CrossMark

Episteme, 15, 2 (2018) 154–165 © Cambridge University Press
doi:10.1017/epi.2016.52

C. I. LEWIS AND THE BENACERRAF PROBLEM

BOB FISCHER
fischer@txstate.edu

ABSTRACT

Realists about modality offer an attractive semantics for modal discourse in terms of possible worlds, but standard accounts of the worlds – as properties, propositions, or causally isolated concreta – invoke entities with which we can't interact. If realism is true, how can we know anything about modal matters? Let's call this "the Benacerraf Problem." I suggest that C. I. Lewis has an intriguing answer to it. Given that we're willing to disentangle some of Lewis's insights from his phenomenalism, we can take the following line. If the Benacerraf Problem is a genuine one, then it threatens all knowledge – not just modal knowledge. But then it leads to a general and implausible form of skepticism, not a limited and more plausible debunking argument. Hence, whatever we're willing to say about skepticism we should say about the Benacerraf Problem.

INTRODUCTION

According to Benacerraf's Dilemma, there's a tension between the semantics and the epistemology of mathematics. We'd like to adopt a Tarskian semantics for mathematics, as it allows us to have a unified account of truth across a number of different domains. However, that leads to to Platonism, which involves positing entities to which we bear no causal relations – namely, abstracta. The logicist drops the Tarskian semantics, and therefore gets a straightforward epistemology. However, her account of mathematical truth is objectionable for familiar reasons.

There is an analogous dilemma for modality. Realists about modality offer an attractive semantics for modal discourse in terms of possible worlds, but standard accounts of the worlds – as properties, propositions, or causally isolated concreta – invoke entities with which we can't interact. Conventionalists can offer straightforward epistemologies, but they manage this only by analyzing modal claims in ways that undermine much of their interest. In mathematics and modality, a plausible account of truth doesn't sit easily with a plausible account of knowledge.[1]

1 The dilemma stems from Benacerraf (1973), and for a helpful discussion of its intricacies, see Clarke-Doane (Forthcoming). It isn't clear who first applied the dilemma to modality. Various philosophers discuss the problem in the 1980s, but those concerned with modality in the recent literature tend to focus instead on Christopher Peacocke's (1999) "integration challenge." The integration challenge is just the generalized version of Benacerraf's Dilemma: there's often a tension in metaphysics between truth and knowledge, and theories in a domain have to attend to both concerns. Of course, the Benacerraf Problem isn't just a problem in mathematics and modality – e.g., it affects logic too – and

https://doi.org/10.1017/epi.2016.52 Published online by Cambridge University Press

Let's suppose that we're committed to realism about modality (i.e., the denial of conventionalism and its ilk, not necessarily David Lewis's modal realism), and so face that horn of the dilemma.[2] Let's call it, "the Benacerraf Problem." What, exactly, is the issue? Well, Gettier cases show that knowledge is incompatible with (one kind of) luck. One plausible solution to the luck problem requires some causal connection between the knower and the known, where this connection explains the knower's epistemic success. But if realism about modality is correct, then we bear no causal relations to the truthmakers for modal claims. Therefore, if the realist can't provide an alternate solution to the luck problem, she makes our epistemic success unintelligible; on her view, it is unclear how we can have any modal knowledge whatever.

In what follows, I suggest that C. I. Lewis has an intriguing answer to the Benacerraf Problem. Given that we're willing to disentangle some of Lewis's insights from his phenomenalism, we can take the following line. If the Benacerraf Problem is a genuine one, then it threatens all knowledge – not just modal knowledge. But then it leads to a general and implausible form of skepticism, not a limited and more plausible debunking argument. Hence, whatever we're willing to say about skepticism we should say about the Benacerraf Problem.

I begin with Lewis's challenge to pure extensionalists like Russell and Whitehead. That challenge helps us appreciate Lewis's two related routes to the idea that knowledge of actuality has modal content. I then return to the Benacerraf Problem, outlining a Lewisian solution, and conclude by replying to three objections.

