Inference to the Only Explanation

**Introduction—Underdetermination, Evidence, and Inference to the Best Explanation**

Peter Lipton’s model of Inference to the Best Explanation (IBE) has the following features:

(a) IBE is a two stage process: (i) the first filter draws up a shortlist of potential explanations of a relevant phenomenon; (ii) the second filter chooses from the shortlist the favoured, best explanation.

(b) The choice in both stages is driven by *explanatory loveliness*. In stage (ii) the best explanation is the loveliest of the potential explanations. In stage (i) the choice of shortlist is driven by background beliefs, which themselves will have been selected thanks to IBE, i.e. on the basis of explanatory loveliness.

(c) The explanations in question are contrastive.

In the above loveliness is a qualitative character that is distinct from a quasi-quantitative assessment of likeliness or probability. Indeed, the point of IBE is that loveliness is a guide to likeliness. In chapter 5 Lipton presents a case study, Semmelweis’s investigation of puerperal (or childbed) fever, which permits a comparison with Hempel’s discussion of the same case in support of his hypothetico-deductive account of confirmation. One interesting feature of Lipton’s presentation is that loveliness is not discussed. This is because, as Lipton says, Semmelweis converted his problem into the question of the only explanation of the contrastive facts (p.90). In what follows I want to explore the prospects for regarding Lipton’s framework as permitting an account of abductive inference in which Inference to the Only Explanation plays a central role.

It might be thought that Lipton himself, as just reported, has given at least an example of Inference to the Only Explanation. But, as we shall see, that is not how Lipton understands things. By Inference to the Only Explanation (IOE) I intend something quite specific, that at the end of inquiry we can be in the position to infer the truth of some hypothesis $h$ since it is the only possible hypothesis left unfuted by the evidence. It is the form of inference advocated by Sherlock Holmes in his famous dictum “Eliminate the impossible, and whatever remains, however improbable, must be the truth.” Of course one requires the auxiliary hypothesis that there is an explanation of the phenomenon in question. But often we know that there is something rather than nothing that explains a phenomenon (it is not the case that nothing explains combustion or the demise of the dinosaurs). Furthermore, if we adopt Timothy Williamson’s equation of knowledge with evidence, then that knowledge is another piece of evidence, and so our total evidence entails the truth of the remaining hypothesis.

---

Lipton does *not* regard this as the correct description of Semmelweis’s inference. This is for three related reasons. First, even if Semmelweis had refuted every other hypothesis that he had considered, it does not follow that every possible alternative hypothesis had been refuted. Secondly, it appears that Lipton may have a restricted conception of evidence combined with a restricted view of the epistemological efficacy of evidence—in which case refutation is rarer than some might think. Thirdly, Lipton does not think that the rejected hypotheses in Semmelweis’s case are refuted.

**Evidence and refutation**

I’ll first briefly consider Lipton’s views on evidence. I should point out immediately that Lipton does not present a theory of evidence and it would be unfair to impute to him a theory on the basis of one or two remarks. Nonetheless, I shall argue that if Lipton were to have a particular theory of evidence, one which contrasts with that suggested by his remarks, then he would be in a position to present a rather more robust account of Inference to the Best Explanation, one which is able to incorporate the Holmesian variety of Inference to the Only Explanation sketched above.

Here is how Lipton characterizes the important notion of a *potential explanation*. It is important because the set of potential explanations is the set from which we must pick that explanation we will regard as the loveliest and hence the best. Lipton says, “We might say that a potential explanation is any account that is logically compatible with all our observations (or almost all of them) and that is a possible explanation of the relevant phenomena” (p.59) Note two things:

(i) It is compatibility with *observations* that determines the set of potential explanations.

(ii) Being incompatible with an observation is nonetheless not *sufficient* for ruling out a possible explanation. Lipton regards observation as non-factive.

We can readily see why on Lipton’s view it is difficult to refute a hypothesis. First, the restriction of ‘refuters’ to observations means that a hypothesis will escape refutation even though it is incompatible with a wider range of known propositions. If a proposed new theory of everything has the consequence that electrons are positively charged or that the universe is only a million years old, we would expect to be able to reject it immediately. However, Lipton’s view implies that this is too hasty. We need to find some observations with which that hypothesis is inconsistent—and it is not at all clear that there are any. Secondly, even if we do find an observation with which some hypothesis is incompatible, that is not enough to refute it; the hypothesis survives as long as it is consistent with most observations.

Of course, Lipton is not suggesting that we do consider all the hypotheses consistent with our observations—there are too many and many of them are absurd. The first filter is a shortlist
mechanism “where our background beliefs help us generate a very limited list of plausible hypotheses, from which we then choose.” (p.149) As a description that seems right. But it is not entirely clear what Lipton thinks is the epistemic status of the hypotheses that don’t get through the first filter and so are never considered. It seems to me that he thinks that many may in fact be consistent with our evidence but are implausible in the light of background beliefs. In my view many of the ‘unconsidered’ hypotheses are in fact incompatible with and hence refuted by our evidence.

