Inferring as a Way of Knowing*

Nicholas Koziolek†

March 25, 2019

Abstract

Plausibly, an inference is an act of coming to believe something on the basis of something else you already believe. But what is it to come to believe something *on the basis of* something else? I propose a disjunctive answer: it is *either* for some beliefs to rationally cause another—where rational causation is understood as causation that is either actually or potentially productive of *knowledge*—*or* for some beliefs to "deviantly" cause another, but for the believer mistakenly to come thereby to believe that the former have rationally caused the latter. The result, I argue, is both a theoretically satisfying account of the act of inferring and a demonstration of the power of a knowledge-first approach to the philosophy of mind.

*This is a post-peer-review, pre-copyedit version of an article published in *Synthese*. The final authenticated version is available online at http://dx.doi.org/10.1007/s11229-017-1632-4

†Email: nkoziolek@gmail.com || Website: http://nkoziolek.net

1 Introduction

Epistemologists commonly assume that one way of coming to know something is by inferring it from something else, something you already know. They assume, in other words, that inferring is a way of knowing.[1] But when it is asked (usually by the philosopher of mind) what it *is* to infer something from something else, this connection with knowledge (so important to the epistemologist) is almost always ignored—and for apparently good reason: for the obvious answer is that to infer something from something else is to come to *believe* the former on the basis of the latter, which you already *believe*—no mention of knowledge required.[2]

Of course, to say that this answer is obvious is not to say that it is clear. And, indeed, it is widely recognized that no such answer will be complete until we can say what it *is* to come to believe something *on the basis of* something else you already believe—until, that is, we can explain the nature of the *connection* between the premise-beliefs and the conclusion-belief of an inference. But most philosophers assume that the story we need to tell here has nothing much to do with knowledge, or, therefore, with the fact (as, I will assume, it is) that inferring is a way of knowing. The standard assumption is thus that, with respect to inference, the epistemologist and the philosopher of mind have different tasks. The task of the philosopher of mind is to say what it is to infer something from something else. The task of the epistemologist, on the other hand, is to say when an inference—as characterized by the philosopher of mind—results in knowledge, as opposed to mere belief.

[1] Or, more precisely, of *coming* to know. On ways of knowing more generally, see Cassam 2007.

[2] In saying that inferring is the act of coming to believe something on the basis of something else you already believe, I may seem to be assuming that all reasoning must both begin from and conclude in belief—thus ruling out the possibility of (for example) reasoning under a mere hypothesis or supposition. But I mean to be doing no such thing. I thus wish to distinguish *inferring* (coming to believe something on the basis of something else you already believe) from *reasoning* more generally ("operating on contents," as John Broome puts it [2013: 231–232]). Inferring, on my way of speaking, must, indeed, both begin from and conclude in belief; but reasoning more generally need not.

This way of dividing our philosophical labor rests on an uncontroversial distinction between inference in general—the genus of inference—and one of its species, which I will call its *K-species*:

> **Genus:** An act of coming to *believe* something on the basis of something else you already *believe*.
>
> **K-species:** An act of coming to *know* something on the basis of something else you already *know*.

In these terms, the standard assumption, as I have called it, is that the task of the philosopher of mind is to investigate the whole genus, while the task of the epistemologist is to investigate only the K-species. But, while this distinction itself (between genus and K-species) is uncontroversial, it fails, on its own, to support the described division of labor. Specifically, in assuming that we can provide an account of the genus of inference without relying on the fact that inferring is a way of knowing—and so without employing the concept of knowledge—philosophers of inference assume that the genus of inference is *prior* to its K-species, in the sense that we must explain the K-species in terms of the genus. But this assumption is questionable. For it is possible to reverse the usual order of explanation, and to explain the genus of inference in terms of its K-species. More importantly, the resulting knowledge-first theory of inference[3] provides an elegant solution to one of the central problems facing the theorist of inference, namely, the problem of causal deviance. My aim in this essay, then, is to develop this alternative theory of inference in some detail. For reasons of space, I will not spend any time arguing against more standard approaches to the theory of inference, though I will discuss the relation between my approach and other extant approaches in §§2 and 3. Instead, I will try to motivate the view on its own terms, leaving a detailed discussion of its relative merits for another occasion.

[3] On knowledge-first epistemology and philosophy of mind more generally, see, e.g., Hossack 2007, McGlynn 2014, and Williamson 2000.

2 Good inferences first

The idea of explaining the genus of inference in terms of one of its species is not entirely novel. For some philosophers have suggested that fallacious inferences are parasitic on—and are thus to be explained in terms of—good inferences.[4] Good inferences, as understood by these philosophers, are those in which both (i) your conclusion does in fact follow from your premises and (ii) you draw your conclusion precisely *because* it follows from your premises. If we say that, in such a case, you *rationally* come to believe your conclusion, then we can say that, on this view, the genus of inference is to be explained in terms of what I will call its *R-species*:

> **R-species:** An act of *rationally* coming to believe something on the basis of something else you already believe.[5]

Since the idea here, again, is to explain the genus of inference in terms of its R-species, hence in terms of specifically good inferences, let's call this approach to inference *the good inferences-first approach*.

