

# IS THERE A RELIABILITY CHALLENGE FOR LOGIC?\*

Joshua Schechter

Brown University

## Introduction

There are many domains about which we think we are reliable. That is, we think that our beliefs about the domains are by and large true (or at least are true much more often than chance alone would predict).<sup>1</sup> For some of these domains, our reliability is not very puzzling. For instance, we understand how it is that we are reliable about facts about medium-sized objects in our environment. We possess psycho-physical explanations (or explanation sketches) of how our perceptual faculties work that explain how these faculties yield true beliefs about our environment. We also possess evolutionary explanations (or explanation sketches) of how we came to possess reliable perceptual faculties. For other domains, our reliability is more puzzling. This is especially true of domains that are typically thought of as a priori—domains where our beliefs do not depend on perception or any other kind of causal interaction with the world. Domains where our reliability is puzzling include logic, pure mathematics, basic facts about necessity and possibility, and basic facts about morality.

The *reliability question* about a domain is the question of how it is that our beliefs about the domain are reliable. When there is prima facie reason to think that there may not be a satisfying answer to this question given our background views, we arrive at a challenge to our reliability about the domain or to our background views. This is what is often called the *reliability challenge* about the domain.

More carefully, the abstract form of the reliability challenge for a domain is as follows:<sup>2</sup> We think that the beliefs we have about the domain are by and large true, or at least are true much more often than chance alone would predict.<sup>3</sup> This is a striking fact, one that “cries out” for explanation. To use Roger White’s term, the fact is “explanatorily urgent”.<sup>4</sup> For many domains, there is prima facie reason to think that there is no satisfying explanation that can be provided.

---

\* Forthcoming in *Philosophical Issues* (2018), *Philosophy of Logic and Inference*.

<sup>1</sup> To this, it should be added that our disbeliefs are by and large false (or are false much more often than chance alone would predict). And similarly for other cognitive attitudes, such as credences and beliefs under suppositions. It could also be added that our beliefs (and disbeliefs) about the domain are wide-ranging, reasonably powerful, and so on. Notice that the claim that we are reliable is a claim about our actual attitudes, not about how our attitudes would match the facts in counterfactual scenarios where the truths are different.

<sup>2</sup> See Schechter (2010) and (2018) for more detail on how to understand reliability challenges. Reliability challenges are analogous to the Benacerraf-Field argument against mathematical Platonism. See the Introduction and title essay of Field (1989). Also see Benacerraf (1973) for an important precursor. See Enoch (2011), chapter 7, for a reliability challenge aimed at moral realism, and for a response to that challenge.

<sup>3</sup> Additional hedges may be needed here. For instance, perhaps our beliefs are reliable only when given sufficient reflection and discussion. Any necessary hedges should be understood to be implicit.

<sup>4</sup> See White (2005).

This tends to be true when (i) we think that the domain is a domain of objective, mind-independent fact; and (ii) we think that the facts of the domain do not play a causal role in explaining our beliefs. This is because, given (i) and (ii), our beliefs do not explain the facts and the facts do not explain our beliefs. (Typically, there is also reason to think there is no third factor that somehow explains both our beliefs and the facts.) According to a general epistemic principle, it is a cost of a theory if it treats a striking phenomenon within the scope of the theory as accidental or otherwise inexplicable. This yields a tension within our overall package of views. There is rational pressure to resolve this tension by giving up the claim that we are reliable about the domain, giving up the claim that the domain is objective, giving up the claim that facts of the domain do not play a causal role in explaining our beliefs, or giving up some of our background views of the world that help to make it difficult to provide a satisfying explanation of our reliability. For some domains, giving up one of these claims will be much more plausible than giving up the rest. For other domains, there may be responses to the challenge that alleviate the tension without requiring us to give up any claim. (For instance, perhaps we can find an explanation of our reliability that is compatible with the objectivity and acausal nature of the domain.) For still other domains, the right thing to do in response to a reliability challenge is to live with the tension.

Reliability challenges do not only arise for beliefs, but also for kinds of reasoning. There are several kinds of reasoning that we think are reliable—deduction, induction, explanatory reasoning, analogical reasoning, and so forth. That is, we think that reasoning in those ways tends to yield true beliefs from true beliefs.<sup>5</sup> The reliability challenge for a kind of reasoning is closely analogous to a reliability challenge to beliefs about a domain. Given a kind of reasoning that we employ and that we take to be reliable, there is the question of how it is that we employ such a reliable kind of reasoning. In many cases, there is *prima facie* reason to think that there is no satisfying explanation that can be provided. That we reason in a reliable way is a striking fact, one that cries out for explanation. Again, according to a general epistemic principle, it is a cost of a theory if it treats a striking phenomenon within the scope of the theory as accidental or otherwise inexplicable. This yields a tension within our overall package of views. It generates rational pressure to resolve this tension.

In previous work,<sup>6</sup> I discussed the reliability challenges for logic and for deductive inference.<sup>7</sup> I argued for four main claims: First, there are reliability challenges for logic and for deduction. At least *prima facie*, there is reason to think that there are no satisfying explanations of how it is that we are reliable about logic, or how it is that we employ reliable deductive rules of inference. This

---

<sup>5</sup> This should be extended to cognitive attitudes other than beliefs (e.g., disbeliefs, credences, and beliefs under suppositions). It should also be added that the forms of reasoning apply broadly, are reasonably powerful, etc.

<sup>6</sup> See Schechter (2010) and (2013).

<sup>7</sup> For present purposes, we can understand logic to consist of a collection of truths—the logical truths. By contrast, deductive inference is a kind of reasoning—including, for instance, the inference from a conjunction to one of its conjuncts. The idea that logic is a body of truths rather than a kind of reasoning is emphasized in Harman (1986), chapter 2, though he goes on to claim that there is no such thing as deductive inference.

puts pressure on the claim that we are reliable, the claim that logic is a domain of objective fact, or on various background views we have about the world. Second, these reliability challenges cannot be answered merely by providing an explanation of how it is that we have the logical beliefs and employ the deductive rules that we do. Providing explanations of how we ended up with our beliefs and rules does not explain how it is that we have reliable beliefs and rules. Third, we can explain our reliability about logic by appealing to our reliability about deduction. Our logical beliefs (or a core subset of our logical beliefs) are the outputs of deductive reasoning. These beliefs are reliable because the deductive rules that we employed in arriving at the beliefs are reliable. Fourth, and finally, there is a good prospect for providing an evolutionary explanation of the reliability of our deductive reasoning.

In my earlier work, I presented a toy explanation that appeals to evolution by natural selection. In broad outline, the idea is that the trait of being a reliable deductive reasoner conferred a heritable survival or reproductive advantage upon our ancestors. We inherited this trait, and that is what explains our reliability. The explanation I presented was not meant to be the correct explanation of our deductive reliability. Rather, it was meant to be something like a proof of concept. The explanation shows that there is no in principle reason to think that evolutionary explanations—whether appealing to evolution by natural selection or to some kind of social or cultural evolution—are incapable of explaining our reliability about deduction. Indeed, I suspect that the correct explanation of our deductive reliability depends both on natural selection and social phenomena, where natural selection accounts for the reliability of some core parts of our deductive capacity and the social phenomena account for the rest. But that is an empirical matter which cannot be settled using purely philosophical methods.