Here is a contrasting view of evidence. A hypothesis is refuted if it is incompatible with any known proposition.3 We can assume that what is known includes propositions that are not observations, for to deny this would be to hold a sceptical view of what science can achieve according to which both IBE and IOE are hopelessly doomed from the start. Let us imagine then that S has correctly deduced that hypothesis h is inconsistent with a proposition p, and S knows p. Closure implies that S knows that h is false.4 In general the known incompatibility of a hypothesis with a proposition (observational or not) that is known to be true gives one knowledge that the hypothesis is false and hence refuted. So the restriction to observations is unwarranted. Furthermore, known incompatibility with a known proposition is sufficient for refutation, which contrasts with (known) incompatibility with Lipton’s observations, which we saw was insufficient ((ii) above).

What about hypotheses that have not been considered? There are two approaches one may take here. First we may say that this class includes those hypotheses such that if we did consider them, we would see that they are incompatible with our evidence. Take for example a sceptical scenario of the kind with which we philosophers are familiar. If we have lots of the ordinary knowledge we take ourselves to have, then the sceptical scenarios are inconsistent with what we know. These are the cases where closure is controversial. The alternative approach argues that since the possible worlds in which such scenarios obtain are distant, it is it reliable, in a sense sufficient for knowledge, to ignore such possibilities. And the same may be said for outré scientific possibilities even if they are not the familiar sceptical scenarios of philosophy. Such unconsidered possibilities either are in fact incompatible with much that we know, or hold only in sufficiently distant possible worlds that it is reliable to ignore them. In effect Lipton’s first filter filters out hypotheses such that either (i) were one to consider them, one would know them not to be the explanation, or (ii) ignoring them in reasoning is epistemically reliable.

3 This proposal is motivated by the connection with closure, as I discuss. But it has obvious connections with Williamson’s equation of evidence with knowledge.

4 Unrestricted closure is controversial. For those who do not accept unrestricted closure we can bracket off the difficult cases that concern scepticism, and accept a restricted version that will apply in ordinary scientific cases. That p is not an observation is not a reason for denying the applicability of closure.
If we allow ourselves all that we know as ammunition in attempting to refute a hypothesis, we allow ourselves many potential refuters excluded by Lipton’s restriction to observation, each of which is sufficient to refute a hypothesis with which it is incompatible. Thus, speaking loosely, refutation of a hypothesis (both considered and unconsidered) ought to be easier according to this proposal than according to Lipton.

Elimination of hypotheses

That Lipton’s view of evidence limits the power of inference is reflected in the following quotation:

“Mill’s methods have their share of liabilities … if the methods are to apply at all, the requirement that there be only a single agreement or difference must be seen as an idealization, since this condition is never met in real life.” (p.19)

Imagine that I am a chemist and have prepared two identical litmus solutions. Into one I drop a small quantity of a test substance. Immediately the solution turns red. Mill’s method of difference allows us to infer that because the introduction of the test substance was the only difference between the two solutions, it is the cause of the change in its colour. Lipton’s complaint is that strictly it is not the only difference. The two solutions are in different positions on the bench and so also stand in different relations to all sorts of things. Furthermore, can we be sure that the two solutions are chemically identical? The natural response to such complaints is that we know that position on the bench makes no difference in experiments like this, and that if the litmus solutions were carefully prepared, e.g. by dividing a well stirred single solution, I know that there is no relevant difference in the constitution of the solutions.

Such knowledge claims can, of course, be wrong, but short of general scepticism, there is no good reason to suppose they can never be correct. In circumstances where they are correct, I am in a position to know that the only relevant difference between the solutions is the introduction of the test substance, and hence know that it is the cause of the changing colour. It is true that the knowledge in question is not observational. That rules it out as a refuter on Lipton’s conception. Hence, on that view, my evidence is not sufficient for me to infer ‘directly’ that the substance is the cause of the colour change—I’ll need to consider the loveliness of alternative explanations.

The conclusion of the preceding section was that, in the application of an inference procedure in science, non-observational evidence is crucial in the refutation of possible explanations. In the case just considered, that knowledge refuted hypotheses that were never ‘live’ for a real chemist, and hence the refutation operates at the stage of Lipton’s first filter. In the more complex Semmelweis case, non-observational knowledge is used in the refutation of live
possibilities, hence at Lipton’s second stage. As is well known, Ignaz Semmelweis considered and rejected a number of hypotheses concerning aetiology of puerperal fever, and in particular the stark difference in mortality between the two wards (or ‘divisions’) of Vienna’s maternity hospital.

