoning is insufficient for appropriately believing a causal hypothesis in medicine, thanks in part to the so-called problem of masking. But standalone mechanistic reasoning is not the only type of mechanistic reasoning. *When exactly then is it appropriate to believe a causal hypothesis on the basis of mechanistic reasoning?* In this paper, I argue that it has proved difficult to provide a satisfying answer to this question. I also argue that this difficulty is predicted by recent work in knowledge-first epistemology. I think this shows that recent work in epistemology has important implications for practice in the philosophy of science. It is therefore worth paying closer attention in the philosophy of science to this recent work in knowledge-first epistemology.

Keywords Mechanistic reasoning · The problem of masking · Knowledge-first epistemology · The philosophy of science

1 Introduction

At least historically, it was common for medical practitioners to believe causal hypotheses on the basis of standalone mechanistic reasoning. A famous example involves the mechanistic reasoning that led to believing that bloodletting was an effective therapy for a variety of diseases (Howick 2011b: pp. 136–137). However, it is now widely acknowledged that such mechanistic reasoning is insufficient for appropriately believing a causal hypothesis in medicine (Russo and Williamson 2007: 162). After coming to believe a causal hypothesis on the basis of such reasoning, it remains possible that the putative cause does not make an overall difference to the relevant effect, because there may exist further so-called *masking mechanisms* by which the putative cause also cancels out any difference made by way of

---

✉ Michael Wilde
  m.e.wilde@kent.ac.uk

1 Department of Philosophy, University of Kent, Canterbury, Kent CT2 7NF, United Kingdom

the established mechanism. This has been called *the problem of masking* by Phyllis Illari ([2011](#)).

However, *standalone* mechanistic reasoning is not the only type of mechanistic reasoning in medicine. Some other type of mechanistic reasoning may yet be sufficient for appropriately believing a causal hypothesis in medicine. *When exactly then is it ever appropriate to believe a causal hypothesis on the basis of mechanistic reasoning?* One answer has been put forward by Jeremy Howick ([2011a](#), [b](#)). He thinks that it is appropriate to believe a causal hypothesis on the basis of mechanistic reasoning if and only if the reasoning counts as *high-quality mechanistic reasoning*, where high-quality mechanistic reasoning meets certain completeness and complexity conditions.

In this paper, I argue that high-quality mechanistic reasoning is also insufficient for appropriately believing a causal hypothesis in medicine. More generally, I argue that it has proved difficult to provide a satisfying answer to the following *principal question*: *When is it appropriate to believe a causal hypothesis on the basis of mechanistic reasoning?* I maintain that this difficulty is predicted by recent work in *knowledge-first epistemology* (Williamson [2000](#)). According to this work, it is simply not possible to provide a satisfying answer to the principal question. I think this shows that recent work in epistemology has important implications for practice in the philosophy of science. And I argue that it is therefore worth paying closer attention in the philosophy of science to recent work in knowledge-first epistemology. I hope that this helps to motivate a novel approach in the philosophy of science, an approach which insists upon a mutually beneficial interaction with this recent work in epistemology.

2 Mechanistic reasoning and the problem of masking

Historically, it was common for medical practitioners to believe causal hypotheses on the basis of *standalone mechanistic reasoning*, where standalone mechanistic reasoning involves believing a causal hypothesis solely on the basis of an established mechanism linking the putative cause and effect (Howick [2011b](#): p. 128).[1] Indeed, a classic defence of standalone mechanistic reasoning in medicine was provided by Claude Bernard ([1865](#)). Of course, this reliance on standalone mechanistic reasoning was also not without its critics (cf. Howick [2016](#)). Nevertheless, it was still common for medical practitioners to rely on such reasoning in order to believe causal hypotheses about the effectiveness of medical interventions at bringing about certain health outcomes (Howick [2011b](#): pp. 154–157). For example, on the basis of established mechanisms concerning vomiting and choking, it was believed that

---

[1] A number of different accounts of the metaphysics of a mechanism have been proposed in the literature. A nice overview of these accounts is given by Jeffrey Aronson ([2020](#)). Given that this paper is more concerned with *epistemological* rather than *metaphysical* issues, I simply take it that a mechanism is an organization of entities and activities that is responsible for some regular phenomenon (Machamer et al. [2000](#)).

putting babies to sleep on their fronts was effective at reducing infant mortality due to sudden infant death syndrome. But observational studies later showed that infant mortality rates were in fact reduced by placing babies to sleep on their backs (Howick 2011b: pp. 137–138). Such failures led to the recognition that standalone mechanistic reasoning is insufficient for appropriately believing a causal hypothesis in medicine (cf. Solomon 2015: pp. 121–124).

Why exactly is it inappropriate to believe a causal hypothesis on the basis of standalone mechanistic reasoning? In part, it is because such mechanistic reasoning suffers from the problem of masking. Phyllis Illari (2011) discusses the following example as an instance of the problem of masking in action. Although there is an established mechanism that links increased exercise to weight-loss by burning calories, this basis is not enough to appropriately believe the hypothesis that increased exercise causes weight-loss, since it remains possible that there exist further masking mechanisms by which increased exercise leads to weight-gain, for example, by increasing appetite (cf. Steel 2008: p. 68). In such a case, in order to appropriately believe the hypothesis that an intervention caused some health outcome, it was not enough to have believed the hypothesis on the basis of evidence that established only that there is a particular mechanism by which the intervention makes a difference to that health outcome. After believing the causal hypothesis on the basis of such standalone mechanistic reasoning, it remains possible that the intervention does not make the relevant *overall difference* to the health outcome, because it remains possible that there exists some masking mechanism by which the intervention also cancels out any difference made by way of the particular established mechanism. This is the problem of masking (Illari 2011: pp. 145–147).