One important feature of the good inferences-first approach is that its account of the genus of inference will be disjunctive: it will provide different accounts of good and fallacious inferences, and understand the genus of inference as a disjunction of the two. This fact is especially clear if we consider Ralph Wedgwood's (2006) account of good inferences, on which the premise-beliefs of good inferences cause their conclusion-beliefs *in virtue of rationalizing* them. For the premise-beliefs of fallacious inferences will not cause their conclusion-beliefs in virtue of rationalizing them, because, by definition, the premise-beliefs of fallacious inferences do not rationalize their conclusion-beliefs. In other words, a fallacious inference is one in which your conclusion does not follow from your premises, and

[4] See, e.g., Grice 2001: 6–8 and Wedgwood 2006: 662–663.

[5] Note that, on this way of speaking, the fact that you've rationally come to believe your conclusion doesn't imply that your conclusion-belief itself is rational, since the rationality of the latter ultimately depends on the rationality of your premise-beliefs. The rationality in question here is thus the (diachronic) rationality of the transition from your premise-beliefs to your conclusion-belief.

so is one in which you cannot draw your conclusion *because* it follows from your premises. So, on this view, the connection between the premise-beliefs and the conclusion-belief of an inference will be *either* one of causation in virtue of rationalization *or* something else. The hope, though, is that we can explain this "something else," i.e., the nature of fallacious inferences, *in terms of* good inferences.

The *knowledge*-first account of inference to be developed below is similarly disjunctive. But it goes a step further than the mere good inferences-first account. In effect, it involves an explanation of the R-species of inference in terms of its K-species (which is a sub-species of the R-species). On this view, a good inference is one with the following feature: if you know your premises, then, by performing that inference, you come to *know* your conclusion. In other words, Wedgwood's causation in virtue of rationalization is understood, more fundamentally, as causation that is productive of knowledge.[6] The connection between the premise-beliefs and the conclusion-belief of an inference will thus be either *productive of knowledge* (if it is a good inference) or, again, something else (if it is fallacious). But what it is for an inference to be productive of knowledge will, in turn, be explained in terms of what I will call *epistemically successful inferences*: acts of coming to *know* something on the basis of something else you already know—hence acts in which you do in fact know your premises. What is distinctive of the knowledge-first account, then, is the treatment of epistemically successful inferences as explanatorily fundamental.

There are thus two tasks facing the knowledge-first theorist of inference: first, to explain good inferences in terms of epistemically successful ones; second, to explain fallacious inferences either in terms of good inferences or, more directly, in terms of epistemically successful ones (the latter, more direct, approach is the one adopted below). I undertake the first task in §§4–6 and the second in §§7 and 8. Before I turn to my positive account, though, I want to say

[6]One advantage of this account, which I won't discuss in detail here, is that it avoids treating the notion of rationalization (or, equivalently, the notion of something's following from something else, in a sense that allows for non-deductive inferences) as a philosophical primitive, instead explaining it in terms of knowledge. On this issue, cf. Hossack 2007: 26–27.

something more about the more traditional approach to the theory of inference, and in particular about what has come to be known as *the Taking Condition*.

3 The Taking Condition

A study of the recent literature on inference reveals a broad, if not complete, consensus that the connection between the premise-beliefs and the conclusion-belief of an inference must be mediated by some additional mental state or act, i.e., by some state or act other than the premise- and conclusion-beliefs.[7] As Paul Boghossian has put it, in stating what he calls *the Taking Condition*: "[i]nferring necessarily involves the thinker *taking* his premises to support his conclusion and drawing his conclusion *because* of that fact" (2014: 5, his emphases). In these terms, the broad consensus is that, when you infer something from something else, you come to believe the former, not just because you believe the latter, but also, in part, because you take the latter to support the former. More precisely, the consensus is that some mental state or act that is not among your premise-beliefs is among the *causes* of your conclusion-belief.[8]

Interestingly, something like the Taking Condition would seem to be required by any non-disjunctive theory of inference, i.e., by any theory that treats the genus of inference as explanatorily prior to its species.[9] For if the connection between the premise-beliefs

[7] For sympathetic discussions, see Audi 1993: 233–273, Boghossian 2014, Broome 2013: 229–231, Broome 2014, Chudnoff 2014, Fumerton 2004: 165, Hlobil 2014, Hlobil MS, Leite 2008, Marcus 2012: Chapter 1, Neta 2013, Neta MS, Pauer-Studer 2014, Stroud 1979, Thomson 1965, Tucker 2010, Tucker 2012: 333-334, and Valaris 2014. For critical discussions, see Lavin 2011, McHugh and Way 2016, Wedgwood 2006, and Wright 2014.

[8] It may be possible to provide an interpretation of the Taking Condition on which the 'because' in its second conjunct is taken in a non-causal sense. But since most defenders of the Taking Condition accept the causal interpretation, and since my own theory of inference will be a causal one, I will here simply set aside the possibility of a non-causal (e.g., merely normative) theory of inference.