1. LEWIS AND THE EXTENSIONALISTS

Here's a no-frills summary of Lewis' challenge to Russell and Whitehead. Recall that their logic includes a very familiar relation: material implication. As Lewis points out, material implication has some odd features that make it ill-suited to capture the logic we use in everyday reasoning – characteristics that result from the logic being purely extensional. The first is that every false sentence implies every sentence whatever; the second is that every sentence whatever implies every true sentence. Common sense disagrees. So, material implication does not capture every ordinary conditional. This problem is not unique to Russell and Whitehead's system; it plagues any purely extensional logic. So, Lewis argues that we need to adopt an intensional logic that better captures ordinary inference patterns.

Lewis uses many examples to show that a purely extensional logic is ill-suited to capturing some common conditionals, but here's a particularly instructive one. Begin with the observation that we make predictions all the time; they are obviously essential to much of the business of life. Now consider the prediction (expressed using material implication), "If I look around the corner, then I will see the coffee shop." This conditional true is true if the antecedent is false. As it turns out, I don't look around the corner; so, the

it isn't just a problem with abstracta – e.g., it's an issue regarding our knowledge of the future and nonexistent objects.

2 I'm ignoring non-conventionalist metaphysical theories that are designed to avoid the Benacerraf Problem: e.g., D. M. Armstrong's (1989) combinatorialism and Tim O'Connor's (2008) theistic reliablism. This is only because such views don't command wide acceptance. Hence, it seems worth considering alternate strategies for handling the problem.

https://doi.org/10.1017/epi.2016.52 Published online by Cambridge University Press

conditional is true. But if I don't look around the corner, then the following is true too: "If I look around the corner, then I will see Ali Baba and his forty thieves." Since material implication sanctions both, it seems that predictions expressed using material implication are trivial unless the antecedent is true. But surely we want to say that the first prediction is *not* trivial; it expresses a reasonable belief about what would happen in that circumstance, whereas the second is merely the whim of a child who's been reading *One Thousand and One Nights*.[3]

There are two obvious ways to block this argument. The first is to deny that predictions with false antecedents are as important as Lewis believes them to be. So, even if material implication doesn't provide the right results in these cases, its usefulness elsewhere may tip the scales in its favor. The second is to deny that it gives the wrong results in the first place. Lewis thinks that both moves are bound to fail. Why?

In short, he's skeptical about the first move because he thinks that it fails to take seriously the connection between knowledge and agency. He's not optimistic about the second because he contends that it tacitly abandons realism. To say more, we'll need to get a better handle on his views about knowledge, agency, and realism.

1.1 Lewis on Knowledge, Agency, and Realism

Lewis takes knowledge acquisition to be pointless if realism is false, since he thinks we acquire knowledge in order to act as agents in a world that isn't of our making. Here is one provocative statement of this view:

> Independent reality is not something to be proved but an original acknowledgment which all men make confronting the facts of life. . . . To think – that is, to think seriously and to know, as against the idle self-entertainment of the day-dreaming consciousness – is to posit that which is and is other than the apprehension of it, and does not arise with or from the apprehension of it but bespeaks another ground. If independent factuality did not force itself upon us, we should have to invent it in order to exist as being who think and wish to do. (Lewis 1955: 237)

I think that we can appreciate the upshot of the above by asking a question about knowledge *de re*: What's the thing of which we have knowledge? Suppose that the object of your knowledge isn't the sort of thing that – as Lewis puts it in a different paper – "transcends the content of any experience in which it may be known" (Lewis 1954: 194). Suppose instead that the object of your knowledge is nothing more than the experience you have of it, that the content of what you're experiencing doesn't "bespeak another ground." Then, you see no reason to suppose that what you're experiencing is somehow independent of you, and you take it to be a day dream, or hallucination, or optical illusion, or whatever. You certainly don't take the experience to be of "independent factuality," even if you take either the experience itself or its content to be somehow prompted by perfectly real entities.