Is the problem of masking really a problem for standalone mechanistic reasoning? An analysis of causality in terms of mechanisms has been given by Stuart Glennan (1996). According to this analysis, a medical intervention and some health outcome are causally connected if and only if they are linked by some appropriate mechanism. In this sense, establishing a mechanism linking an intervention and some health outcome simply is establishing a causal hypothesis involving the intervention and that outcome. It might therefore be thought that an established mechanism does provide a basis for appropriately believing a causal hypothesis. In the weight-loss case, evidence that establishes a mechanism linking increased exercise to weight-loss thereby establishes the hypothesis that increased exercise causes weight-loss. Given this analysis, it might be argued that this established mechanism is enough to provide a basis for appropriately believing that increased exercise causes weight-loss.

However, two types of causal hypothesis have been distinguished by Christopher Hitchcock (2001). The problem of masking depends upon a similar distinction (cf. Illari 2011: p. 146). On the one hand, there are causal hypotheses that claim a *component effect* for a mechanism. They claim only that there is a particular mechanism by which the putative cause makes a difference to the effect, even if there may be other mechanisms that preclude the putative cause from making an overall difference to the effect. On the other hand, there are hypotheses that claim a *net effect*. They claim that the putative cause does make an overall difference to the effect, where this overall difference takes into account

all of the mechanisms linking the cause and effect. In the medical context, the point of believing a causal hypothesis is often in order to predict the effects of interventions, for example, the effect of the increased exercise on weight-loss (Russo and Williamson 2011: p. 567). Given this, the relevant causal hypothesis does not concern only the component effect, since a medical intervention is beneficial only if it makes an overall difference to a health outcome, rather than just a difference along a particular mechanism. In other words, establishing that there is a particular mechanism by which the intervention makes a difference to the health outcome may be enough to appropriately believe a hypothesis about a *component effect* of the intervention. However, this is not enough to appropriately believe the hypothesis relevant in medicine, which requires also establishing that the intervention has a *net effect* on the relevant health outcome.

How significant is the problem of masking? Historically, the problem of masking has proved a significant difficulty in medicine. A number of cases in which a harmful or non-beneficial intervention was proposed on the basis of an established mechanism have been listed by Jeremy Howick (2011b: pp. 154–157). In many of these cases, the problem of masking provides a plausible explanation of these unsuccessful interventions. One example here is the controversy regarding the benefits of screening for breast cancer with mammography (Solomon 2015: pp. 210–217). Miriam Solomon points out that some medical researchers recommended mammography on the basis of the established mechanism that breast cancer is a progressive disease, and that therefore the earlier that the cancer is detected by mammography then the more effectively the disease may be treated (Solomon 2015: p. 212). However, certain systematic reviews of clinical studies have suggested that mammography leads to no overall reduction in breast cancer mortality in certain age ranges [e.g., Gøtzsche and Olsen (2000)]. This is sometimes called *the mammography paradox* (Retsky et al. 2005). One possible explanation of this paradox is that any difference to breast cancer mortality made by screening is being cancelled out by the action of some unknown masking mechanism. Indeed, Solomon mentions the hypothesis that 'the injury caused by biopsy produces increased blood flow and stimulates angiogenesis in the area, and that this nourishes the cancer and enables it to spread' (Solomon 2015: p. 212). This hypothesis has been considered elsewhere as an explanation of the mammography paradox (Baum et al. 2005; Retsky et al. 2005).

In conclusion, the problem of masking is a *real* and *significant* problem for standalone mechanistic reasoning in medicine. It is now widely recognized that standalone mechanistic reasoning is therefore insufficient for appropriately believing a causal hypothesis in medicine. This fact is perhaps best reflected by the positioning of standalone mechanistic reasoning at the bottom of the evidence hierarchy of evidence-based medicine (Clarke et al. 2014). However, standalone mechanistic reasoning is not the only type of mechanistic reasoning. Some other type of mechanistic reasoning may yet be sufficient for appropriately believing a causal hypothesis, as long as that reasoning satisfies additional constraints to standalone mechanistic reasoning.

3 High-quality mechanistic reasoning

When is it ever appropriate to believe a causal hypothesis on the basis of mechanistic reasoning? Given the preceding discussion, a natural answer is that it is appropriate to believe a causal hypothesis on the basis of mechanistic reasoning if and only if the mechanistic reasoning overcomes the problem of masking. However, Howick has pointed out an additional problem for certain instances of standalone mechanistic reasoning: 'the probabilistic nature of mechanistic relationships presents a serious problem for inferring the overall effect of an intervention from knowledge about mechanisms' (Howick 2011a: p. 937). In other words, the stochastic nature of some mechanisms makes it difficult to infer the overall effectiveness of a medical intervention on the basis of standalone mechanistic reasoning, even if that reasoning overcomes the problem of masking. Let us call this additional problem *the problem of complexity*.

Howick therefore thinks that it is appropriate to believe a causal hypothesis on the basis of mechanistic reasoning if and only if the reasoning counts as *high-quality mechanistic reasoning*, where high-quality mechanistic reasoning meets the following conditions:

(1)   The knowledge of mechanisms upon which the mechanistic reasoning is based is *not incomplete*, i.e., there are no obvious gaps in our knowledge of the inferential chain linking the intervention and the patient-relevant outcome.