[9] Such accounts needn't necessarily endorse the Taking Condition as formulated by Boghossian. For example, Broome (2013) says (in effect) that inferring necessarily involves the thinker's following a rule of inference, and McHugh and Way (MS)

and the conclusion-belief of an inference is the same in both good and fallacious inferences, then, because that connection holds in the case of a fallacious inference, it must be weaker than a rationalizing or knowledge-producing one. But since you can *take* your premises to support your conclusion whether or not they do, it turns out that, if the Taking Condition is true, then the difference between a good and a fallacious inference is simply (at least at first approximation) that, when you perform a good inference, you *correctly* take your premises to support your conclusion, while, when you perform a fallacious inference, you do so *mistakenly*.

The other side of this particular coin, though, is that disjunctive theories of inference don't require the Taking Condition (or anything like it)—at least not for an account of *good* inferences. The good inferences-first approach opens up the possibility that a good inference is, quite simply, an act in which some beliefs cause another in a certain way—say, in virtue of rationalizing it. On this view, the connection between the premise-beliefs and the conclusion-belief of a *good* inference is *not* mediated by any additional mental state or act. Instead, the premise-beliefs directly cause the conclusion-belief all on their own—they simply do so in a special way. What is distinctive of good inferences, then, is not that they have special causes (e.g., acts or states of taking it that your premises support your conclusion) but, instead, that they involve a distinctive variety

say that it necessarily involves the thinker's being guided by the aim of getting fitting attitudes. Either of these claims could plausibly be combined with a denial that inferring necessarily involves the thinker's taking her premises to support her conclusion. I'm less sure, however, that these theories don't endorse the broader consensus view that (I've claimed) lies behind the Taking Condition, namely, that, as I've put it in the text, "some mental state or act that is not among your premise-beliefs is among the *causes* of your conclusion-belief." The question here is whether "following a rule of inference" and "having the aim of getting fitting attitudes" are to be understood as mental states or acts that are both (i) additional to the premise- and conclusion-beliefs and (ii) among the causes of the conclusion-belief. If they're not, then the claim in the text is too strong. But it isn't obvious to me that there's any way of making these views work that doesn't treat the relevant addition to the premise- and conclusion-beliefs as an additional mental state or act that is among the causes of the conclusion-belief. (Thanks to an anonymous referee for *Synthese* for pressing me on this issue.)

of causation, namely, causation in virtue of rationalization—or, as I will later call it, *inferential rational causation*.

A pressing question for such disjunctive theories of inference, however, is whether they can be extended to cover fallacious inferences as well as good ones. As Ram Neta has said of Wedgwood's account: "if there is a way to generalize his account of reasoning so that it covers both good and bad reasoning, then it is not at all obvious how to do this" (2013: 392). In particular, it is not obvious how to do it without reintroducing the Taking Condition (or something like it). What could possibly distinguish fallacious inferences from mere associations, if it's not that, in the former, you mistakenly take your premises to support your conclusion (or something similar), while, in the latter, you make no such mistake? Admittedly, the defender of the disjunctive account has one fairly obvious reply: he might suggest that, since his account is already disjunctive, he can simply explain fallacious inferences in terms of "takings," *à la* the Taking Condition, while continuing to explain good inferences directly in terms of the special (rational) way in which their premise-beliefs cause their conclusion-beliefs. This move, however, seems *ad hoc*. Indeed, it is scarcely credible that fallacious inferences, *but not good inferences*, involve (for example) taking your premises to support your conclusion and drawing your conclusion because of that fact. If this is how fallacious inferences work, then, surely, it's how good inferences work, too.

The problem with that last suggestion in defense of the disjunctive account is that it treats the connection between the premise-beliefs and the conclusion-belief of a fallacious inference as mediated by some additional mental state or act, while, at the same time, treating the connection between the premise-beliefs and the conclusion-belief of a *good* inference as *un*mediated by any additional mental state or act. And that seems *ad hoc*. Let's assume that this view is indeed *ad hoc* and therefore implausible. Still, there's another option: perhaps the connection between the premise-beliefs and the conclusion-belief of an inference is *never* mediated by an additional mental state or act, *even* when the inference is fallacious.

This alternative approach involves a particular view of the nature of the additional, but now non-mediating, mental states or acts in question—which, for ease of reference, I'll simply call "takings." On the required view, takings are second-order beliefs about inferences (hence about the premise- and conclusion-beliefs of inferences and the causal connections between them), beliefs that are formed *in response to* the inferences themselves. This view of takings can be motivated by the plausible idea that, as Wilfrid Sellars has put it, "[e]ven our consciousness of what is going on in our own mind is a conceptual *response* which must be distinguished from that which evokes the response" (1967: 280, his emphasis). This idea is, admittedly, controversial.[10] But it makes a kind of sense. The basic idea is just that, even when it comes to our knowledge of our own minds, what is known is both causally and epistemically prior to our knowledge of it.[11] If this idea is right, then it may be that, while fallacious inferences do involve the thinker's taking her inference to be a good one, that taking is an *effect* of her drawing her conclusion, rather than the cause of it. And it may also be that, when *good*