In the last fifteen or twenty years, a number of arguments have appeared in the literature that can be applied against one or more of these four theses. Some of these arguments predate my earlier work on the subject and some are more recent. Some of the arguments involve considerations that are specific to the case of logic and some are more general. What I would like to do in this paper is to respond to some of these arguments. In what follows, I will discuss arguments by Paul Horwich, Jack Woods, Dan Baras, Justin Clarke-Doane, and Hartry Field. These arguments do not exhaust the arguments which have appeared in the literature on this topic. They do, however, provide a good sample of interesting and important objections. So in what follows, I'll focus on them.

### **Overgeneralizing from the Physical**

The first objection I'd like to consider is due to Paul Horwich. Horwich argues that the apparent metaphysical and epistemological problems facing domains such as mathematics, modality, and morality are really pseudo-problems. He writes:

[T]he basic thought is that we tend to over-generalize. Most of the facts that we encounter—most of the facts that we care about—involve material objects and their empirically detectable properties. As a consequence, these facts become our paradigms. So much so, that the very idea of a fact comes to seem conceptually bound up with the characteristics of physical facts. Thus the acknowledgment of genuine ethical facts, arithmetical facts, modal facts, and so on, becomes paradoxical: on the one hand they are unlike physical facts; but, on the other hand—in order to qualify as facts at all—they must resemble physical facts. Hence the aura of weirdness.

And for similar reasons their epistemological status can seem problematic. For knowledge of the physical requires certain causal or counterfactual relations between the facts and our beliefs about them—relations that cannot hold in a variety of other domains. Thus generalization from the physical case—over-generalization again—will lead to skepticism with respect to those areas.<sup>8</sup>

Horwich does not explicitly discuss the case of logic. He also doesn't discuss reliability challenges. But his general line of thought presumably applies to the reliable challenge for logic as well.<sup>9</sup> Applying his line of thought to this case, the idea then would be this: Our reliability about an empirical domain is a striking fact. Compelling reason to think that no satisfying explanation of our reliability is possible would yield a tension in our overall package of views. But the same is not true of our reliability about logic. Logic is not an empirical domain, and the epistemological principles that apply to our beliefs about the physical world do not apply to our beliefs about non-empirical matters. In particular, the fact that we are reliable about logic is not a striking fact—it is not explanatorily urgent.<sup>10</sup> (And similarly for our reliability about mathematics, modality, and morality.)

It is certainly true that there are important differences between facts about the physical world and facts about logic (or mathematics, modality, and morality). And there are important differences between our empirical beliefs and our beliefs about logic. But Horwich's response does not show that the reliability challenge for logic is without force. An initial problem with Horwich's line of thought is that it makes the defense of patently silly views much too easy. Consider someone who claims that he has copious knowledge of facts about empirically undetectable non-material gremlins that have complex social practices. (Suppose he claims that his beliefs about the gremlins are justified because they are based on rational intuitions about these creatures.) It would be natural to object that there is no way that he could have such knowledge, since there is no explanation of how he could have come to have true beliefs about empirically undetectable non-material gremlins. If Horwich's line of thought were correct, however, this objection would

---

<sup>8</sup> Horwich (2006), p. 199.

<sup>9</sup> In personal communication, Horwich endorsed extending his line of thought to reliability challenges, too.

<sup>10</sup> Alternatively, one might claim that our reliability is a striking fact but that having reason to think there is no satisfying explanation of this fact does not yield a tension in our package of views. This is a less plausible avenue to pursue, since the strikingness of a fact seems to consist in the fact crying out for explanation.

not have any force. The silly view could be defended against the objection merely by claiming that because the gremlins are empirically undetectable and non-material, there is no need to explain how one could have ended up with true beliefs about them. This defense seems absurd. And that suggests something is wrong with Horwich's approach.

There is a second, more illuminating problem with Horwich's line of thought. Suppose we grant that much of our understanding of metaphysics and epistemology comes from our observations of the physical world. Suppose we also grant that the metaphysical and epistemological principles we endorse on this basis only apply to empirical facts and empirical beliefs.<sup>11</sup> A notable feature of the reliability challenge is that the general principle at its base—namely, that it is a cost of a view if it treats striking phenomena as accidental or otherwise inexplicable—need only be applied to empirical phenomena. The empirical phenomena in question are our beliefs (and other cognitive attitudes). What is striking is that our beliefs about the relevant domain are by and large true. This is a highly non-random pattern in a collection of empirical phenomena. So the general principle is not being applied to anything non-empirical. Thus, even if Horwich is right that in some cases we overgeneralize from the physical to the non-physical, the reliability challenge does not rely on such an overgeneralization.

An analogy might be useful here. Suppose one flips an apparently fair coin a thousand times. Suppose the coin comes up heads when (and only when) the number of the flip in the sequence is a prime number. (Alternatively, suppose the pattern of heads and tails spells out a proof of some mathematical theorem in Morse code, or ....) This would be a very striking empirical fact. Because the pattern of flips was so striking, there would be rational pressure to come to believe that the coin was somehow being manipulated, or at least, that something funny was going on. The reliability challenge depends on the same basic thought.

## **Self-Effacement**

The second objection I'd like to consider is due to Jack Woods.<sup>12</sup> Woods argues that in raising a particular kind of epistemological challenge to any domain, we must rely on elementary parts of logic and mathematics. Thus, argues Woods, those parts of logic and mathematics are immune to that kind of challenge, unlike (say) morality and other a priori domains.

Woods's discussion does not focus on the reliability challenge, but on a challenge to moral realism due to Gilbert Harman.<sup>13</sup> Harman's challenge is (roughly) that the best explanation of why we have the moral beliefs that we do does not involve any moral facts. Or, more precisely,

---

<sup>11</sup> I'm not convinced that we should grant this. As Berry (ms) points out, mathematicians treat unexplained correlations as crying out for explanation, too. This has led to many successful research projects in mathematics.

<sup>12</sup> See Woods (2018).