(2)   The probabilistic and complex nature of the mechanisms are explicitly taken into account when inferring from mechanisms to any claims that a particular intervention has a patient-relevant benefit (Howick 2011b: p. 144)

Let us call the first condition *the completeness condition*, and the second condition *the complexity condition*. The completeness condition is intended to address the problem of masking. And the complexity condition is intended to address the problem of complexity (Howick 2011a: pp. 937–938).

I want to argue that high-quality mechanistic reasoning is also *insufficient* for appropriately believing a causal hypothesis in medicine. I think that there is a counterexample showing that it is possible to inappropriately believe a causal hypothesis on the basis of high-quality mechanistic reasoning. One such counterexample concerns the effect of hormone replacement therapy on coronary heart disease (cf. Howick (2011a: pp. 927–930) and Solomon (2015: pp. 123–124)). In this case, there was extensive mechanistic evidence that hormone replacement therapy reduced the risk of complications due to blood clotting by lowering the lipid concentration of the blood (Mendelsohn and Karas 1999). In other words, the reasoning in this case meets the completeness condition, because the extensive evidence meant that there were no obvious gaps in the knowledge of the relevant mechanisms. Moreover, in this case there was epidemiological evidence that provided evidence of a net effect in order to account for the probabilistic and complex nature of the mechanisms involved in the reasoning [cf. Petitti et al. (1987) and Stampfer and Colditz (1991)]. In other words, the reasoning in this case also meets the complexity condition on

high-quality mechanistic reasoning. However, it was still not appropriate on this basis to believe that the hormone replacement therapy lowers the risk of coronary heart disease. Indeed, it was surprising to find that a large randomized trial later suggested that it was *false* that hormone replacement therapy lowered the risk of coronary heart disease (Rossouw et al. 2002). I think that this surprise was partly the result of there being no obvious gaps in the available mechanistic knowledge. Here it was the *non-obvious* gaps that stood in the way of appropriately believing the causal hypothesis. And it was this surprise that led to further investigation into a non-obvious potential masking mechanism involving age-dependent changes in vascular pathology (Dubey et al. 2005). Given this, it looks like it is possible to meet the proposed conditions on high-quality mechanistic reasoning without thereby appropriately believing the causal hypothesis, thanks to the possibility of *non-obvious* masking mechanisms.

Of course, Howick himself introduced this hormone replacement therapy case as an example of failed mechanistic reasoning (Howick 2011a: pp. 928–930). He therefore acknowledges that the mechanistic reasoning in this example was not enough to appropriately believe the hypothesis that hormone replacement therapy lowered the risk of coronary heart disease. Given this, Howick will likely maintain that in this case the mechanistic reasoning simply did not meet the conditions on high-quality mechanistic reasoning. *How exactly did the reasoning fail to meet those conditions?* One suggestion would be to maintain that in this case there still remained some *obvious gaps* in the relevant mechanistic knowledge. But it would then become difficult to explain why it was *surprising* to learn that it was false that hormone replacement therapy lowered the risk of coronary heart disease. A better suggestion therefore would be to amend the proposed conditions on high-quality mechanistic reasoning so that they require not only that there are no *obvious* possible masking mechanisms, but also that there are no *non-obvious* possible masking mechanisms.

In effect, the suggestion here is to maintain that it is appropriate to believe a causal hypothesis on the basis of mechanistic reasoning if and only if that reasoning counts as *super-high-quality mechanistic reasoning*, where super-high-quality mechanistic reasoning meets the following conditions:

(1) The knowledge of mechanisms upon which the mechanistic reasoning is based is *not incomplete*, i.e., there are *neither obvious nor non-obvious gaps* in our knowledge of the inferential chain linking the intervention and the patient-relevant outcome.
(2) The probabilistic and complex nature of the mechanisms are explicitly taken into account when inferring from mechanisms to any claims that a particular intervention has a patient-relevant benefit (cf. Howick 2011b: p. 144)

However, the problem now is that it is difficult to carry out such super high-quality mechanistic reasoning, since it is difficult to avoid such *non-obvious gaps* in our knowledge of the relevant mechanisms. Here is Miriam Solomon:

A general problem with mechanistic accounts is that they are typically incomplete, although they often give an *illusion* of a complete, often linear, narra-

tive. Incompleteness is the consequence of there being mechanisms underlying mechanisms, mechanisms inserted into mechanisms, background mechanisms that can fill out the mechanistic story, and mechanisms that can hijack regular mechanisms. That is, there is complex interaction of multiple mechanisms in a chaotic and multidimensional system. There are possible hidden mechanisms everywhere in mechanistic stories, despite an easy impression of narrative or causal completeness. (Solomon 2015: pp. 131–132).

In other words, if it is appropriate to believe a causal hypothesis by mechanistic reasoning only if that reasoning is super-high-quality mechanistic reasoning, it would then become next to impossible to appropriately believe a causal hypothesis by mechanistic reasoning, thanks to the extent and complexity of the mechanisms involved in most medical conditions. Of course, Howick may simply accept that it is next to impossible to appropriately believe a causal hypothesis in medicine by mechanistic reasoning. Indeed, Howick acknowledges 'there is much to stand in the way of mechanistic reasoning being of high quality since there are limits to our knowledge of bodily mechanisms and their interactions' (Howick 2011a: p. 939).