Lewis is on board with those empiricists who hope to translate talk about the world into talk about experience. He balks, however, at the idea that this can be done with

3 The preceding argument is paraphrased from Lewis (1970a). He makes an analogous argument in his (1912). Note, moreover, that the extensionalist cannot salvage her view by appealing to some metalinguistic analysis; that move only works for definitions and the like.

https://doi.org/10.1017/epi.2016.52 Published online by Cambridge University Press

only the resources of a purely extensional logic. The empiricists that Lewis criticizes only appeal to past, present, or future experiences – not possible ones. So, when they render "All birds have wings" using the material conditional, the result is "(x)(Bx ⊃ Wx)," which says nothing about what birds are like when they aren't being observed. Lewis's complaint is that "All birds have wings" is false if there is a bird that lacks wings at some point during its existence *whether or not it is ever observed*.

On Lewis's view, by contrast, "All bird have wings" is true just in case all *observable* birds have wings. It doesn't matter whether every bird is actually observed:

> The sense meaning of any verifiable statement of objective fact, is exhibitable in some set of terminating judgments each of which is hypothetical in form; it is a judgment that a certain empirical eventuality will ensue if a certain mode of action be adopted. Such judgments may be decisively verified or found false by adopting the mode of action in question and putting them to the test. And it is by such conclusive verification of terminating judgments, constituent in the meaning of it, that the objective belief – the non-terminating judgment – receives its confirmation as more or less highly probable. (Lewis 1946: 211)

Take the sentence, "A book is on the table." This is an objective judgment. He thinks that all objective judgments like this one entail an indeterminate (but very large) number of "terminating judgments" – namely, counterfactuals of the form, "S being given, if A were the case, then E would be the case."[4] Therefore, "A book is on the table" entails (*inter alia*) "Sensory experience as of a book on a table being given, if I were to turn my head to the right, then the book and the table would appear to be displaced." And, if the meaning of a sentence that expresses an objective belief *is* equivalent to the set of terminating judgments it entails, then Lewis seems to have a shot at carrying out the translation project. The result of all this is a tempered verificationism. Although the meaning of every sentence transcends all actual experience, there is no sentence whose meaning outstrips experience entirely. All legitimate claims are verifiable, even if none is actually verified.

The task now is to make explicit the connection between the above and Lewis' thought about agency. Lewis' line of thought – developed most clearly in his (1970b) – goes something like this. If you can use a sentence meaningfully, then you have some grasp of how you might determine whether it's true.[5] But there are, for many sentences, countless possible ways to determine whether a particular sentence is true – or, perhaps more perspicuously, there are indeterminately many possible tests that you could perform to convince yourself of a sentence's truth value. So, using a sentence meaningfully involves having a grasp of at least some of those possibilities.

This view lends itself to the development of an attractive view of agency. Suppose that you believe, as Lewis does, that being an agent involves selecting between possible courses

4 In the transition from *Mind and the World Order* to *An Analysis of Knowledge and Valuation*, Lewis (officially, though not consistently) trades the counterfactual rendering for a probabilistic one: "S being given, if A were the case, then with high probability, E." This formulation is of less interest here, and it isn't clear to me that Lewis actually prefers it.

5 It's reasonable to worry that "having some grasp" will be difficult to explain without relying on meaning, thereby rendering the account viciously circular. But perhaps it can be cashed out in terms of dispositions to assent to some set of propositions of the form, "if A were to be the case, then E would be case," or by appealing to some other, sub-doxastic state.

https://doi.org/10.1017/epi.2016.52 Published online by Cambridge University Press

of action. Then, we need some theory on which agents are aware of the possibilities open to them so that they can make the appropriate selections. Lewis thinks that he has such an account. As I've mentioned, he holds that sentences entail an indeterminately large number of terminating judgments of the form, "S being given, if A were the case, then E would be the case." Following this schema, if you know that S is true (or you justifiably believe that S is true), then you know that the consequent is true: "If A were the case, then E would be the case." So, knowing that S is true involves knowing the truth values of counterfactuals that represent possible courses of action. For example: If I know that I am standing in front of the stove, then I know the counterfactuals:

1. My standing in front of the stove being given, if I were to turn up the burner, then I would see a higher flame.
2. My standing in front of the stove being given, if I were to turn down the burner, then I would see a lower flame.