<sup>13</sup> See Harman (1977), chapter 1.

the best explanation of our moral beliefs does not involve moral facts as the moral realist conceives of them. This provides reason to give up either our moral beliefs or moral realism.<sup>14</sup>

Woods considers the analogous argument for the cases of logic and mathematics. He notes that claims about what is the best explanation of a fact—and, in particular, what is the best explanation of our beliefs about a domain—rely on elementary parts of logic and mathematics. This is because “the notion of *best* in *best explanation* is not independent of our mathematical and logical beliefs—our conclusions about which explanation is best will depend on what mathematical and logical theory we accept in the background.”<sup>15</sup> For instance, in assessing which explanation is best, considerations about ontological simplicity play a role. To assess the simplicity of a candidate explanation requires determining whether the candidate explanation logically entails the existence of any kinds of entities, and if so, how many kinds of entities are so entailed. This involves facts about logical entailment and facts about numerical comparisons. More generally, the assessment of the goodness of a candidate explanation plausibly depends on facts about logical consistency and logical entailment, cardinality comparisons, elementary probability theory, and elementary statistics. So a Harman-like challenge for those parts of logic and mathematics would rely on those parts of logic and mathematics. Thus, argues Woods, “Harman-style debunking arguments against mathematics and logic are self-effacing; coming to believe that we should give up our mathematical and logical beliefs on their basis undermines the premises on which this conclusion is based.”<sup>16</sup> So a Harman-like challenge cannot be raised for those parts of mathematics and logic.

Woods’s line of thought would seem to extend to reliability challenges as well. The claim that we are reliable about a domain involves the claim that our beliefs are correct more often than chance alone would predict. This relies on elementary facts about probability and statistics. Similarly, the claim that there is no satisfying explanation of our reliability depends on the evaluation of candidate explanations, which again plausibly depends on elementary parts of logic and mathematics. So, by an analogous line of thought, if a reliability challenge for those parts of logic and mathematics were successful, it would undermine its own premises. Therefore, one might argue, there cannot be a successful reliability challenge for elementary parts of logic and mathematics.

This objection is arresting, but I think that it ultimately fails to work. There are two main issues with Woods’s line of thought. First, there can be good arguments that undermine their own premises. Indeed, that is how *reductio ad absurdum* arguments work. We temporarily grant that some claims are true, end up with the conclusion that they are not, and so have good reason to reject the claims. The fact that the premises of the argument are undermined by its very success is no count against the force of the argument. So even if the reliability challenge (or Harman’s

---

<sup>14</sup> See Schechter (2018) for an interpretation of Harman’s challenge.

<sup>15</sup> Woods (2018), p. 59, emphasis his.

<sup>16</sup> *Ibid.*, p. 61.

challenge) for elementary parts of logic and mathematics ends up undermining its own premises, that is not a problem with the argument.<sup>17</sup>

Second, and more importantly, the reliability challenge for elementary parts of logic and mathematics does not undermine its own premises. We should understand the argument not as aimed against the truth of our logical and mathematical beliefs, but as aimed against views about their status. A reliability challenge shows that there is a tension in one's overall package of views. One way to alleviate the tension is to reject the claim that one is reliable about logic and mathematics. But that's not the only way to alleviate the tension. Indeed, it is not the most plausible way. It would be a radical shift in one's views if one were to give up elementary parts of logic and mathematics. It is far more plausible to give up the view that logic and mathematics are objective or the view that logic and mathematics are acausal, or to give up some other part of our overall view of the world. So it is best to understand the reliability challenge for elementary parts of logic and mathematics as having the ultimate conclusion that either (i) logic and mathematics are not objective, or (ii) we have causal interactions with logical and mathematical objects, or (iii) some other part of our picture of the world needs modification. If we understand the reliability challenge in this way, if the challenge is successful, it does not undermine its own premises, and so is not self-effacing in the way that Woods suggests. And the same would seem to be true of a Harman-like challenge for elementary parts of logic and mathematics, as well.

### **High Probability Explanations; Safety and Sensitivity**

The third kind of objection I'll consider applies to reliability challenges for any domain in which the facts are necessary truths. The idea is that we can explain our reliability about such a domain simply by explaining how it is that we came to have the beliefs about that domain that we have. We can provide an evolutionary, psychological, or sociological explanation of how we ended up with our beliefs, and this answers the challenge.

There are several different ways to argue for this conclusion, each involving a claim about what it is to explain our reliability. Dan Baras argues that it suffices to explain our reliability to provide what he calls a "high probability explanation" of our reliability. Such an explanation "consists of information about initial conditions that imply that the explanandum (i.e., the striking correlation) was highly probable."<sup>18</sup> Given this conception of what is required to explain our reliability, Baras argues that it suffices to provide an account of how we ended up with the beliefs that we have. That is because the claim that our beliefs are reliable logically follows from the conjunction of the claim that we have the beliefs that we do and the claim that the contents of

---

<sup>17</sup> One might argue that the claims that are undermined by reductio arguments are not premises, but suppositions. I think that's right, but it does not affect the point here. We can re-interpret an argument that undermines its premises as an argument that shows that some suppositions are false.

<sup>18</sup> Baras (2017), p. 201.

many of these beliefs are true.<sup>19</sup> Since the domain in question is a domain of necessary truths, the probability of the second conjunct is 1.<sup>20</sup> So the probability of the conjunction is simply the probability of the first conjunct. If information can be provided that shows that the first conjunct is highly probable, this information will show that the conjunction is highly probable, and thus will show that our reliability about the domain is highly probable. A causal explanation of how we ended up with our beliefs about the domain will therefore suffice to explain our reliability.

Justin Clarke-Doane provides a different argument for the claim that an explanation of how we came to have our beliefs about the domain suffices to explain our reliability.<sup>21</sup> Clarke-Doane argues that there are only two legitimate explanatory demands concerning our reliability about a domain. The first is the demand to show that our beliefs about the domain are *sensitive* in the sense that had the contents of our beliefs about the domain been false, we would not still have believed them (using the very same methods we actually used to form our beliefs).<sup>22</sup> The second is the demand to show that our beliefs about the domain are *safe* in the sense that we could not have easily had false beliefs about the domain (using the very same methods we actually used to form our beliefs).<sup>23</sup> Since the domain is a domain of necessary truths, our beliefs are automatically sensitive—it is impossible for the contents of our beliefs to be false. If we have an evolutionary, psychological, or sociological explanation of how we came to have our beliefs about the domain, it is plausible that this explanation will show that we could not easily have had different beliefs. So we could not easily have had false beliefs. This shows that our beliefs are safe, as well.

In reply to Baras, it seems to me that explanation should not be understood in merely probabilistic terms. For example, a thinker may have observed very many snowflakes and seen that they all exhibit six-fold symmetry.<sup>24</sup> This information implies that the probability that all snowflakes have six-fold symmetry is very high. But it does not explain why snowflakes have six-fold symmetry. To be fair, the observational information about the snowflakes is not information about initial conditions, at least when that is construed narrowly. But I don't see that this matters. Baras does not tell us why a high probability explanation should involve information about initial conditions. And, indeed, the explanation of the six-fold symmetry of snowflakes does not clearly involve initial conditions. (The explanation has to do with the fact

---

<sup>19</sup> The second conjunct is meant to be understood *de re* and not *de dicto*.

<sup>20</sup> This may not be true for all necessary domains. But it is certainly true for the specific case of logic.