However, I want to argue that there is an additional problem with the conditions proposed by super-high-quality mechanistic reasoning. I think another counterexample shows that the conditions on super-high-quality mechanistic reasoning are not *necessary* for appropriately believing a causal hypothesis. Here is an example put forward by Howick (2011a, b). It was long ago established that there is a mechanism linking a reduction in the size of obstructions in the airway to improved respiratory function. More recently, it was established that a type of radiotherapy leads to a reduction in the size of large nodular goiters, which are an enlargement of the thyroid gland (Nielsen et al. 2006). Altogether then, it had been established that there is a mechanism by which radiotherapy improves respiratory function in patients with large nodular goiters, since large nodular goiters are simply a type of obstruction in the airway. Moreover, Howick maintains that in this case the conditions on high-quality mechanistic reasoning were satisfied: 'there are no obvious gaps in the mechanistic knowledge linking the intervention with the patient-relevant outcome, and the possibility of serious adverse events (complexity) was made unlikely by the clinical trials in the various stages of the inferential chain' (Howick 2011b: p. 144). He therefore thinks that this is one of the presumably rare cases in medicine where the mechanistic reasoning was sufficient for appropriately believing a causal hypothesis, namely, the hypothesis that radiotherapy improves respiratory function in patients with large nodular goiters: 'Mechanistic reasoning should therefore allow us to conclude that radiotherapy will improve respiratory function' (Howick 2011b: pp. 144–145).

I agree that this radiotherapy case provides a convincing example of appropriately believing a causal hypothesis on the basis of mechanistic reasoning. However, I don't think that the reasoning in the radiotherapy case meets the conditions on super-high-quality mechanistic reasoning. Indeed, an appeal is made only to the conditions on *high-quality mechanistic reasoning*. Moreover, the researchers in this case believed that it was still possible that there existed non-obvious masking mechanisms by which the radiotherapy worsened respiratory function, for example, by

inducing a swelling of thyroid tissue (Nielsen et al. 2006: pp. 1481–1482, Bonnema et al. 2008: p. 3981). As a result, these researchers also thought that the mechanistic reasoning available at the time did not provide a basis for appropriately believing that the radiotherapy improves respiratory function. Indeed, they were so concerned about the possibility of such non-obvious mechanisms that they believed it was only appropriate to believe that causal hypothesis once an additional randomized controlled trial had been carried out, where this trial systematically performed tracheal imaging and pulmonary function tests in order to provide a direct measure of the *net effect* of the radiotherapy on respiratory function (Bonnema et al. 2007, 2008). Of course, Howick's point is that this later randomized trial was unnecessary, presumably because although *non-obvious* masking mechanisms were still *possible* in this case, they were sufficiently *improbable* given the available evidence that their mere possibility failed to preclude appropriately believing the causal hypothesis (Howick 2011a: p. 938). Indeed, if one is committed to the idea that such improbable possibilities do in fact preclude appropriately believing a hypothesis, then it is likely that one is also committed to the idea that it is appropriate to believe a hypothesis only on the basis of evidence that logically entails that hypothesis. And this latter idea is surely overly demanding. Given this, the radiotherapy case is plausibly an example where it was appropriate to believe a causal hypothesis by mechanistic reasoning even though that reasoning failed to meet the conditions on super-high-quality mechanistic reasoning, thanks to the remaining possibility of non-obvious masking mechanisms. In other words, the conditions proposed by super-high-quality mechanistic reasoning are not *necessary* for appropriately believing a causal hypothesis on the basis of mechanistic reasoning.

Howick is perhaps more charitably interpreted then as maintaining that it is appropriate to believe a causal hypothesis by mechanistic reasoning if and only if the reasoning sits somewhere between high-quality mechanistic reasoning and super-high-quality mechanistic reasoning. *But where exactly?* A satisfying answer to this question will need to propose conditions on appropriately believing a causal hypothesis of the *apposite* logical strength, namely, strong enough to preclude appropriately believing the causal hypothesis in cases like the hormone replacement therapy example, but weak enough to allow for appropriately believing the causal hypothesis in cases like the radiotherapy example. However, that is likely to continue to prove a difficult task, given the poor track record in attempting to specify such conditions.

Where does all this leave us? At this stage, the take-home message is that it has proved difficult to provide a satisfying answer to the principal question: *When is it appropriate to believe a causal hypothesis on the basis of mechanistic reasoning?* On the one hand, the answer that appealed to high-quality mechanistic reasoning was unsatisfying, since the proposed conditions on high-quality mechanistic reasoning were not *sufficient* for appropriately believing a causal hypothesis. Those conditions were not strong enough to preclude appropriately believing the causal hypothesis in the hormone replacement therapy case. On the other hand, the answer that appealed to super-high-quality mechanistic reasoning was also unsatisfying, since the conditions on super-high-quality mechanistic reasoning were not *necessary* for appropriately believing a causal hypothesis. Those conditions were not weak enough to allow for appropriately believing the causal hypothesis in the radiotherapy case. A

satisfying answer to the principal question will instead propose conditions on appropriately believing a causal hypothesis of the apposite logical strength. But specifying such conditions is likely to continue to prove a difficult task. Of course, one way to complete this task would be by brute force, that is, to maintain that it is appropriate to believe a causal hypothesis on the basis of mechanistic reasoning if and only if that reasoning is a way of appropriately believing the causal hypothesis. However, the result is another unsatisfying answer to the principal question, but this time not because the answer proposes conditions of the inapposite logical strength, but rather because the conditions it proposes are simply *circular*. In sum, it has proved difficult to provide a satisfying answer to the principal question.