Lipton’s discussion of Semmelweis is intended to achieve two things:

1. to show the importance of contrastive explanation for inference, and
2. to show the virtues of IBE over hypothetico-deductivism.

Note that these are distinct aims, for one could support IBE as applied to non-contrastive explanations. And one could apply hypothetico-deductivism to contrastive cases. Thus when Lipton shows the shortcomings of Hempel’s hypothetico-deductivist account of non-contrastive inference, we need to be clear whether the blame lies with the hypothetico-deductivism or with its application to non-contrastive cases.

Hempel’s use here of hypothetico-deductivism is essentially Popperian. Semmelweis infers the falsity of a hypothesis, because he deduces from it some proposition in conflict with the evidence. He then moves next hypothesis. I argue that Hempel is entirely correct, at least as concerns contrastive hypotheses. Lipton, however, complains that in fact the hypotheses are not refuted by the evidence. Of certain hypotheses concerning puerperal fever (miasma and epidemic causes), overcrowding, general care, diet etc.), Lipton says, “These hypotheses are rejected on the grounds that, though they are consistent with the evidence, they would not explain the contrast between the divisions. Epidemic influences, for example, still might possibly be part of the causal history of the deaths in the First Division, say because the presence of these influences is a necessary condition for any case of childbed fever.” (p.75)

Lipton says this because he is focussing on non-contrastive hypotheses. Let T1 be “Miasma is a causal factor in puerperal fever” and let T2 be “Miasma is a cause of the difference in mortality between the two divisions”. It may be that the non-contrastive T1 is not refuted by the evidence. But the case for saying that the contrastive T2 is thus refuted is strong. The miasma theory held that disease originates with decaying organic matter spread through the befouled air. It is a consequence of the theory that the miasmic influence on geographically similar and proximate locations should not differ greatly. And in particular, the first and second divisions were not different in this or related respects (e.g. ventilation etc.). But there was a significant difference in disease. Thus miasmic factors cannot explain that degree of difference. Thus, while this example does show the importance of contrastive inference, it does not show that hypothetico-deductivism, or, more importantly, the refutation of hypotheses, cannot account for contrastive inference. If we are to assess them in that regard, clearly we must assess them with respect to contrastive hypotheses. And they succeed well enough in that respect.

---

It is one thing to say that a deductivist approach can refute hypotheses, another to say that it does as well as Lipton’s IBE in confirming the correct hypothesis, that the cause of the difference in mortality is an infection from the ‘cadaveric’ matter on the hands of the medical students visiting the mothers in the First Division. (They did not visit the Second Division, where mothers were examined by midwives, who did not carry out dissections.) The salient fact is that when Semmelweis ensured that students disinfected their hands with chlorine before examining the mothers in the First Division (from May 1847), mortality rates were reduced to the levels of the Second Division. More precisely, the percentage mortality rates for the six years 1841-1846 were: I—9.92, II—3.88, and for the twelve years 1847-1858 were: I—3.57, II—3.05. This looks as if it ought to permit a straightforward application of Mill’s method of difference, just as in the litmus case discussed above. And in essence I think it does. Given the numbers of patients involved (over 125,000), this change in mortality is dramatic to say the least. There had been other differences between the two divisions, such as birth position and the priest who passed through the First Division to administer the sacrament of last rites. But elimination of these differences had not been followed by any statistically significant change in the mortality rates. Thus, like the miasma hypothesis, they cannot explain the large difference in mortality rates. Of course, all of these could have had some slight influence on mortality rates. Thus strictly, the relevant hypotheses must be of the form ‘X is a primary cause of the difference in mortality’ where a ‘primary cause’ is one that accounts for the bulk of the phenomenon observed. Clearly none of the factors could be a primary cause of the phenomenon (the large difference in mortality), if that phenomenon persists without the factor.

To reach Semmelweis’s conclusion we need also to consider the following possibilities:

(i) there is no single cause of the difference—because there is a multiplicity of independent, parallel causes operating cumulatively;
(ii) there is no single cause of the difference—because it has no causes at all;
(iii) some other hypothesis explains the difference, that we have not yet considered.

We can exclude (i) for two reasons. First, the possible contributory factors had mostly been removed (difference in birth positions, presence of the priest, etc.). So the cause could not be a compound of plausible factors. This leaves factors not considered (which we will look at under (iii)). Secondly, if this compound of factors does explain the difference, then it would have to have disappeared at the same time as Semmelweis introduced the hand disinfection policy. Apart from this being a coincidence, there was no change in a large number of conditions in the First Division.

The proposal under (ii) is in effect that the difference is a blunt fact that has no explanation. As regards singular facts, we may have to accept this possibility at the quantum level, but at

---

the macro level we do not. The detective does not have to consider the null hypothesis that the bullet just materialized in the victim’s brain, that there was no cause of its getting there. However, in the statistical macro case, we must consider the possibility that a statistic is a mere blip. Nonetheless in this case it is clear that Semmelweis and others knew that there was a genuine difference and a sizeable difference at that.\(^7\) Hence what Semmelweis knew is inconsistent with (ii).