[10] Sellars, who attributes the idea to Kant, himself continues: "Kant tends to limit this point to the introspection of sense impressions and other sensory states of the empirical self" (1967: 280), refusing to apply it to rational acts like inferences. Many contemporary Kantians seem to agree with Kant; see, e.g., Boyle 2015. (Sellars, for his part, is somewhat equivocal.) But their defenses of Kant's view on this issue aren't terribly compelling, resting as they do on the still-obscure Kantian notion of pure apperception. And there are, in any case, good reasons to question the relevant aspects of Kant's view, reasons forcibly developed by Longuenesse 2017. (With respect to Longuenesse's objections to Kant, cf. Sellars 1967: 280, note 10, which, in effect, anticipates a defense of Kant that Longuenesse argues is, for systematic philosophical reasons, extremely dubious.)

[11] I don't mean to be suggesting that defenders of the Taking Condition (or of non-disjunctive theories of inference more generally) necessarily reject this view. So far as I can see, their theories of inference are compatible both with this theory of self-knowledge and with its rejection. I introduce the idea here only to help motivate the view that takings *can* plausibly be understood as second-order states that, in the best case, constitute knowledge of (causally and epistemically prior) acts of inferring. Even if the defender of the Taking Condition accepts this view of self-knowledge, however, he'll need to argue that inferences *also* involve "taking your premises to support your conclusion" in some *other* sense—one in which the taking is a cause, rather than an effect, of the conclusion-belief. (Thanks to an anonymous referee for *Synthese* for pressing me on this issue.)

inferences involve the thinker's taking her inference to be a good one (which, for all I will say here, could be—though it needn't be—always), that taking will again be an effect, rather than a cause, of her drawing her conclusion.[12] With respect to *takings*—though not necessarily with respect to the connection between premise-beliefs and conclusion-belief—such an account of inference would thus be uniform (i.e., non-disjunctive), in the sense that it would give the same account of the role of takings in both good and fallacious inferences.

I will develop these ideas more fully in §§7 and 8. For now, it is important to note only that the view I will be developing there involves a rejection of the second conjunct of the Taking Condition (according to which takings are among the causes of the conclusion-beliefs of inferences), but not necessarily of the first (according to which inferences necessarily involve takings). It does even this, however, only on a particular interpretation of the Taking Condition, and, more specifically, a particular understanding of takings. Thus, on the usual understanding of takings, they are about the logical or evidential connections between propositions. On my understanding, they will instead be about the premise- and conclusion-beliefs of inferences, and the causal connection between them.[13] It is thus possible—and I will not try to settle this issue here—that the account I provide is compatible with the Taking Condition on other understandings of takings.[14] So, really, what I'll be doing in §§7 and 8 is providing a particular, and novel, interpretation of takings and then arguing that, *on that interpretation*, the Taking Condition—its second conjunct in particular—is false. The question of the relation

[12] Defenders of the Taking Condition might object, at this point, that the view I'm proposing will fail to acknowledge the thinker's *agency* in performing her inference. I don't think that this objection sticks, but the issue is a large one, discussion of which will thus need to be left for another occasion. For some relevant discussion, though, see, e.g., Setiya 2013 and Strawson 2003.

[13] The view adopted by Neta MS, however, is in at least partial agreement with the view I adopt here: he, too, takes takings to be about the inferences that, on his view, they help constitute. So this aspect of my account is not entirely novel, either.

[14] For example, I'm fairly confident that my view is compatible with the Taking Condition as understood by Hlobil MS, whose view is rather deflationary both of takings and of the sense in which they cause their subjects to draw their conclusions.

between the view I propose and other interpretations of the Taking Condition will be left for another occasion.

So much by way of background. I turn now to the task of developing my positive view. I will begin at a fairly high level of abstraction: not with inference in particular, but with epistemic acts more generally. Epistemic acts are acts that, at least sometimes, or at least potentially, result in knowledge. They include not just acts of inferring, but also, e.g., acts of perceiving and of receiving testimony. Since the account I will be developing is a knowledge-first account, however, I won't begin with just any epistemic acts. I'll begin, instead, with the successful ones: those that actually result in knowledge.

4 Successful epistemic acts

My account begins, again, with what I call *successful epistemic acts*: acts in which knowledge is either acquired or confirmed. As I've said, I will treat the concept of a successful epistemic act as explanatorily fundamental. What I take that to mean, though, is simply that the concept is to be explained through examples, i.e., through acquaintance with its objects, and not through definition. So I begin with examples. The point of these examples is simply to bring into view the explanatorily fundamental phenomenon of a successful epistemic act.

First Example: Suppose you're sitting at your desk, in front of your computer, answering emails. And suppose that, on the wall, behind and above your computer screen, sits a small spider. Assuming that the spider is in your field of vision, and that you're capable of recognizing spiders (of this kind) for what they are, you are, at this point, in a position to know that there's a spider on the wall. But, though you're in that position, you might fail to avail yourself of it: you might, for a time, fail to *notice*, to *register*, the spider—because, for example, you're so absorbed in the task of answering those emails. But you might then hit "send," let your eyes stray from the screen, and, finally, notice the spider. At that point, you

avail yourself of the position you were in all along: you come to *know* that there's a spider on the wall.