4 Knowledge-first epistemology

I think that the difficulty with providing a satisfying answer to the principal question is predicted by recent work in the knowledge-first epistemology advocated in part by Timothy Williamson (2000).

Here is a quick overview of this work. A standard way of thinking in epistemology is from within a *belief-first framework*. This framework takes the presence of beliefs as unexplained, and puts forward specific conditions on belief in an attempt to explain other epistemological phenomena. In particular, one programme within the belief-first framework consists of attempts to appeal to the notion of belief in providing necessary and sufficient conditions for knowledge, where these conditions are specified without making reference to knowledge. In other words, this programme attempts to provide a non-circular analysis of knowledge. And perhaps the historically most famous example of such a proposed analysis of knowledge is the so-called *tripartite account*, according to which a subject knows a proposition p if and only if they have a justified true belief that p [cf. Gettier (1963: p. 121)].

Now, the problem with this programme is that it has proved difficult to provide such an analysis of knowledge. Arguably, all proposed analyses of knowledge have been subject to counterexamples. On the one hand, for some proposed analyses in terms of necessary and sufficient conditions, it has been possible to show that a belief may meet the proposed conditions without counting as knowledge. Indeed, Gettier showed that the tripartite account of knowledge was incorrect, by presenting counterexamples to the effect that a subject's justified true belief that p was not *sufficient* for that subject's knowing that p (Gettier 1963). On the other hand, when the proposed conditions are made sufficiently strong enough to exclude those counterexamples, it has been possible to show that those conditions are then overly strong, since they also exclude ordinary cases of knowledge from counting as cases of knowledge. That is, there are counterexamples that show that the revised conditions are just not *necessary* for knowing that p (Williamson 2000: p. 4). It has therefore proved difficult to provide an analysis of knowledge in terms of necessary and sufficient conditions of the apposite logical strength. Moreover, Williamson maintains that '[o]ther analyses are circular rather than false; if someone insists that knowledge *is* justified true belief on an understanding of ''justified'' strong enough to exclude Gettier cases but weak enough to include everyday empirical knowledge,

the problem is likely to be that no standard of justification is supplied independent of knowledge itself' (Williamson 2000: p. 4). In other words, it is likely that any proposed analysis of knowledge either has the inapposite logical strength, or else it has the apposite logical strength, but thanks only to being a *circular* analysis.

Some have therefore drawn the moral that it is just not possible to provide non-circular necessary and sufficient conditions for knowledge (Williamson 2000: pp. 27–33). Instead, they think about epistemology from within a *knowledge-first framework*. They recommend taking *knowledge* rather than *belief* as unexplained, and they then attempt to explain other epistemological phenomena in terms of knowledge. For example, an attempt can be made to appeal to the notion of knowledge in explaining what makes it appropriate to believe a hypothesis: it is appropriate to believe a hypothesis on the basis of an instance of reasoning if and only if that reasoning is a way of *knowing* the hypothesis. This is a version of a so-called *knowledge norm of belief*.[2]

One version of the knowledge-first framework is therefore committed to the following tenets: (i) that knowledge is the norm of belief; (ii) that it is not possible to provide non-circular necessary and sufficient conditions for knowledge. It follows from these two tenets that it is just not possible to provide a satisfying answer to the question: *When is it appropriate to believe a hypothesis?* According to the first tenet, it is appropriate to believe a hypothesis by some instance of reasoning if and only if the reasoning is a way of *knowing* the hypothesis. But this preliminary answer is arguably not a *satisfying* answer to the question. It immediately prompts the further question: *When exactly is an instance of reasoning a way of knowing a hypothesis?* A satisfying answer will therefore involve providing non-circular necessary and sufficient conditions for knowing a hypothesis. However, the second tenet then implies that it is just not possible to give such a satisfying answer to this further question, because the second tenet maintains that it is not possible to provide non-circular necessary and sufficient conditions for knowledge. It is just not possible to provide a satisfying answer to the question: *When is it appropriate to believe a hypothesis?* Instead, the knowledge-first approach counsels us ultimately to rest content with the preliminary answer to the question: it is appropriate to believe a hypothesis by some instance of reasoning if and only if the reasoning is a way of *knowing* the hypothesis.

According to this knowledge-first approach, it is likewise not possible to give a satisfying answer to the principal question: *When is it appropriate to believe a causal hypothesis on the basis of mechanistic reasoning?* By the first tenet, it is appropriate to believe a causal hypothesis on the basis of mechanistic reasoning if and only if that reasoning is a way of knowing that causal hypothesis. Again, this preliminary answer is arguably unsatisfying. It prompts the further question: *When exactly is an instance of mechanistic reasoning a way of knowing a causal hypothesis?* A satisfying answer will therefore involve providing non-circular necessary and

---

[2] A full discussion of the knowledge-first approach is beyond the scope of this paper. Further discussion is given in Greenough and Pritchard (2009), McGlynn (2014). And further discussion of the knowledge norm for belief in particular is given by Williamson (2000), Sutton (2007), Littlejohn (2013), Hawthorne and Srinivasan (2013).

sufficient conditions for knowing a hypothesis by mechanistic reasoning. However, the second tenet of this knowledge-first approach again implies that it is not possible to give such a satisfying answer. It is just not possible to provide non-circular necessary and sufficient conditions for knowing a causal hypothesis by mechanistic reasoning. As a result, this knowledge-first approach predicts that any attempt to provide such an answer to the principal question will prove unsatisfying.