<sup>21</sup> See Clarke-Doane (2017). Clarke-Doane's view is a sophisticated version of an idea that is sometimes attributed to Lewis (1986), pp. 114–5. The idea is that if one's belief has a content that is necessarily true, it is not possible for one to be wrong, and so one's reliability does not need explanation. In Schechter (2010), I attributed this idea to Lewis, but I now believe that the attribution is mistaken. Lewis presents this idea in the course of providing a *reductio* argument. Views that resemble Clarke-Doane's can be found in Burgess and Rosen (1997), pp. 45–7, and Pust (2004).

<sup>22</sup> See Clarke-Doane (2017), p. 25.

<sup>23</sup> *Ibid.*, pp. 23–4.

<sup>24</sup> I borrow this example from Peacocke (2004), p. 75, who uses it for a different purpose.



that the most efficient way to pack spheres in a plane is in a hexagonal array.)<sup>25</sup> More generally, explanatory relations are dependence relations. The explanandum depends in some sense on the explanans. Providing information about initial conditions that raises the probability of some phenomenon is not a dependence relation. So Baras's account is not a good candidate for being the right account of explanation.

In reply to Clarke-Doane, it is worth noting an odd feature of his discussion. He argues that the only legitimate explanatory demands concerning our reliability are to show that our beliefs are safe and sensitive. But these explanatory demands are particular to facts about beliefs. One would like to have a general conception of explanation that applies to the special case of explanations of our reliability. Clarke-Doane does not provide such a conception. It is not clear what general conception could be provided that, when applied to the claim that our beliefs are reliable, yields the result that the only explanatory demands are to show that our beliefs are safe and sensitive. Unfortunately, the nature of explanation is a vexed topic, and I do not have a general account of explanation to provide. But my strong suspicion is that showing that our beliefs are safe and sensitive is not closely connected to explaining their reliability—safety and sensitivity are counterfactual notions, and what we need to explain is a fact about the actual correlation between our beliefs and the truth, not anything about what happens in nearby modal space.

There is a more general problem that applies to both Baras and Clarke-Doane's arguments. The underlying intuitions supporting their views seem to rest on the claim that explanation is closed under logical consequence. If we possess an explanation for A and an explanation for B, then putting those explanations together yields an explanation for the conjunction A&B. Similarly, if we possess an explanation of A and A entails C, then the explanation of A can be turned into an explanation of C. The reason that both Baras and Clarke-Doane seem to presuppose these claims is that they both seem to accept that if we have an explanation of why certain claims are true—e.g., because they are necessary or because they are logically entailed by basic necessary truths or the like—and we have an explanation of why we believe those claims—e.g., an explanation that relies on some evolutionary, psychological, or sociological account—then we can put those explanations together to come up with an explanation of our reliability.

The problem is that these general claims about explanation are false. For example, suppose one flips an apparently fair coin one thousand times and the coin turns up heads every time. Suppose for each flip, one has information about the exact motion of the coin, the air currents in the room, and so forth, that explains why that very flip ended up heads. It seems to me that even if the information about each flip explains why that flip ended up heads, the total body of information about the flips does not explain why the coin landed heads every time. Assuming the flips were not all made in the same way, the air currents were not all the same, and so forth, it remains

---

<sup>25</sup> There are other plausible cases of explanations that do not involve information about initial conditions, including explanations that appeal to conservation laws, explanations that involve homeostatic properties, functional explanations, and explanations of physical phenomena appealing to mathematical theorems.

striking that the coin landed heads every time. This is a counterexample to the claim that explanation is closed under logical consequence.<sup>26</sup>

In reply to this kind of objection, Baras argues that we should distinguish single-premise and multi-premise closure principles for explanation. The single-premise closure principle states that if we possess an explanation of a fact A and A entails some other fact C, then we possess an explanation of C. The multi-premise closure principle states that if we possess an explanation of each of  $A_1, \dots, A_n$ , and these facts jointly entail some other fact C, then we possess an explanation of C. Baras claims that the single-premise principle is true and the multi-premise principle is false, but that all he needs is the former principle.

I don't see why Baras's argument (or Clarke-Doane's argument) only requires the single-premise principle. The background thought presupposed by their arguments seems to be that if we possess an explanation of the truth of various claims and an explanation of our beliefs about those claims, then the two explanations can be turned into an explanation of our reliability. This involves an appeal to a multi-premise closure principle for explanation. Moreover, the single-premise principle seems to be just as problematic as the multi-premise principle. Consider the coin flipping case. It seems to me that we have a good explanation of the conjunction: the first flip was heads and the second flip was heads and ... and the one thousandth flip was heads. But we do not have a good explanation of why every flip was heads. If that's right, the single-premise closure principle is false, as well.

## Modal Security

The next objection I'd like to discuss is also due to Clarke-Doane. He (tentatively) presents a general principle governing undermining defeat. Recall the distinction between rebutting and undermining defeat.<sup>27</sup> Very roughly, a rebutting defeater of a belief provides direct evidence that the belief is false. An undermining defeater of a belief provides direct evidence that (some of) the grounds for the belief do not in fact support the truth of the belief. For instance, suppose I believe that my desk drawer contains blank paper. Seeing that my desk drawer is empty rebuts this belief. By contrast, suppose that I believe that a wall is red (and not white) on the basis that it looks red. Learning that the wall is being illuminated by red lights undermines this belief.

Clarke-Doane states the general principle as follows:

---

<sup>26</sup> See Schechter (2010), p.447 for a somewhat more complicated example. It is worth distinguishing two claims here. First, it may be true that we have apparent explanations of each of several facts, but when we consider the facts together, we have reason to doubt the explanations. Second, it can be true that we have genuine explanations of each of several facts, but lack an explanation of some logical consequence of the facts. I think both of these claims are true, and coin cases can be used to illustrate each of them.

<sup>27</sup> See Pollock (1986). Pollock distinguishes between rebutting and undercutting defeaters. My way of presenting the distinction differs from his.

Modal Security: If information E undermines all of our beliefs of a kind F, then it does so by giving us reason to doubt that our F-beliefs are both sensitive and safe.<sup>28</sup>

Here, sensitivity and safety are defined just as above.

Notice that this principle is restricted to undermining defeat and does not apply to rebutting defeat. Moreover, as presented, it only concerns information that undermines all of our beliefs on some subject matter (or that otherwise hang together as a kind). But, presumably, a more general principle will apply to information that undermines an individual belief. It is also important to note that the principle does not merely say that undermining information provides us with reason to doubt that the relevant beliefs are sensitive and safe. Rather, the principle makes a stronger claim: the way that undermining information works is by providing us with reason to doubt that the relevant beliefs are both sensitive and safe.