Did what he knew exclude all other unconsidered hypotheses? Plausibly it did. One may have a sufficient understanding of a field that anything one has not thought of would have to be pretty outré. And as I discussed above the outré possibilities may all be incompatible with what one knows. Of course, one might not be in such a fortunate position, but, equally, a good scientist might be—and I think that Semmelweis was. Semmelweis had eliminated all possibly relevant differences between the wards. Other differences he knew of he also knew not to be possible explanations of the difference in mortality. So any unconsidered hypothesis, must concern a difference that Semmelweis was not able to discern. Furthermore, as in the discussion of (i), that difference would have to have disappeared coincidentally with Semmelweis ordering the student to disinfect their hands. Such an unconsidered hypothesis would, I suggest, be of the outré or sceptical kind that is incompatible with background knowledge and which may be reliably ignored.

I have argued that a deductivist refutation of hypotheses can eliminate all but one hypothesis, if we restrict ourselves, as Lipton enjoins, to contrastive explanation and so to contrastive hypotheses. It is worth asking whether any conclusions may be drawn about non-contrastive hypotheses. Clearly some conclusions can be drawn. Since Semmelweis had shown that the failure to disinfect hands that had been in contact with decaying cadaveric matter was a primary cause of the difference between the two divisions, and since that difference accounted for two-thirds of the deaths, it is clear that Semmelweis has identified a primary cause tout court of puerperal fever in that hospital, and by extension many other hospitals employing similar procedures. Had he eliminated other potential primary causes? Lipton says not, because the relevant hypotheses are consistent with what Semmelweis discovered. But consider the miasma hypothesis. It is part of the miasma hypothesis that for susceptible individuals the miasma is not merely necessary but also sufficient for illness, in the way that a poison is. Semmelweis had gone a long way to refuting the non-contrastive miasma hypothesis; but what prevented a conclusive refutation is the fact that Semmelweis could not eliminate puerperal fever altogether.\(^8\) The miasma theorist could thus claim that the remaining deaths (about a third of the original number, and 3% of the patients overall) are due to miasma. If

---

\(^7\) Popper would have regarded this as a conventional decision. C.f. K. Popper 1959 *The Logic of Scientific Discovery* London: Routledge and Kegan Paul; pp.104-5, 198ff. That is because he could not acknowledge that scientists knew in such situations without admitting that this is an instance of inductive knowledge.

\(^8\) In so saying we must remember that the fever was identified by its symptoms.
those deaths had been eliminated too, then miasma could not be a factor in deaths in the maternity wards.

_Inference to the Only Explanation_

Although Lipton is arguing for Inference to the Best Explanation, both his general approach and his example support the possibility, in some cases, of Inference to the Only Explanation, where the latter is such that one’s evidence includes:

(i) the fact the phenomenon in question has an explanation;
(ii) facts sufficient to refute all explanations bar one.

Note the following:

(a) The refutation of other hypotheses in (ii) might be either through Lipton’s first filter and are not consciously considered; or through his second filter, where experimental and other evidence may be used to refute live possibilities.
(b) The phenomenon in question may be a contrastive one, as in the Semmelweis case.
(c) The evidence is question is not exclusively observational. Non-observational knowledge may be required to know that the phenomenon has an explanation, or to refute a hypothesis.

We may think of considerations of loveliness coming into IBE when evidence is insufficient for us to carry out an IOE. Thus the proposal above is not intended to supplant Lipton’s account, but rather to point out an important limiting possibility it encompasses.

Although Lipton mentions the phrase ‘inference to the only explanation’ in relation to Semmelweis’s inference, he does not acknowledge the latter as a case of IOE as characterized above. Why not? There are two reasons. The lesser reason is that in discussing Hempel’s hypothetico-deductivism he concludes that Hempel cannot show that the hypotheses he considers are refuted. But Hempel is considering hypotheses concerning non-contrastive phenomena (hypotheses concerning the cause of puerperal fever _tout court_). And so Lipton’s argument does not rule out (‘hypothetico-deductive’) refutation by the same evidence of hypotheses concerning the _contrastive_ phenomenon (the difference in puerperal fever between the two divisions).

The second, perhaps weightier, reason is this. Lipton is impressed by considerations of a general problem of underdetermination. If IOE can occur, then we have a case where a theoretical inference is not underdetermined (even deductively) by the evidence. And I surmise that Lipton is impressed by underdetermination because he has a notion of evidence that is restricted to the observed. But once we remember that in general we can make knowledge-generating inferences from non-observational knowledge, a restriction of evidence to the observed is implausible.