Of course, this prediction was borne out in the above discussion concerning mechanistic reasoning. It there proved difficult to provide a satisfying answer to the principal question: *When is it appropriate to believe a causal hypothesis on the basis of mechanistic reasoning?* In particular, the appeal to high-quality mechanistic reasoning was one attempt to provide a satisfying answer to this question, because it attempted to provide necessary and sufficient conditions for appropriately believing a causal hypothesis by mechanistic reasoning. It maintained that it is appropriate to believe a causal hypothesis by mechanistic reasoning if and only if the reasoning is an instance of high-quality mechanistic reasoning. But it was argued above that high-quality mechanistic reasoning is not *sufficient* for appropriately believing a causal hypothesis, because it is possible to meet those conditions without appropriately believing the relevant causal hypothesis, for example, in the hormone replacement therapy case. An attempt was then made to strengthen the conditions proposed by high-quality mechanistic reasoning by appealing instead to super-high-quality-mechanistic reasoning. But this appeal to super-high-quality mechanistic reasoning also failed to provide a satisfying answer to the principal question. In particular, it was argued above that super-high-quality mechanistic reasoning is not *necessary* for appropriately believing a causal hypothesis by mechanistic reasoning, since it is possible to appropriately believe a causal hypothesis by mechanistic reasoning without meeting those conditions, for example, in the radiotherapy case. Of course, such counterexamples were inevitable from within a knowledge-first framework. In order to provide an answer to the principal question of the apposite logical strength, a brute force approach was then entertained, according to which it is appropriate to believe a causal hypothesis on the basis of mechanistic reasoning if and only if that reasoning is a way of appropriately believing the causal hypothesis. However, this also provided an unsatisfying answer to the principal question, but this time not because the conditions it proposed were prone to counterexamples, but rather because it avoided those counterexamples only by proposing conditions that were simply *circular*.

I think this shows that recent work in epistemology has important implications for the philosophy of science. In this case, it makes clear a potential blind alley; if the knowledge-first framework is correct, then at least one prominent activity within the philosophy of science was misguided, namely, the attempt to provide non-circular necessary and sufficient conditions for appropriately believing a causal hypothesis on the basis of mechanistic reasoning. Moreover, this activity has serious consequences: if one wrongly believes that meeting a certain set of non-circular conditions is sufficient for appropriately believing a causal hypothesis, then one may end up inappropriately believing that hypothesis when they believe those conditions to be satisfied. For example, it was argued above that if one wrongly believes that high-quality mechanistic reasoning is sufficient for appropriately believing a causal

hypothesis in medicine, then one may end up inappropriately believing a causal hypothesis about the effectiveness of hormone replacement therapy in lowering the risk of coronary heart disease. And if one wrongly believes that super-high-quality mechanistic reasoning is necessary for appropriately believing a causal hypothesis on the basis of mechanistic reasoning, then one may end up inappropriately *not* believing a causal hypothesis about the effectiveness of radiotherapy at improving respiratory function. Instead, the knowledge-first approach counsels us to ultimately rest content with the preliminary answer to the principal question: it is appropriate to believe a causal hypothesis on the basis of mechanistic reasoning if and only if that reasoning is a way of knowing that causal hypothesis.

It might be objected that the knowledge-first approach therefore wrongly recommends giving up on trying to answer the principal question. *Surely, it's a mistake to give up on trying to answer a question of such importance to medical practice.* However, I don't think the knowledge-first approach really does recommend giving up on the principal question. Rather, it just counsels us ultimately not to expect a completely *satisfying* answer to that question. In effect, it just recommends for philosophers of science to manage their expectations.[3]

Now, it might also be objected that the attempt to provide non-circular necessary and sufficient conditions for appropriately believing a causal hypothesis is not exactly a blind alley, because the attempt has resulted in specifying certain *necessary* conditions for appropriately believing a causal hypothesis on the basis of mechanistic reasoning. In particular, it is appropriate to believe a causal hypothesis on the basis of mechanistic reasoning only if that reasoning at least meets the completeness and complexity conditions. And such conditions may be appealed to in providing an explanation of why an instance of mechanistic reasoning was insufficient for appropriately believing a causal hypothesis. However, at best this provides a partial explanation, because there is more to be said about why exactly those conditions are necessary for appropriately believing a causal hypothesis. Anyway, such partial explanations are also available on the knowledge-first approach, because it remains possible to specify necessary conditions on knowing a causal hypothesis, and such conditions may also be appealed to in explaining why an instance of mechanistic reasoning was insufficient for appropriately believing a causal hypothesis. One possible advantage of the knowledge-first approach here is that it can also provide the more complete explanation, since it can explain why a given condition is necessary for appropriately believing a causal hypothesis, namely, because it is a necessary condition for *knowing* the causal hypothesis.

A more central objection is that the knowledge-first approach is itself misguided. In particular, it might be objected that knowledge is not the norm of belief. One worry here is that one's knowledge is not *transparent* or *luminous*, in the sense that it is possible to mistake the extent of one's knowledge (Williamson 2000). It therefore becomes difficult to follow the knowledge norm of belief; it is difficult to believe a

---

[3] A similar point has been made in a different context by John Hawthorne and Amia Srinivasan (2013). They take a knowledge-first approach to argue that it is not possible to provide a satisfying answer to the question: '*What ought one to do, epistemically speaking, when faced with a disagreement?*'.

hypothesis only if one knows that hypothesis, when one cannot always tell whether one knows that hypothesis. It may seem overly demanding therefore to maintain that it is inappropriate to flout the norm by believing an unknown hypothesis (cf. Hawthorne and Srinivasan 2013).