These aren't the most exciting counterfactuals, but they are certainly useful. If I justifiably believe that "I am standing in front of the stove" is true, then, presumably, I can use it meaningfully. If I can use it meaningfully, then I have some grasp of those conditions the satisfaction of which would help me to determine its truth value. (1) and (2) are just such counterfactuals, and having some grasp of them just is understanding what I can expect in various circumstances based on my knowledge. In this way, knowledge of a sentence's truth includes some modal knowledge – namely, all those counterfactuals entailed by it, the grasp of which is part and parcel of understanding the sentence's meaning.[6]

So Lewis' commitment to the view that all knowledge has some potential practical value is probably a corollary of his verificationism. Every legitimate claim to know has verifiable consequences; in other words, everything that I know tells me what will happen if I act in different ways. Hence, everything that I know is such that I can, in principle, use it to direct my actions. Of course, knowledge wouldn't be so valuable if the world didn't push back on us; so, the purpose of empirical knowledge is to help us to act in a world that doesn't depend on us. We need to be able to decide what to do, so we need to know what would happen were we to act in various ways. These predictions allow us to (try to) avoid various undesirable states of affairs and realize the desirable ones. As he puts it, the "significance of knowledge can be justly phrased by saying that what an objective fact *means* is certain possibilities of experience which are open to realization through our action" (Lewis 1946: 208).

1.2 *Back to Lewis on Material Implication*

Recall that Lewis is attacking the view that you could use material implication to translate the prediction, "If I look around the corner, then I will see the coffee shop." The problem, again, is that if I do not look around the corner, then the following is no less true than the correct prediction: "If I look around the corner, then I will see Ali Baba and his forty thieves." To get around this objection, you can argue that even if material implication

6 Lewis' views about meaning are no longer in vogue, but the connection between agency and counterfactual knowledge is still popular enough. Lucy O'Brien (2007: 120), for example, argues that a subject is an agent only if her actions are based on her "consideration of possibilities, grasped as possibilities."

https://doi.org/10.1017/epi.2016.52 Published online by Cambridge University Press

doesn't provide the right results when the antecedent is false, its usefulness elsewhere may tip the scales in its favor. Alternately, you can deny that it gives the wrong results in the first place.

It should now be clear why Lewis isn't impressed with either option. The first response ignores the intimate connection between knowledge and action. The point of acquiring knowledge is to help us act, and would not be helped at all if we couldn't distinguish between legitimate and illegitimate predictions. The second response is perhaps worse. Not only does it break the connection between knowledge and action in the same way that the first does, but it also is guilty of the naïve sort of subjectivism that Lewis decries. We should, therefore, acknowledge that our beliefs have modal content.

2. PROBLEMS FOR LEWIS

These are important challenges to extensionalism. However, Lewis has troubles of his own. His brand of verificationism, for example, is likely to fail on its own terms. Lewis is committed to the translation project that has occupied many empiricists. Thomas Baldwin points out a deep problem for it:

> [W]hen Lewis mentions the kinds of hypothetical experience relevant to the truth of the judgment 'This penny is round' he does so by means of a conditional such as 'If I take two steps to the right, it will look elliptical'. The obvious point to note is that the antecedent of this conditional concerns an objective change and not an experience at all: so by relying on conditionals of this kind Lewis rules out the possibility of any form of reduction. (Baldwin 2007: 181)

Lewis comes close to acknowledging this difficulty, and he tries to dodge this bullet by stipulating that terminating judgments be interpreted as being solely concerned with experience:

> It may well be that no language is available to us, for such expressive use, except language which in its more usual signification would refer to objective things and states of affairs. [... But it] is not this physical event which, in my terminating judgment, I intend to predict, but merely the passage of experience of itself. (Lewis 1946: 203)

The stipulation won't do. The translation project depends entirely on the availability of the experiential language; if we lack the appropriate vocabulary, then the translation project can't get off the ground.