Second Example: Suppose that, on Tuesday, your friend Sophie mentions in passing that she's flying to New York on Friday to give a talk at NYU. And suppose that, on Thursday, your mutual friend Martin invites you to a party on Friday evening. Suppose, finally, that, viewing the invitation (a mass email), you notice that Sophie was invited, too. At this point, you're in a position to know that Sophie won't be at the party. But, though you're in that position, you might have forgotten that she'll be out of town. And so you might find yourself hoping that you'll see her at the party. And it might not be until Friday evening, at the party—perhaps just before you ask Martin if he knows whether Sophie is planning to come—that you remember: Sophie is in New York—which means that she won't be at the party. At that moment, you avail yourself of the position you were in all along: you come to *know* that Sophie won't be at the party.

In the first of these examples, you come to know that there's a spider on the wall on the basis of your visual perception of the spider on the wall. In the second, you come to know that Sophie won't be coming to the party on the basis of your knowledge that she's in New York. These are both successful epistemic acts: acts in which knowledge was either acquired or confirmed. In fact, as described, they are both acts in which knowledge was acquired, not merely confirmed.[15] But it's possible for such acts to confirm knowledge instead: you might, for example, come to know something on the basis of testimony, and later confirm it on the basis of perception. The later perceptual act, like the earlier testimonial act, is a successful epistemic act, for it is an act in which knowledge is confirmed.

[15] They are also both *self-conscious* epistemic acts, hence acts that involve taking it that your newly acquired belief is epistemically well-grounded. I say more about self-conscious epistemic acts in §§7 and 8.

5 Theoretical rational causation

The key idea of the approach to inference that I'll be developing in §§6–8 is that, using the notion of an epistemically successful *inference* (the kind of successful epistemic act on display in the second example from §4), we can define a special kind of causation, which I will call *inferential rational causation*. This strategy, however, is simply an instance of a more general explanatory strategy, one that can be employed in connection with any epistemic act (including, for example, the act of coming to know something on the basis of perception). At this higher level of abstraction, then, the idea is to use the more generic notion of a successful epistemic act to define a special kind of causation of a more generic kind, namely, *theoretical rational causation*. Inferential rational causation is thus a species of theoretical rational causation. But, importantly, the idea here is not to define inferential rational causation in terms of theoretical rational causation by adding an appropriate differentia. Rather, inferential rational causation is to be understood directly in terms of epistemically successful inferences.[16] (The same will be true for other epistemic acts—including acts of self-consciousness, which will be discussed in §7.) I think it will be helpful, however, to begin by discussing these issues at the higher level of abstraction. So my aim in this section is to explain the more generic notion of theoretical rational causation. In doing so, however, I will be using perception as an example. So there is, in fact, a sense in which what I'll be doing here is sketching an account of *perceptual* rational causation. Still, the discussion should help to reveal the generic shape of the general strategy I'll be employing in my account of inference, which I'll begin developing in the next section. (I plan to develop the account of perception itself more fully in future work.)

The definition of theoretical rational caustion in terms of successful epistemic acts is given in three stages. First, from examples of successful epistemic acts (like those given in §4), we can abstract out a notion of *causation that's actually productive of knowledge*. Sec-

[16] Inferential rational causation is thus what Ford (2011) calls a *categorial species* of theoretical rational causation.

ond, in terms of causation that's *actually* productive of knowledge we can define another kind of causation, causation that's *potentially* productive of knowledge. Finally, we can define theoretical rational causation as the disjunction of the two kinds of causation thus defined, i.e., as causation that's either actually or potentially productive of knowledge.

My aim in this section is just to explain these ideas in a bit more detail. I want to stress here at the outset, though, that, while some of what I say in this section may sound like an attempt to explain or define theoretical rational causation otherwise than in terms of successful epistemic acts (specifically: in terms of knowledge and causation), that is not my intent. What follows is merely an attempt to draw out a bit more of what I take to be contained in the notion of a *successful* epistemic act, and then to use the resulting materials to define what we might call (diachronically) *rational* epistemic acts: epistemic acts in which knowledge is neither acquired nor confirmed, but only because the candidate cause of knowledge is not of the right type to produce knowledge—so that there is a mismatch, of sorts, between the cause and the effect. In other words, my aim here is to explain causation that's merely potentially productive of knowledge in terms of causation that's actually productive of knowledge.

For causation to be *actually* productive of knowledge, in the intended sense, is for it to *explain* knowledge, and, moreover, to explain it *as* knowledge. For causation to be merely *potentially* productive of knowledge, on the other hand, is for it to be merely *such as* to explain knowledge as knowledge.