However, an important third tenet of the knowledge-first framework is worth mentioning in this respect, namely, the claim that all non-trivial norms for belief will be difficult to follow in this respect (Williamson 2000: pp. 222–223). In other words, for any non-trivial norm of belief, it is possible to mistake whether one is conforming to that norm. In this sense, this difficulty is not a special mark against the knowledge norm of belief; all non-trivial norms for belief will be similarly difficult to follow. Again, this prediction is borne out in the debate about mechanistic reasoning in medicine. In particular, Solomon has pointed out that it would be difficult to follow a norm appealing to high-quality mechanistic reasoning, because '[i]t is difficult, however, to assess our knowledge of the completeness of relevant mechanisms' (2015: pp. 122–123). In other words, it is difficult to follow the norm of high-quality mechanistic reasoning, since our mechanistic evidence is non-transparent. Indeed, Solomon thinks that the problem in the hormone replacement therapy case was not only that we did not know the relevant complexity of the mechanisms, but that also '*we did not know that we did not know* the relevant complexity of mechanisms, and thus exaggerated the little evidence there was for expecting hormone replacement therapy to reduce cardiac mortality' (2015: p. 123).

Of course, this falls well short of a conclusive argument in favour of this third tenet of the knowledge-first framework. But I am not attempting here a full-blown defence of any of the tenets of the knowledge-first framework. This has been done elsewhere in large part by Timothy Williamson (2000). A proper appraisal will involve close engagement with the relevant arguments on its behalf as they are laid out in the epistemology literature. I am content merely to point out some important implications of the knowledge-first approach for a prominent recent activity in the philosophy of science: if the knowledge-first framework is correct, then at least one prominent activity within the philosophy of science is misguided. Given such implications, I think it is worth paying closer attention in the philosophy of science to recent arguments from epistemology to the effect that knowledge is the *explanans* rather than the *explanandum*. A recent advocate of this approach is Alexander Bird (2010). But the approach goes against a current trend in the philosophy of science, which attempts to practice independently of results in epistemology. For example, John Worrall says that '"*S* knows that *p*" epistemology … is entirely irrelevant for any sort of science' (Worrall 2007: p. 1019).

5 Conclusion

In this paper, I have argued that it has proved difficult to provide a satisfying answer to the principal question: *When is it appropriate to believe a causal hypothesis on the basis of mechanistic reasoning?* An appeal to high-quality mechanistic reasoning was one attempt to provide a satisfying answer to that question. However, I appealed to a historical case study involving hormone replacement therapy in order

to argue that this answer was prone to a counterexample. I also appealed to a more recent historical case study to argue that an answer appealing instead to super-high-quality mechanistic reasoning was also susceptible to a counterexample. Moreover, I argued that this difficulty in answering the principal question is predicted by recent work in *knowledge-first epistemology*. According to this work, it is not possible to provide a satisfying answer to the principal question. I think this shows that recent work in epistemology has important implications for practice within the philosophy of science. It is therefore worth paying closer attention in the philosophy of science to this work in knowledge-first epistemology.

Of course, a post-positivist philosophy of science takes it that any epistemological framework available to armchair reflection upon imagined cases is likely to prove a crude tool for understanding the diversity of methods employed in actual scientific practice (cf. Kuhn (1970)). Instead, epistemological morals should be drawn only from a historically-informed philosophy of science: an epistemology fit for science pays close attention to individual case studies of actual scientific practice. I agree that historically-informed case studies in the philosophy of science provide a more down-to-earth testing ground for competing epistemological frameworks. Indeed, I have appealed to a case study involving hormone replacement therapy in order to draw an epistemological moral about high-quality mechanistic reasoning. But an adverse side-effect of this development has been that philosophers of science pay less attention to work in recent epistemology. I hope to have shown that recent results in epistemology have important implications for the philosophy of science. In particular, if the knowledge-first approach is correct, then at least one prominent activity within the philosophy of science involves going down a blind alley, namely, the attempt to provide conditions for appropriately believing a causal hypothesis by mechanistic reasoning. In order to better determine whether such an activity is indeed heading down a blind alley, it is worth paying closer attention in the philosophy of science to work in knowledge-first epistemology. I hope that this helps to motivate an alternative approach in the philosophy of science, an approach which insists upon a mutually beneficial interaction with recent work in epistemology. To misquote Imre Lakatos: 'Epistemology without the philosophy of science is empty. The philosophy of science without epistemology is blind' (cf. (1981: p. 107)).

References

Aronson, J. (2020). Defining aspects of mechanisms: evidence-based mechanism (evidence *for* a mechanism), mechanism-based evidence (evidence *from* a mechanism), and mechanistic reasoning. In A. LaCaze & B. Osimani (Eds.), *Uncertainty in Pharmacology: Epistemology, Methods, and Decisions* (pp. 3–38). Cham: Springer.

Baum, M., Demicheli, R., Hrushesky, W., & Retsky, M. (2005). Does surgery unfavourably perturb the natural history of early breast cancer by accelerating the appearance of distant metastases? *European Journal of Cancer, 41,* 508–515.