Why endorse Modal Security? Clarke-Doane puts forward two reasons. The first is that the principle fits familiar examples of undermining defeat. For instance, when I learn that a wall that looks red is being illuminated by red lights, this shows that my earlier belief that the wall is red is not sensitive. If the content of my belief had been false—and the wall was not in fact red—I still would have believed that it was red on the basis of its appearance. The second reason is more general. If some information does not rebut a belief, and if it does not show that the belief is unsafe or insensitive, then it is hard to see how the information could in any way tell against retaining the belief. The belief would seem to be secure, at least with respect to nearby regions of modal space.

Clarke-Doane goes on to argue that if Modal Security is true, then the reliability challenges for logic, mathematics, and morality are not effective. Reliability challenges are meant to provide undermining rather than rebutting defeaters. They do not provide direct evidence that our beliefs are unreliable. Moreover, argues Clarke-Doane, a reliability challenge does not work by providing us with reason to doubt that the relevant beliefs are sensitive or safe. Take the case of logic. Suppose we start out believing that (i) logical truths are necessary truths, and so our logical beliefs are safe; and (ii) we could not have easily come to have false logical beliefs, and so our logical beliefs are sensitive. That we have reason to think that there is no satisfying explanation of our reliability about logic does not directly tell against either (i) or (ii), and so according to Modal Security cannot undermine our logical beliefs.

There are several issues with Clarke-Doane's argument. As a preliminary point, notice that Clarke-Doane understands the reliability challenge for a domain to provide an argument against our reliability about that domain. This is to misconceive the challenge. Reliability challenges should not be understood as arguments to the conclusion that our beliefs are unreliable. Rather, they generate a tension in our overall package of views. In many cases, rejecting our reliability is

---

<sup>28</sup> Clarke-Doane (2017), p. 30. This principle has been much discussed (and objected to) in the recent literature. See, for example, Berry (ms), Dogramaci (2017), Faraci (ms), Klenk (2018), ch. 7, Tersman (2016), and Woods (2018).

implausible. A better response is to give up the objectivity of the domain, or its apriority, or something else entirely. So for Modal Security to show that the reliability challenge is ineffective, it would have to be argued that the reliability challenge does not work by casting doubt on the safety or sensitivity of (say) our belief that the relevant domain is objective. Perhaps Clarke-Doane's argument could be reformulated in that way. It's not entirely clear to me whether this can be done.<sup>29</sup>

A different problem with Clarke-Doane's argument concerns the formulation of Modal Security. It is not clear that the principle should require that undermining information works *by* providing us with reason to doubt that the relevant beliefs are sensitive and safe. The two motivations for Modal Security that Clarke-Doane puts forward do not motivate that aspect of the principle. And if that aspect of the principle is removed, Modal Security will no longer have the impact that Clarke-Doane suggests. That is because a proponent of a reliability challenge could argue that the challenge calls into question the truth of the relevant beliefs, which indirectly casts doubt on their safety and sensitivity. (If a belief is false, it is neither safe nor sensitive.) Indeed, if that modification were made to the principle, Modal Security would no longer provide any real constraint on undermining defeat.

A further problem with Modal Security is that the principle is not well-motivated. Clarke-Doane's first, example-based motivation does not clearly support Modal Security as opposed to some other principle. And his second motivation is very close to a restatement of the principle.

A natural way to try to motivate Modal Security is by presenting a conception of the force of evidence and how it can be undermined. Consider the following picture of what it is for evidence E to be sufficient to justify belief in claim C: E justifies belief in C because it indicates the truth of C, where this is understood as the combination of the following two claims. (i) It would not have been easy for E to be present and C to be false; and (ii) If C were false, E would not have been present. If this were the correct conception of the force of evidence, then Modal Security would be a very natural principle to accept. Modal Security, in effect, says that information undermines the force of evidence when it directly provides reason to doubt (i) or (ii). The problem with this approach to motivating Modal Security is that the proposed conception of the force of evidence is false. This can be seen by considering a case where the evidence is misleading and justifies belief in a false proposition. In such a case, it would have been easy for E to be present and for C to be false, since that's the actual situation. Moreover, it is not true that if C were false, E would not be present, since again it is actually the case that C is false and E is present. In a case of misleading evidence, then, both (i) and (ii) are false. It is not clear what alternative conception of the force of evidence could be provided that would motivate Modal Security.<sup>30</sup>

---

<sup>29</sup> Warren (2017), fn. 9, makes a similar point.

<sup>30</sup> There is another difficulty with the motivation for Modal Security, at least for certain domains. Plausibly, our basic logical (and mathematical and moral) beliefs are not justified by evidence. What could such evidence be?

A final problem with Modal Security is that there are intuitive counterexamples to this principle.<sup>31</sup> For example, suppose a thinker is initially justified in believing some alleged necessary truths (e.g., concerning mathematics, modality, morality, or the constitution of physical kinds). Perhaps the thinker is justified on the ground that the claims seem intuitively correct and the thinker's seemings have a good track record. Suppose the thinker then learns that having those seemings (and only those seemings) was a side effect of having ancestors that were selected for having some very different trait—for instance, having bipedal locomotion. (Suppose, for instance, that the easiest way for the thinker's ancestors to have evolved bipedal locomotion involved coming to have a brain organization that generated those intuitions.) Learning this information should undermine the thinker's justification for those beliefs (or for taking the domain to be objective, etc.). This is so despite the fact that the information does not directly tell against the safety or sensitivity of the beliefs. It does not directly tell against the necessity of the contents of the beliefs. And it does not directly tell against the claim that the thinker could not easily have had different beliefs. If anything, it provides further support for that claim.

For a second counterexample, suppose that a thinker learns that the reason that people believe that water is composed of H<sub>2</sub>O is that a secret society originally formed in the early 1200s has as a core tenet the claim that water is composed of H<sub>2</sub>O (on the basis of no evidence whatsoever). Suppose the thinker learns that the secret society is very large and has spent the past centuries infiltrating every major organization, faking scientific experiments, and spreading its core belief about water. Learning this information should undermine the thinker's belief that water is composed of H<sub>2</sub>O. Again, this is so despite the fact that the information does not directly tell against the safety or the sensitivity of the belief that water is composed of H<sub>2</sub>O. If that's right, then Modal Security is false.

### **Counterlogical Counterfactuals**

The final two objections that I'd like to consider are due to Hartry Field. They are both specific to the case of logic. The first concerns the intuitive motivation behind the reliability challenge for logic.

---

Rather, the epistemological story in such basic cases has to be different. Our basic logical beliefs are (perhaps) default justified, or (perhaps) justified in virtue of being meaning constituting, or (perhaps) justified in virtue of being indispensable to rationally required projects, or something similar. (See Schechter (forthcoming) for a view of the justification of basic deductive rules that does not involve evidence.) If that's right, then even if a principle like Modal Security applies to underminers of beliefs that are justified by evidence, it may not apply to underminers of basic logical, mathematical, and moral beliefs. I don't want to put much weight on this problem, however, since it seems to depend on understanding the reliability challenge as undermining our logical (or mathematical or moral) beliefs, which is not how we should conceive of the challenge. But if I'm wrong about that, this problem for Modal Security may be significant.