For causation to explain knowledge *as* knowledge is for it to explain that knowledge in such a way as to reveal *how* it is that the subject knows what she knows. The latter requirement is important, because there are explanations of knowledge that *don't* reveal how it is that the subject knows what she knows. Thus, suppose that Quentin believes that all cats are black (perhaps because, by some strange coincidence, every cat he has ever seen or heard about, whether real or depicted, has been black). And suppose that his older sister, who has been planning to adopt a cat, learns of his

surprising belief, and so decides to adopt a black cat. Suppose, finally, that, when Quentin visits his sister and sees the cat, he (quite reasonably) comes to believe that her cat is black. Indeed, let's suppose that he comes to *know* that her cat is black—for his sister, let's assume, is generally quite trustworthy, and so there is no reason for him to suspect that she may be trying to deceive him. The result is that Quentin now *knows* that his sister's cat is black. Further, his belief that all cats are black is part of the causal explanation of his knowledge. So we have the following—true, though partial—explanation of Quentin's knowledge: Quentin knows that his sister's cat is black because he believes that all cats are black.

Now, although this is an explanation of Quentin's knowledge that his sister's cat is black, it is not an explanation of it *as* knowledge. It does not explain his knowledge *as* knowledge because it does not reveal *how* he knows. The explanation that reveals how he knows that his sister's cat is black is, instead, that he saw that it was black. The latter, then, *is* an explanation of his knowledge as knowledge. For, again, an explanation of knowledge as knowledge tells us how it is that the subject of the explanation knows what she knows.

This example—Quentin's coming to know that his sister's cat is black by seeing that it is black—thus serves to identify a special form of explanation: explanation of knowledge as knowledge. Explanations of this form are, I assume, efficient-causal explanations. For to say that someone came to know something in some particular way is to say that something brought something else about. Quentin, for example, saw that his sister's cat was black, and it was through this perceptual act that he came to know that his sister's cat was black; his perceptual act thus brought about his knowledge. Where we have such an explanation, then, I will say that the causation cited in the explanation is actually productive of knowledge. Such causation, importantly, is involved in all and only successful epistemic acts.

For causation to be merely *potentially* productive of knowledge, on the other hand, is for it to be merely *such as* to explain knowledge as knowledge. To see what this means, consider a variation on the example involving Quentin. This time, suppose that it is Quentin's

aunt who learns of his surprising belief that all cats are black. And suppose that his aunt, being a bit of a prankster, decides to dye her white cat black whenever Quentin visits. In this case, when Quentin first visits and sees the cat, he will come to believe, mistakenly but reasonably, on the basis of his perception of the cat, that his aunt's cat is black. The crucial point is that, in this case, Quentin has not come to know, but merely to believe, that his aunt's cat is black—for the cat is, in fact, white. But the reason he fails to acquire knowledge here is simply that the appearance of the cat is misleading as to its true color. In particular, given how the cat appears to Quentin, it is quite reasonable for him to believe, on the basis of his perceptual experience, that it is black. In fact, had things been as they appeared to Quentin to be (and, perhaps, had Quentin's aunt been less prone to playing tricks on her nephew), he *would* have come to know that the cat was black. In the present case, then, (the case in which Quentin is deceived) the *causal connection* between Quentin's perceptual experience and his subsequent belief is, in itself, of the right kind to produce knowledge. It is just that the *cause* of his belief—his perceptual experience—is not. In such a case, I will say, the causation in question is *potentially* productive of knowledge. For it is such as to explain knowledge, and to explain it as knowledge, in the sense that it will explain knowledge as knowledge when the cause is also of an appropriate type. In the perceptual case, this means: if the perceptual state is non-accidentally veridical (and any relevant "background beliefs" are knowledgeable). In the inferential case, it means: if the premise-beliefs are knowledgeable. Importantly, causation that is merely potentially productive of knowledge is involved in all and only unsuccessful but still (diachronically) rational epistemic acts.

In the context of the problem of causal deviance (the problem of distinguishing inferences from mere associations), it's important to see that the notion of causation that's productive of *knowledge* is stronger than the notion of causation that's productive of *belief*. In fact, my suggestion here is that the former picks out a special *kind* of causation (i.e., a special kind of causal *relation*), and that the latter does not. One way of putting the point is to say that causation that

is productive of knowledge necessarily rationalizes what it causes. Causation that is merely productive of belief, on the other hand, may or may not rationalize what it causes. Sometimes, for example, you believe something because you're biased in some way, or because you desperately want what you believe to be true. In these cases, your bias or your desire help (at least) to bring about your belief. But they don't rationalize it (or if, by chance, they do, they do so only accidentally). And so they can't (or at least don't) bring it about in virtue of rationalizing it. The result is that we cannot define theoretical rational causation in terms of belief. We must instead define it in terms of knowledge.[17]

6 Good inferences

Theoretical rational causation, again, is causation that is either actually or potentially productive of knowledge. More precisely, it is the kind of causation involved in successful epistemic acts, acts in which knowledge is either acquired or confirmed. Correspondingly, then, *inferential* rational causation is the kind of causation involved in epistemically successful *inferences*, i.e., acts of coming to know something on the basis of something else you already know. This kind of causation is thus, in effect, causation that is either actually or potentially productive of *inferential* knowledge. It is actually productive of inferential knowledge when the premise-beliefs are known. It is merely potentially productive of inferential knowledge when the subject fails to know at least one of her premises.