What's more, while Lewis's "realism" is an improvement over his opponents', it's still a brand of phenomenalism. It lacks the resources to accommodate objects that are not observable at all. For Lewis, the meanings of objective judgments are sets of counterfactuals, the consequents of which specify what you would experience in various circumstances. Now suppose you want to make a claim about some unobservable. That claim's meaning is a set of counterfactuals. But what do the consequents of those counterfactuals describe? Well, either they don't describe anything at all or they describe generic conditions (i.e., those experiential conditions that you would experience whether or not the unobservable object is present). In either case, those counterfactuals won't distinguish the meaning of your claim about unobservable object x and my claim about unobservable

https://doi.org/10.1017/epi.2016.52 Published online by Cambridge University Press

object *y*. For Lewis, meanings are only as fine-grained as human experience. So, to countenance unobservables would be to commit himself to a finer distinction between meanings that his semantics allows. Lewis's rhetoric suggests a full-blooded realism, but he doesn't have the resources to secure it.

The problem here is Lewis's semantics; the meanings of sentences aren't what he took them to be. However, Lewis seems to be right about what's involved in knowing all sorts of ordinary claims about the world. That knowledge underwrites all sorts of counterfactuals. If not via semantics, then how?

3. BEYOND LEWIS: MODELS AND MODAL KNOWLEDGE

Consider two truths:

1. I own a green mug.
2. It's false that, if I were to wiggle my nose, my mug would turn blue.

Why wouldn't my mug turn blue if I were to wiggle my nose? Plainly, my mug's color doesn't depend on (at least some) changes in its environment – my nose-wiggling being just one example. There is nothing special about this counterfactual; we could swap out its antecedent with countless others, all equally banal, and the products of these substitutions would be just as obviously false. It isn't the case that, if I were to crack my knuckles, my mug would turn blue. Nor would it turn blue if I were to eat a salad for lunch. Nor if a ladybug were to crawl across the windowsill. And so on. It's striking that we know all these counterfactuals to be false; it's equally striking that we have so little trouble explaining why – at least in some provisional way.

How could we argue for the falsity of these counterfactuals? As a matter of fact, the colors of objects don't change despite a great number of changes in their environments. There are two possible explanations of this fact. The first is that it might be some grand coincidence that the colors of objects remain stable; the second is that the colors are independent of these variations. Suppose that the coincidence hypothesis is true. If so, then counterfactuals like this are true: had my chair been a half-inch closer to the window, then my mug would not have remained green. However, it would be an absolute miracle if things just happened to unfold in such a way that, though tiny changes would have resulted in Technicolor dishware, we've found the one path on which there's the appearance of stability. The coincidence hypothesis is simply unbelievable, and this is partly based on a point that Lewis stressed. Beliefs are for navigating the world, and so those beliefs needs to delimit the space of relevant possibilities. Leaving open the possibility that my mug's color might be so sensitive to environmental changes doesn't help on this score; it would be counterproductive to the aim of acting. Navigation requires ruling possibilities out (at least for all practical purposes) – not just recognizing what possibilities there are.

If all that's right, then we can think of ordinary claims as part of a more complex picture of the world, an evolving explanation of why we have the experiences we have – and, ultimately, why things behave as they do.

How do we explain phenomena? In part, we do so by providing a model that makes sense of it. The coincidence hypothesis is a model, albeit a poor one. The independence

https://doi.org/10.1017/epi.2016.52 Published online by Cambridge University Press

hypothesis is a better one, though it needs to be filled out. Models represent the world's structure: its parts, the locations of its parts, the relations between its parts. In so doing, they specify the truth values of a very large but (usually) indeterminate number of counterfactuals. If you know that the independence hypothesis is true, it only remains to determine whether it specifies the truth value of the counterfactual in question. If it does, then you know the counterfactual because you know the hypothesis that specifies its truth value.

Here is a simple example. I come home at 4:30 PM; my wife isn't there. I try to explain her absence: she isn't home because she's still at work. I say that this explanation is a very simple model: it represents two locations (home and work) and supposes that my wife is either at one or the other. This is obviously a toy model – most obviously, it ignores her commute between the two locations. Still, it's a useful model for making certain judgments about my wife's location, and it also specifies the truth values of some counterfactuals. If this hypothesis is true, then it's also true that, if my wife weren't at work, she'd be home.

Let's complicate the story. Suppose that my wife works from 8 AM to 4 PM. She stays late about as often as she doesn't, but never later than 5 PM. Moreover, she always goes to the gym after work for an hour or so. On the day in question, she's still at work at 4:30, although I don't know this. I come home at 4:30, I notice that my wife isn't there, and I try to explain her absence: she isn't home because she's still at work. Is this explanation any good?