<sup>31</sup> A particularly nice counterexample appears in Berry (ms).

Field argues that “[t]he ... problem in mathematics or logic seems to arise from the thought that we would have had exactly the same mathematical or logical beliefs, even if the mathematical or logical facts were different; because of this, it can only be a coincidence if our mathematical or logical beliefs are right, and this undermines those beliefs.”<sup>32</sup> But, Field argues, we cannot make sense of a counterfactual whose antecedent is that the logical facts are radically different. He writes, “But in the logical case, we have no idea how to determine what we would have believed had the logical facts been different: reasoning about what our beliefs would be in alternative circumstances requires logic, and if we contemplate a radically altered logic we have no idea how to conduct the reasoning. This seems to undermine the intelligibility of the counterfactual (about what we would have believed given different logical facts); in which case we have undermined...the problem itself.”<sup>33</sup>

The idea here is not that the supposition that the logical facts are radically different from how they actually are is necessarily false, and so any counterfactual with that supposition as its antecedent is senseless. Field grants that there can be sensible (and non-trivial) counterfactuals with necessarily false antecedents. Indeed, he thinks we can make sense of what would be the case if various mildly non-classical logics were true. For instance, if intuitionistic logic were correct, the law of excluded middle would be invalid, several classical theorems would be false, and so forth. It is false that if intuitionistic logic were correct, the physical world would be a very different place. The problem, rather, is that we cannot make sense of counterlogical counterfactuals where the antecedent states that logic is radically different from how it actually is. We cannot reason about such scenarios since our reasoning depends on the logic in play, and we have no clear sense of how we should reason if a radically different logic were correct. So we cannot make sense of the counterfactuals. Since such counterfactuals are what provide the reliability challenge with its bite, the reliability challenge for logic has no force. (And presumably the same is true for the reliability challenge for deductive inference.) Or so Field argues.

There are a few things that are unclear to me about Field’s argument. I do not know if the problem is supposed to be epistemic—we do not know the truth value of the relevant counterfactuals—or semantic—the counterfactuals are literally senseless. If the latter is Field’s view, I don’t know why the view isn’t instead that the counterfactuals are trivially true. On the standard accounts of counterfactual conditionals, if there are no worlds (either possible or impossible) at which the antecedent of the counterfactual is true, the entire counterfactual turns out to be trivially true no matter what the consequent is. Of course, Field’s argument could be modified to handle the claim that the relevant counterlogical counterfactuals are all trivially true. All he would have to say is that for the reliability challenge for logic to have force, it has to be false that if logic were radically different, we would have the corresponding true logical beliefs. Finally, I’m not fully convinced that when we consider a counterfactual scenario in which some

---

<sup>32</sup> Field (2005), p. 81.

<sup>33</sup> Ibid. Also see Field (1996), pp. 373–7 for an earlier presentation of this line of thought.

alternative logic is correct, we should reason in the alternative logic (or, more precisely, employ the deductive rules corresponding to the alternative logic). Instead, it seems to me that we should use our actual deductive rules to examine what would be true if the logical facts were different.

Putting all of those issues aside, however, there is a crucial error in Field's argument. For a reliability challenge to have force does not require that we endorse the claim that we would have had exactly the same beliefs (or inference rules) if the relevant facts were different. Instead, the challenge arises from the thought that we could have gotten it wrong. In general, what gives the reliability challenge its bite is the thought that there are counterfactual scenarios in which we are not reliable. Such scenarios come in two main varieties—the facts could be very different without us having correspondingly different beliefs (or rules), or the facts could be the same and we could have very different beliefs (or rules). For the case of logic, perhaps we cannot make sense of counterfactual scenarios of the first variety. But we certainly can make sense of counterfactual scenarios of the second. It's easy to imagine (and reason about) scenarios in which the logical facts are the same but in which we have false beliefs about logic or employ non-truth-conducive deductive rules. And that is sufficient for there to be a challenge.

### **Selection of and Selection for**

The final objection that I'd like to consider is also due to Field.<sup>34</sup> It targets the prospect of an evolutionary response to the reliability challenge for logic (and deduction). The objection is perhaps the most interesting of the objections that I have canvassed in this paper. Field's discussion is rather compressed. In what follows, I'll expand on his discussion, staying close to what I consider to be the crux of the problem.

Consider the simplest kind of explanation by natural selection of our possession of a trait. Such an explanation will say that our ancestors were selected for having the trait. The trait is heritable. And that explains why we, their descendants, have the trait, too. (There are, of course, much more complex evolutionary explanations.) In understanding how such an evolutionary explanation works, it is important to distinguish between there being *selection of* a trait and there being *selection for* a trait.<sup>35</sup> There is selection of a trait in a population when organisms with that trait on average have more evolutionary fitness than organisms without that trait, and the trait therefore increases in frequency in the population. There is selection for a trait when having the trait is what explains why the trait increases in frequency.

The canonical illustration of this distinction concerns the heart. Hearts have many properties. For instance, they pump blood. They also make a characteristic lub-dub noise. In our ancestral population, there was selection of both the trait of having an organ that pumps blood and the trait

---

<sup>34</sup> See Field (2005), pp. 80–1. Also see Field (1998), p. 19 for an earlier presentation of the point.

<sup>35</sup> Sober (1984), pp. 98–101.

of having an organ that makes a lub-dub noise. This is because the two traits are tightly correlated. So, for either trait, organisms in the population with that trait on average had more evolutionary fitness than organisms without the trait. Of course, the trait that was selected for was the trait of having an organ that pumps blood. Having an organ that pumps blood is what caused our ancestors to have greater success in survival and reproduction. Pumping blood—and not making lub-dub noises—is what explains the increased frequency of hearts in the population. That is why it is the trait that was selected for.<sup>36</sup>

Given two correlated traits that there was selection of, which one was selected for? There is a test for determining the answer to this question. Consider the following counterfactual: If pumping blood was not correlated with making lub-dub noises, the trait of having an organ that pumps blood would have increased in frequency in the population but the trait of having an organ that makes lub-dub noises would not have increased in frequency. It is the truth of this counterfactual, or one of roughly the same shape, that tells us that what was selected for was the trait of having an organ that pumps blood. More generally, the test is to look at counterfactuals that concern what would have happened if the two traits had been uncorrelated. If one trait would have increased in frequency but not the other, the former trait was the trait that was selected for. For ease of reference, let us call such counterfactuals “correlation counterfactuals”.

In the simplest kind of explanation by natural selection of our reliability about logic (or deduction), the relevant trait is that of having reliable logical beliefs (or deductive rules). Such an explanation works by claiming that our ancestors were selected for having the trait of having reliable logical beliefs (or deductive rules)—a heritable trait—and this explains why we, their descendants, have that trait, too.