The account of inferential rational causation is exactly along the lines of the more generic account of theoretical rational causation. We begin here, however, with examples of epistemically successful *inferences* in particular. From these examples (like the second

[17] You might object here that we could define it in terms of *justified* belief. I can't get into this issue here, but I don't think that this definition will work any better than the one in terms of belief, unless justification is itself defined in terms of knowledge. And that definition would effectively vindicate the account provided here. (Thanks to both Rebeka Ferreira and Reza Hadisi for pressing me to think more about this objection.)

example given in §4), we can abstract out a notion of causation that's productive (i.e., actually productive) of specifically *inferential* knowledge. We can then use this notion to define causation that's merely potentially productive of inferential knowledge. And, finally, we can define inferential rational causation as the disjunction of the two kinds of causation thus defined, i.e., as causation that's either actually or potentially productive of inferential knowledge. And we can then define good inferences in terms of inferential rational causation.[18]

What we have here is, in effect, a definition of what I have called the R-species of inference (i.e., the act of rationally coming to believe your conclusion on the basis of your premises) in terms of its K-species (the act of coming to know something on the basis of something else you already know). But the definition is in two steps. In the first step, we begin with the K-species, i.e., with epistemically successful inferences, and we abstract out the notion of inferential rational causation, i.e., causation that is either actually or potentially productive of inferential knowledge. We then define merely good inferences—those that do not issue in knowledge (but merely in belief), but only because the subject fails to know at least one of the premises—as acts in which some beliefs cause another in a way that is merely potentially productive of inferential knowledge. Good inferences in general are thus defined disjunctively: as acts in which some beliefs cause another either in a way that is actually productive of inferential knowledge or in a way that is merely potentially productive of inferential knowledge.

Given that inferential rational causation, so defined, will be actually productive of inferential knowledge just in case the subject of the inference knows the premises, we can also say that inferential

[18] You might object here that an account of inference in terms of inferential knowledge is going to be viciously circular. But my aim here, again, isn't to explain epistemically successful inferences. Those, I'm assuming, are primitive. The aim, rather, is to explain both merely good (and so generically good) and fallacious inferences in terms of epistemically successful inferences. Any appearance of circularity thus arises from the (false) assumption that my aim is to explain epistemically successful inferences (and so epistemic knowledge) in terms of inferential rational causation, rather than the other way around. (I discuss the methodology at work here in more detail in Koziolek 2015: chapter 2.)

rational causation is causation that is productive of knowledge *when the cause is knowledge*. It remains only to show that the notion of inferential rational causation, so understood, gives the right results.

On the account of good inferences just given, to infer $\langle p \rangle$ from $\langle q \rangle$ and $\langle r \rangle$, and for your inference to be good, is for your beliefs that q and that r together to inferentially rationally cause you to believe that p. Here, though, there are two cases. In the first case, you *know* both that q and that r. Given that you know both that q and that r, and given that inferential rational causation is causation that's productive of knowledge when the cause is knowledge, it follows that you know that p. But it could follow that you know that p, given that you know both that q and that r, only if you *inferred* $\langle p \rangle$ from $\langle q \rangle$ and $\langle r \rangle$. Thus far, then, the account gives the right results.

In the second case, you do not know, but *merely* believe both that q and that r (that is, you believe both, but you know at most one). So we cannot conclude that you know that p. Nonetheless, given that inferential rational causation is causation that's productive of knowledge when the cause is knowledge, we *can* conclude that, *had* you known both that q and that r, then you *would* have known that p. Again, though, your beliefs that q and that r could not be such as to transmit their epistemic standing to your belief that p unless you had *inferred* $\langle p \rangle$ from $\langle q \rangle$ and $\langle r \rangle$. So, again, the account gives the right results.

Such is my account of good inferences. There are, to be clear, two kinds of good inferences: (i) those that issue in knowledge; and (ii) those that issue merely in belief, but *only* because the subject fails to know the premises. Both of these, however, are acts in which the premise-beliefs inferentially rationally cause the conclusion-belief. What remains, then, is only to explain the nature of fallacious inferences.

7 Self-conscious inferences

I've said that inferential rational causation is causation that's productive of knowledge when the cause is knowledge. To say this, however, is to imply that *fallacious inferences involve deviant, i.e.,*

non-rational, causation. For, in a fallacious inference, the connection between the premise-beliefs and the conclusion-belief is not even potentially productive of knowledge: you could not come to know the conclusion of a fallacious inference by inferring it from the premises, even if you knew the premises.

But if fallacious inferences involve deviant causation, and if (as I have claimed) the connection between beliefs that is distinctive of inference is the relation of inferential rational causation (which is non-deviant), how can fallacious inferences count as inferences? The short answer is that, although fallacious inferences do not *involve* inferential rational causation, they are nonetheless explained in terms of it, and in a way that reveals their nature as inferences—or, at least, shows why it makes sense to *call* them inferences. The explanation, however, requires that we first expand the account of inference provided above.