Plainly not. Here is a better explanation: it's before 5 PM and she's never home any earlier. The model in the present case is more complex. Instead of representing two locations, it needs to represent three; instead of the supposition that my wife is either at home or work, it includes the supposition that my wife is at work before 4 PM, either at work or the gym between 4 PM and 5 PM, and either at the gym or home from 5 PM onward. Hence, this hypothesis specifies different truth values for certain counterfactuals than the previous one: it's not true, according to this hypothesis, that if my wife weren't at work, she'd be home (she could just as well be at the gym). In addition, it also specifies truth values for a broader range of counterfactuals than the previous hypothesis: it's true, according to this hypothesis, that if my wife weren't at the gym, she'd be home. Also, it says that, even if my wife had left work promptly, she wouldn't be home yet.

I think that all explanations work in roughly this way. Most explanations are, of course, vastly more complex; they include many more variables and many more assumptions about the behaviors of and relations between the variables. Some of them introduce new objects, properties, and relations; some simply organize the objects, properties, and relations that we already countenance. Regardless, though, an explanation involves a model of whatever we'd like to explain – a model that has modal implications. Admittedly, it isn't always easy to determine whether a model specifies the truth value of a particular counterfactual, but there are always plenty of counterfactuals the truth values of which it obviously entails.

Crucially, models don't represent possible experiences. Rather, models represent possibilities directly – or, more accurately, they represent sets of possibilities. So Lewis was correct that ordinary knowledge underwrites modal knowledge. But he was wrong about how it does this. Our beliefs aren't about possible experiences, but about mind-independent objects and their configurations. However, they aren't exclusively about *actual* objects and their configurations, since they wouldn't be of any use to us if they were. The purpose of our beliefs is to represent the world, and so to provide us with

https://doi.org/10.1017/epi.2016.52 Published online by Cambridge University Press

tools to act. Our beliefs manage this by forming, jointly, a complex model of the environment we encounter, positing structure that has implications for all sorts of modal claims. In virtue of adopting certain models of how the world operates, we're committed to some things being possible, to some counterfactuals being true and false, and perhaps even to some necessities.

4. AGAINST BENACERRAF

Lewis offers us two arguments for thinking that our ordinary beliefs have modal implications. The first is that if none of our claims about the world have any modal implications, then realism is false. But realism is true. So, some of our claims about the world have some modal implications. The second argument is based on the conviction that beliefs have some practical value; they are the means by which agents navigate the world. If beliefs are to serve agents in this way, then they must provide some information about non-actual possibilities, since agential action involves selecting between possibilities. So, all beliefs provide some information about non-actual possibilities.

We can accept these arguments without going in for Lewis's semantics. And if we don't, then his arguments help us respond to the Benacerraf problem. Again, the Benacerraf problem demands some anti-luck condition on knowledge, which we seem to be able to satisfy when it comes to our knowledge of actual phenomena. When it comes to non-actual possibilities, however, we can't appeal to a causal relation to explain why it isn't an accident that our beliefs are accurate, and no other relation seems to be forthcoming. Are we stuck?

No. Based on the above, it's clear that the Benacerraf problem doesn't merely threaten our modal knowledge, it also threatens our knowledge of actuality. This is because our models of actuality are themselves modal – they sanction some modal claims and not others. So, either our hypotheses about the world aren't justified, in which case we don't even have ordinary knowledge about whether mugs are green, or they *are* justified, in which case we have mundane modal knowledge in spades. But surely the latter is true; to say anything else would be to go in for global skepticism. The Benacerraf problem is supposed to be limited to modality, and if it extends as far as I'm suggesting, then we can dismiss it however we're inclined to dismiss skepticism generally. My own view is that we should follow Lewis here: "Independent reality is not something to be proved but an original acknowledgment which all men make confronting the facts of life." We *begin* theorizing with a commitment to independent reality, and that commitment isn't shakable – it "is not something to be proved." And insofar as our commitment to independent reality is bound up with a commitment to non-actual possibilities, they aren't to be proved either.