In his discussion, Field uses the phrase “accepting the correct logic” to stand for this trait and, in the remainder of this section, I’ll follow his usage. Field points out that there were two correlated traits that there was selection of: (i) the trait of accepting the correct logic, whatever it is; and (ii) the trait of accepting a given logic, which happens to be correct.<sup>37</sup> So, for instance, if classical logic is the correct logic, the latter trait is the trait of accepting classical logic.

Which of these two traits was selected for? The answer that a proponent of an evolutionary explanation has to provide is that the trait of accepting the correct logic, whatever it is, is the trait that our ancestors were selected for. Otherwise, the evolutionary explanation provides an explanation of why we accept some specified logic. It doesn’t provide an explanation of why we accept the correct logic.

This is where the problem arises. Field argues that we cannot make sense of a counterfactual scenario in which the trait of accepting the correct logic, whatever it is, is uncorrelated with the trait of accepting a given logic which happens to be correct. Again, the problem is not that this

---

<sup>36</sup> Of course, this is a gross oversimplification. Hearts did not appear in the relevant population in one fell swoop.

<sup>37</sup> The two traits correspond to a *de dicto* and a *de re* reading of “accepting the correct logic”.



scenario is an impossible one—Field grants that we can reason about impossible scenarios. Rather, as before, the problem is that, to make sense of a counterfactual scenario in which the correct logic is different, we would have to reason using deductive rules that correspond to the alternative logic in question, and we do not know how to do so. Field writes, “But how are we to argue what would be selected for in a world with an alternative logic? We would apparently need to conduct the argument in the alternative logic in question, and we have so little idea how to do this that the counterfactual begins to look nonsensical.”<sup>38</sup> Thus, Field concludes, an explanation of our reliability about logic that appeals to natural selection simply doesn’t work.

Before I turn to problems with Field’s argument, it is worth noting two initial points. First, for Field’s argument to work, the trouble with the counterlogical correlation counterfactuals cannot be merely epistemic. It cannot be just that we do not know how to directly evaluate such counterfactuals. Rather, the problem had better be semantic. Our reliability about logic is a striking fact. This provides evidence that there is some explanation of our reliability. If there are no good alternatives to an evolutionary explanation, this provides evidence in favor of an evolutionary explanation. If an evolutionary explanation of our reliability requires that what was selected for was the trait of accepting the correct logic whatever it is (and not the trait of accepting some specific logic), this provides evidence in favor of that claim. Thus, for Field’s argument to pose a problem, it had better be that the counterlogical correlation counterfactuals are meaningless, not merely difficult to evaluate. (I’m not convinced that the counterlogical counterfactuals are meaningless, since on the standard accounts of counterfactual conditionals, they turn out to be trivially true. But in what follows, I’ll assume that they are meaningless at least for argument’s sake.) Second, Field’s objection generalizes beyond the case of evolution by natural selection. The crux of the problem is that “accepting the correct logic” cannot play a causal explanatory role, since there is no way to distinguish between the case where accepting the correct logic plays that role and the case where accepting a given logic plays that role. So an explanation that appeals to cultural evolution—or, indeed, any explanation that claims that accepting the correct logic plays a causal explanatory role—faces the very same difficulty.

Let me now turn to the difficulties facing Field’s argument. There are several. One problem is that the objection overgeneralizes. Consider any selective explanation of the presence of a trait in a population. For instance, consider the evolutionary explanation of why we have an organ that pumps blood. Consider the (somewhat contrived) trait of having an organ that pumps blood and is such that classical logic is correct. This trait is correlated with the trait of having an organ that pumps blood. Which trait was selected for—(i) the trait of having an organ that pumps blood, or (ii) the trait of having an organ that pumps blood and is such that classical logic is correct? Following Field’s line of thought, there seems to be no way to answer this question, and so the evolutionary explanation has to be mistaken. That is the wrong verdict.

---

<sup>38</sup> Field (2005), p. 81. Field does not distinguish between logic and deductive inference. If Field were to make this distinction, this passage would have to be slightly modified.

The problem here can be raised for any evolutionary explanation. Indeed, it can be raised for any causal explanation whatsoever. The abstract form of the difficulty is that whenever some fact F plays an alleged causal role in explaining some fact G, there will be an alternative fact F' logically equivalent to F that will also be a candidate for explaining G. (For instance, F' might be the conjunction of F with the claim that everything is self-identical.) We can't rule out F' by claiming that in the counterfactual scenario where F and F' are uncorrelated, F is what does the work. If F' is properly chosen, such a correlation counterfactual will be senseless. So, according to Field's line of thought, F does not play a causal role in explaining G. That cannot be right.<sup>39</sup>

A second problem is that the ultimate conclusion of Field's argument doesn't obviously follow from what leads up to it. Suppose we grant that counterfactual correlation counterfactuals of the form, "If accepting the correct logic, whatever it is, were uncorrelated with accepting a given logic, which happens to be correct, then ..." are senseless. Why should we conclude that neither of the two traits was selected for? Field does not provide a clear argument for this conclusion. An alternative conclusion one could draw is that each of the two traits was selected for. Indeed, a rough and ready test for whether a fact F plays a causal role in explaining some fact G is whether the following counterfactual is true: if it weren't the case that F then it would not have been the case that G. For each of the two traits, it is plausible that if it wasn't the case that our ancestors had the trait, we wouldn't have accepted the correct logic. So there is reason to think that at least one of the traits, and possibly both of them, play a role in the explanation of our acceptance of the correct logic.<sup>40</sup>

A third problem is that, for Field's argument to work, correlation counterfactuals cannot be merely of heuristic value in determining which of various traits were selected for. Otherwise, we could respond to Field by claiming that the heuristic fails in this instance. (It would not be surprising for a heuristic to fail in the case where the relevant traits involve the acceptance of a logic.) But it is implausible that correlation counterfactuals provide necessary conditions on selection for. Selection for is a causal explanatory notion. A trait is selected for when the presence of the trait explains its increase in frequency in the population. The truth of various counterfactuals typically provide useful diagnostics of such explanatory connections, but do not provide necessary conditions on them. (Or, to put the point more cautiously, it has proved difficult to identify counterfactuals which provide necessary conditions on causal explanation.)<sup>41</sup> Correlation counterfactuals may be particularly suspect here, since such counterfactuals involve comparing a trait with other traits. But the question of whether a trait is selected for is wholly a

---

<sup>39</sup> Could Field reply to this objection by arguing that F' has F as a conjunct, and so F' is a worse candidate explainer? Perhaps. But it is not clear to me that this response would work. It is also not clear to me that this response couldn't be applied to the case of explaining our reliability about logic, as we will see below.

<sup>40</sup> The same is true for other tests of whether a fact plays a causal role, such as tests involving conditional probabilities or the manipulation counterfactuals that appear in structural equation-based theories of causation.

<sup>41</sup> Moreover, counterfactual diagnostics seem out of place for other kinds of explanation. Some mathematical proofs are explanatory and others are not. It is difficult to capture this difference using counterfactuals. The same is true for explanations internal to morality, and for grounding explanations.

matter of whether that very trait played an explanatory role. No comparison with another trait is required.