Bernard, C. (1865). *An Introduction to the Study of Experimental Medicine*. Dover, 1957.

Bird, A. (2010). The epistemology of science—a bird's eye view. *Synthese, 175,* 5–16.

Bonnema, S. J., Nielsen, V. E., Boel-Jorgensen, H., Grupe, P., Andersen, P. B., Bastholt, L., & Hegedüs, L. (2007). Improvement of goiter volume reduction after 0.3 mg recombinant human thyrotropin-stimulated radioiodine therapy in patients with a very large goiter: A double-blinded, randomized trial. *The Journal of Clinical Endocrinology and Metabolism, 92*(9), 3424–3428.

Bonnema, S. J., Nielsen, V. E., Boel-Jorgensen, H., Grupe, P., Andersen, P. B., Bastholt, L., & Hegedüs, L. (2008). Recombinant human thyrotropin-stimulated radioiodine therapy of large nodular goiters facilitates tracheal decompression and improves inspiration. *The Journal of Clinical Endocrinology and Metabolism, 93*(10), 3981–3984.

Clarke, B., Gillies, D., Illari, P., Russo, F., & Williamson, J. (2014). Mechanisms and the evidence hierarchy. *Topoi, 33,* 339–360.

Dubey, R. K., Imthurn, B., Barton, M., & Jackson, E. K. (2005). Vascular consequences of menopause and hormone therapy: Importance of timing of treatment and type of estrogen. *Cardiovascular Research, 66*(2), 295–306.

Gettier, E. (1963). Is justified true belief knowledge? *Analysis, 23*(6), 121–123.

Glennan, S. (1996). Mechanisms and the nature of causation. *Erkenntnis, 44,* 49–71.

Gøtzsche, P., & Olsen, O. (2000). Is screening for breast cancer with mammography justifiable? *The Lancet, 355,* 129–134.

Greenough, P., & Pritchard, D. (2009). *Williamson on Knowledge*. Oxford: Oxford University Press.

Hawthorne, J., & Srinivasan, A. (2013). Disagreement without transparency: Some bleak thoughts. In D. Christensen & J. Lackey (Eds.), *The Epistemology of Disagreement* (pp. 9–30). Oxford: Oxford University Press.

Hitchcock, C. (2001). A tale of two effects. *The Philosophical Review, 110*(3), 361–396.

Howick, J. (2011a). Exposing the vanities—and a qualified defense—of mechanistic reasoning in health care decision making. *Philosophy of Science, 78,* 926–940.

Howick, J. (2011b). *The Philosophy of Evidence-Based Medicine*. Oxford: BMJ Books.

Howick, J. (2016). Aulus Cornelius Celsus and 'empirical' and 'dogmatic' medicine. *Journal of the Royal Society of Medicine, 109*(11), 426–430.

Illari, P. (2011). Mechanistic evidence: Disambiguating the Russo-Williamson thesis. *International Studies in the Philosophy of Science, 25,* 139–157.

Kuhn, T. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Lakatos, I. (1981). History of science and its rational reconstructions. In I. Hacking (Ed.), *Scientific Revolutions* (pp. 107–127). Oxford: Oxford University Press.

Littlejohn, C. (2013). The Russellian retreat. *Proceedings of the Aristotelian Society, 113*(3), 293–320.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 67*(1), 1–25.

McGlynn, A. (2014). *Knowledge First?* London: Palgrave Macmillan.

Mendelsohn, M., & Karas, R. (1999). The protective effects of estrogen on the cardiovascular system. *New England Journal of Medicine, 340*(23), 1801–1811.

Nielsen, V. E., Bonnema, S. J., Boel-Jorgensen, H., Grupe, P., & Hegedüs, L. (2006). Stimulation with 0.3-mg recombinant human thyrotropin prior to iodine 131 therapy to improve the size reduction of benign nontoxic nodular goiter: A prospective randomized double-blind trial. *Archives of Internal Medicine, 166*(14), 1476–1482.

Petitti, D., Perlman, J., & Sidney, S. (1987). Noncontraceptiveestrogens and mortality: Long-term follow-up of women in the Walnut Creek Study. *Obstetrics and Gynecology, 70,* 289–293.

Retsky, M., Demicheli, R., & Hrushesky, W. (2005). Does surgery induce angiogenesis in breast cancer? Indirect evidence from relapse pattern and mammography paradox. *International Journal of Surgery, 3,* 179–187.

Rossouw, J., Anderson, G., & Prentice, R. (2002). Risks and benefits of estrogen plus progestin in healthy postmenopausal women: Principal results from the women's health initiative randomized controlled trial. *Journal of the American Medical Association, 288,* 321–333.

Russo, F., & Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in the Philosophy of Science, 21,* 157–170.

Russo, F., & Williamson, J. (2011). Epistemic causality and evidence-based medicine. *History and Philosophy of the Life Sciences, 33,* 563–582.

Solomon, M. (2015). *Making medical knowledge*. Oxford: Oxford University Press.

Stampfer, M., & Colditz, G. (1991). Estrogen replacement therapy and coronary heart disease: A quantitative assessment of the epidemiologic evidence. *Preventive Medicine, 20,* 47–63.

Steel, D. (2008). *Across the boundaries: Extrapolation in biology and social science*. Oxford: Oxford University Press.

Sutton, J. (2007). *Without justification*. Cambridge: MIT Press.

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

Worrall, J. (2007). Evidence in medicine and evidence-based medicine. *Philosophy Compass, 2*(6), 981–1022.