5. OBJECTIONS

Let's consider a few objections. The first is that this Lewisian argument relies on an unwarranted closure principle. Plainly, if the argument assumes closure, then we don't face the Benacerraf problem. So, the argument can't presuppose it. But if the argument doesn't presuppose closure, then it isn't clear how we get the crucial premise: namely, that if we know

https://doi.org/10.1017/epi.2016.52 Published online by Cambridge University Press

various mundane truths, then we know at least some of the counterfactuals that they imply.

This objection misconstrues the reason to postulate a link between mundane truths and their associated counterfactuals. Granted, the truths imply the counterfactuals, but we don't know the counterfactuals simply on the basis of a *general* closure principle – e.g., one to the effect that if you know that *p*, then you know whatever you know *p* to entail. Rather, we know the associated counterfactuals based on a much more limited closure principle, and one that's hard to deny. According to this principle, if you know that an explanation is true, and if you know that that explanation implies certain counterfactuals, then you know the counterfactuals.

Suppose that we deny this principle. Then, you could know (1) and (2):

1. The cigarette's being flicked out the car window caused the forest fire.
2. If the cigarette's being flicked out the car window caused the forest fire, then, if the cigarette hadn't been flicked out the car window, there wouldn't have been a forest fire.

...and yet not know (3):

3. If the cigarette hadn't been flicked out the car window, there wouldn't have been a forest fire.

And that's absurd. To know that the cigarette caused the forest fire is to know what explains there being a forest fire, and to explain is to provide modal information: it's to articulate what depends on what. So, if you know the explanation, then you know the counterfactual – otherwise, you don't understand what you purport to know.

A second objection is that this Lewisian argument doesn't license *knowledge* of the relevant counterfactuals, but only a pragmatic propositional attitude, such as acceptance. After all, recall the passage from Lewis with which we began: "If independent factuality did not force itself upon us, *we should have to invent it in order to exist as being who think and wish to do*" (Lewis 1955: 237, emphasis added). The thought here seems to be that a certain view of the world is unavoidable for agents. But why should the requirements of agency determine what's true? So, while it may be appropriate for us to accept certain counterfactuals for practical purposes, we aren't entitled to claim that we know them.

This objection reveals what's worth saving from Lewis's verificationism. Recall: Lewis thinks that the meaning of ordinary claims are given by sets of counterfactuals that specify the conditions under which we could verify them. He's wrong about that. But it's plausible that *understanding* a claim like "The mug is green" partly involves knowing that certain counterfactuals are relevant to its truth or falsity. If you claim to understand the claim that the mug is green, but don't realize that, were you to look at it in normal lighting conditions, it would appear green, then we doubt your understanding.

After all, appreciating such connections allows searches for evidence to be purposeful rather than random, and in that respect such connections are relevant to investigating the world as agents. That is, we discover whether ordinary claims are true by making observations, and knowing which observations to make requires knowing certain counterfactuals of the form, "If hypothesis *h* is true, then (*ceteris paribus*) if *x* occurs, *y* will occur." These counterfactuals don't simply articulate the evidential link between observations and hypotheses: they also allow us to plan how to acquire the relevant evidence.

https://doi.org/10.1017/epi.2016.52 Published online by Cambridge University Press

Note, however, that we know many more such counterfactuals than we confirm. I might know, for example, that if a container is hollow, then, were I to tap its left side, I would hear an echoey sound. I might also know that if a container is hollow, then, were I to tap its right side, I would hear a similar echoey sound. I opt to tap the left side, I hear the sound, and by "actualizing" that counterfactual, I get some evidence for the hypothesis that the container is hollow. But the hypothesis, if true, supports the other counterfactual. What's more, it matters that it supports the other counterfactual, as that information matters for subsequent purposeful searches for evidence. If someone knows that I tapped the left side of the container and heard the echoey sound, and yet doubts my claim that the container is hollow, I can reasonably infer that he won't accept tapping the right side as evidence that I'm correct: I'll need to find a new way to defend my hypothesis (e.g., removing the lid and looking inside).