There are two additional problems worth noting. The first is that there is reason to think that the trait that plays an explanatory role in an evolutionary explanation of our reliability about logic is the trait of accepting the correct logic, not the trait of accepting some given logic. Consider what the two candidate explanations would look like. One explanation goes like this: Our ancestors were selected for accepting the correct logic. This trait is heritable, and that explains why we accept the correct logic. By contrast, the other explanation goes like this: Our ancestors were selected for accepting some given logic. This trait is heritable, and that explains why we accept the correct logic. This second candidate explanation, unsupplemented, involves a non-sequitur. The fact that our ancestors had the heritable trait of accepting some given logic does not by itself explain why we accept the correct logic. It must be added to the explanation that the given logic is in fact the correct logic. The need for this addition makes the second candidate explanation a worse explanation than the first. This provides reason to think that the trait that was selected for is the trait of accepting the correct logic, whatever it is, and not the trait of accepting a given logic.

The final point I'd like to make is that, if we are sufficiently careful, we may be able to distinguish between selection for accepting (something like) the correct logic and selection for accepting a given logic. The reason is that the right trait to focus on is not the trait of accepting the correct logic, whatever it is. According to a more plausible evolutionary explanation, our ancestors were selected for accepting a logic that is (i) approximately correct and perhaps exactly correct in a core range of applications; (ii) sufficiently powerful to be useful; and (iii) sufficiently simple and computationally tractable to be usable. If we compare the trait of accepting a logic with these three properties with the trait of accepting some given logic, it is plausible that there was selection for the former rather than the latter trait. This is because there is (plausibly) more than one logic that has these three properties. If, for instance, classical logic is the correct logic, both classical and intuitionistic logic plausibly count. If it turns out that, on alternative evolutionary paths, the relevant population would have ended up accepting one of several alternative logics with these three properties, this suggests that the trait that was selected for was the trait of accepting a logic with the three properties. In response, one might propose that what was selected for was the trait of accepting one of some specified list of logics. But if the logics on the list all have the three properties, it is a much more natural view that what was important is what is in common between the logics, not that they appear on some list. This provides a way of distinguishing between the traits.

## Conclusion

As we have seen, there are a number of pressing objections to there being a reliability challenge for logic (and deduction) and to an evolutionary answer to that challenge. If what I've argued here is correct, these objections are not successful. Nevertheless, considering the objections helps to clarify the nature of the reliability challenge, the nature of evolutionary explanations, and other important philosophical issues.<sup>42</sup>

## References

- Baras, Dan. (2017). Our Reliability is in Principle Explainable. *Episteme* 14(2): 197–211.
- Benacerraf, Paul. (1973). Mathematical Truth. *The Journal of Philosophy* 70(19): 661–79.
- Berry, Sharon. (ms). Coincidence Avoidance and Formulating the Access Problem.
- Burgess, John and Rosen, Gideon. (1997). *A Subject with No Object*. Oxford: Oxford University Press.
- Clarke-Doane, Justin. (2017). What is the Benacerraf Problem? In Fabrice Pataut (ed.), *New Perspectives on the Philosophy of Paul Benacerraf: Truth, Objects, Infinity*. Cham, Switzerland: Springer, pp. 17–43.
- Dogramaci, Sinan. (2017). Explaining our Moral Reliability. *Pacific Philosophical Quarterly* 98(51): 71–86.
- Enoch, David. (2011). *Taking Morality Seriously*. Oxford: Oxford University Press.
- Faraci, David. (ms). Knowing What's Necessary: How Modal Conditions Threaten to Trivialize Ethical (and Other) Knowledge.
- Field, Hartry. (1989). *Realism, Mathematics and Modality*. Oxford: Basil Blackwell.
- Field, Hartry. (1996). The A Prioricity of Logic. *Proceedings of the Aristotelian Society* 96(1): 359–79.
- Field, Hartry. (1998). Epistemological Nonfactualism and the A Prioricity of Logic. *Philosophical Studies* 92(1/2): 1–24.
- Field, Hartry. (2005). Recent Debates about the A Priori. *Oxford Studies in Epistemology* 1: 69–88.
- Harman, Gilbert. (1977). *The Nature of Morality*. Oxford: Oxford University Press.

---

<sup>42</sup> Thanks to Dan Baras, Mark Benz, and Michael Klenk for helpful discussion of some of these issues.

- Harman, Gilbert. (1986). *Change in View*. Cambridge, Mass.: The MIT Press.
- Horwich, Paul. (2006). A World without Isms: Life after Realism, Fictionalism, Non-Cognitivism. In Patrick Greenough and Michael P. Lynch (eds.), *Truth and Realism*. Oxford: Oxford University Press, pp. 188–202.
- Klenk, Michael. (2018). *Survival of Defeat: Evolution, Moral Objectivity, and Undercutting* (Doctoral dissertation). Utrecht University, Utrecht.
- Lewis, David. (1986). *On the Plurality of Worlds*. Oxford: Blackwell Publishers.
- Peacocke, Christopher. (2004). *The Realm of Reason*. Oxford: Oxford University Press.
- Pollock, John. (1986). *Contemporary Theories of Knowledge*. First edition. Towota, NJ: Rowman and Littlefield Publishers.
- Pust, Joel. (2004). On Explaining Knowledge of Necessity. *Dialectica* 58(1): 71–87.
- Schechter, Joshua. (2010). The Reliability Challenge and the Epistemology of Logic. *Philosophical Perspectives* 24, *Epistemology*, 437–64.
- Schechter, Joshua. (2013). Could Evolution Explain Our Reliability about Logic? *Oxford Studies in Epistemology* 4: 214–39.
- Schechter, Joshua. (2018). Explanatory Challenges in Metaethics. In Tristram McPherson and David Plunkett (eds.), *The Routledge Handbook of Metaethics*. New York: Routledge, pp. 443–58.
- Schechter, Joshua. (forthcoming). Small Steps and Giant Leaps in Thought: The Epistemology of Basic Deductive Rules. In Magdalena Balcerak Jackson and Brendan Balcerak Jackson (eds.), *Reasoning: New Essays on Theoretical and Practical Thinking*. Oxford: Oxford University Press.
- Tersman, Folke. (2016). Explaining the Reliability of Moral Beliefs. In Neil Sinclair and Uri Leibowitz (eds.), *Ethics and Explanation*. Oxford: Oxford University Press.
- Sober, Elliott. (1984). *The Nature of Selection*. Cambridge, Mass.: The MIT Press.
- Warren, Jared. (2017). Epistemology versus Non-Causal Realism. *Synthese* 194(5): 1643–62.
- White, Roger. (2005). Explanation as a Guide to Induction. *Philosophers' Imprint* 5(2): 1–29.
- Woods, Jack. (2018). Mathematics, Morality, and Self-Effacement. *Noûs* 52(1): 47–68.