So, it isn't that the requirements of agency determine what's true: rather, it's that counterfactuals make our searches for evidence directed and purposeful, and are in that sense related to our agency. Moreover, these counterfactuals are the means by which we both confirm hypotheses and decide how to seek further evidence when it's needed. Hence, if we allow that the tap followed by the echoey sound counts as evidence for the hollowness hypothesis – as we certainly do – then we shouldn't posit a separate epistemic standard for the other counterfactuals associated with that hypothesis, as the objection does. They're confirmed as part of your model of the container and its properties.

This brings us to one final objection. As I argued above, Lewis's realism doesn't lead us to render different epistemic judgments about ordinary factual claims and their associated counterfactuals. However, the objection goes, if a certain view of the world is unavoidable for agents, then while we might not be epistemically blameworthy for believing both ordinary factual claims and their associated counterfactuals, we don't deserve any epistemic credit for them either. So while we might be entitled to *accept* a whole host of propositions – as Crispin Wright uses "entitlement" and "acceptance" – belief that they're true wouldn't be epistemically rational (as, e.g., Jenkins (2007) argues).

This objection doesn't pose a problem for my argument. My aim here is to reduce the Benacerraf Problem to the problem of skepticism generally. If this is the only objection I face, then the objection concedes that my project is successful: the worry about Lewis's realism is tantamount to the skeptic's challenge. So, I don't need to reply to this objection: I can simply concede that there is additional work to do, insisting only that the work doesn't specifically concern the Benacerraf Problem. That's a concession I'm happy to make.[7]

REFERENCES

Armstrong, D. M. 1989. *A Combinatorial Theory of Possibility*. New York, NY: Cambridge University Press.

Baldwin, T. 2007. 'C. I. Lewis: Pragmatism and Analysis.' In *The Analytic Turn: Analysis in Early Analytic Philosophy and Phenomenology*, pp. 178–95. New York, NY: Routledge.

Benacerraf, P. 1973. 'Mathematical Truth.' *Journal of Philosophy*, 70: 661–79.

7 Thanks to Eric Gilbertson, Bill Hart, Burkay Ozturk, and an anonymous referee for thoughtful comments on earlier drafts of this paper.

https://doi.org/10.1017/epi.2016.52 Published online by Cambridge University Press

Clark-Doane, J. Forthcoming. "What is the Benacerraf Problem?" In F. Pataut (ed.), *New Perspectives on the Philosophy of Paul Benacerraf: Truth, Objects, Infinity*. Dordrecht: Springer.
Jenkins, C. 2007. 'Entitlement and Rationality.' *Synthese*, 157: 25–45.
Lewis, C. I. 1912. 'Implication and the Algebra of Logic.' *Mind*, 21: 522–31.
—— 1946. *An Analysis of Knowledge and Valuation*, The Paul Carus Lectures. LaSalle: The Open Court Publishing Co.
—— 1954. 'A Comment (On "The Verification Theory of Meaning").' *The Philosophical Review*, 63: 193–6.
—— 1955. 'Realism or Phenomenalism?' *The Philosophical Review*, 64: 233–47.
—— 1970a. 'Logic and Pragmatism.' In *Collected Papers of Clarence Irving Lewis*, pp. 3–19. Stanford: Stanford University Press.
—— 1970b. 'Meaning and Action.' In *Collected Papers of Clarence Iriving Lewis*, pp. 87–91. Stanford: Stanford University Press.
O'Brien, L. 2007. *Self-Knowing Agents*. New York, NY: Oxford University Press.
O'Connor, T. 2008. *Theism and Ultimate Explanation: The Necessary Shape of Contingency*. Malden: Blackwell.
Peacocke, C. 1999. *Being Known*. New York, NY: Oxford University Press.

BOB FISCHER is an Assistant Professor of Philosophy at Texas State University. He is co-editor of two volumes – *The Moral Complexities of Eating Meat* (Oxford University Press, 2015) and *Modal Epistemology After Rationalism* (Springer, 2016) – and the sole editor of *College Ethics: A Reader on Moral Issues That Affect You* (Oxford University Press, 2016). He is also the author of *Modal Justification via Theories* (Springer, 2017), as well as a number of essays on modal epistemology, animal ethics, and moral psychology.

https://doi.org/10.1017/epi.2016.52 Published online by Cambridge University Press