What we need to add to the above account of good inferences is an account of what it is for an inference to be *self-conscious*. It is, of course, controversial just what it is for a mental state or act in general to be self-conscious. Here, however, I will simply assume the account that I think is correct. On this account, for a state to be self-conscious is for you to know that you're in it, where you know that you're in it in a particular, and special, way, namely, by being in it.[19] Similarly, for an act to be self-conscious is for its performance to involve coming to know that you've performed it, where you come to know that you've performed it precisely *by* performing it.

Can you really (come to) know that you're in a state merely by being in it? Well, suppose you have a headache, and that you know that you do.[20] The question is: *how* do you know that you have a headache? This isn't a question we often press. But that

[19] For the view that self-consciousness is the capacity to know that you're in a state by being in it (and thus that self-consciousness is a special, *sui generis*, way of knowing), see Rödl 2007. For a brief but helpful defense of the possibility of self-conscious knowledge of this sort, see Valaris 2014: §5.4. For an earlier treatment of similar issues, which is also in line with the view taken here, see Burge 1996.

[20] Some philosophers, following Wittgenstein, have denied that you can know that you're in pain. But I don't find Wittgenstein's, or his followers', argument(s) for this view particularly plausible. So I assume that, if you know something, then there is a good answer to the question how you know it.

might be because the answer is (usually) obvious. After all, we also rarely press the question how someone knows something they can plainly see to be the case. If you're looking at your neighbor's new car, and you tell me that it's a lovely shade of blue, I'm unlikely to ask you how you know that your neighbor's car is blue. But that's not because there's no good answer. It's because I know perfectly well how you know. So it might be that we rarely ask people how they know that they're in pain only because the answer is usually obvious: if a person is in pain, then she's in a position to know that she is; and so, if she does know, her knowing is no surprise. (Compare: If someone can see the blue car, then, typically, she's in a position to know that the car is blue; and so, if she does know that it's blue, that's no surprise.) There is of course a step from here to the view that she knows that she's in pain *by being in pain*. But what these considerations suggest is that to be self-conscious is simply to be such that being in a mental state sometimes puts you in a position to know that you're in that state. More precisely, it puts you in a position to know that you're in that state in a special way—not by perception, or by inference, or by testimony, but via self-consciousness, i.e., via a special, and *sui generis*, way of knowing your own mental states. And since what enables your knowledge is just (i) your being self-conscious and (ii) your being in the state in question, it seems reasonable to say that, when you know, via self-consciousness, that you're in a given state, you know that you're in it *by being in it*. For, after all, being in the state is what—in addition to your being self-conscious—enables you to know that you're in it. And so we can say that self-consciousness is the capacity to know that you're in a state by being in it (and, by extension, to know that you've performed an act by performing it).

It will be important, in what follows, that exercises of self-consciousness—attempts, as it were, to come to know that you're in a state by being in it—can be defective. Thus, for example, it's possible (I would argue) to mistake an itch for a pain (and vice-versa). In such a case, however, you still employ the relevant way of knowing, i.e., self-consciousness. This is revealed in the fact that, if you claim, mistakenly, to have felt a momentary pain (when what

you felt was in fact an itch), and someone asks you how you know it was a pain, rather than an itch, your answer will be that you know it was a pain because of how it *felt*—i.e., in effect, because the state you were in was one of *pain*. You'll be wrong, of course, because what you felt was, by assumption, an itch and not a pain. But the false belief you form will nonetheless be formed via an exercise—a *defective* exercise—of self-consciousness, i.e., via an exercise of the capacity to know that you're in a state by being in it. So we can say that, in this case, you believe that you felt a momentary pain because you felt a momentary *itch*, which you mistook for a pain. Here, the itch directly causes your belief that you felt a momentary pain. And it does so in a special way, the way distinctive of self-consciousness. The relation between your itch and your belief that you felt a momentary pain is thus not (for example) a perceptual one or an inferential one; it is, rather, a distinctively self-conscious one. In the terms I've been employing in this paper: the itch *self-consciously rationally* causes your belief that you felt a momentary pain. (Note here that, *had* you in fact felt a momentary pain, you *would* have come to know that you had felt a momentary pain. So the causation in question is potentially productive of knowledge.)

The same kind of mistake can occur with respect to your attempts to come to know that you know something. For example, you might take yourself to know that it's raining, on the basis of the sounds you're currently hearing (i.e., you take yourself to know that it's raining by hearing that it is). If you're wrong, you'll still *think* you know. So you'll *believe* that you know that it's raining (by hearing that it is), and you'll have formed that belief via an exercise of self-consciousness. And although you won't know that you *know* that it's raining, via this exercise of self-consciousness, you *will* know, via that very same exercise of self-consciousness, that you *believe* that it's raining. Thus, just as you can mistake an itch for a pain, so you can mistake mere belief for knowledge. (You can also, in a way, mistake knowledge for belief. There's a sense in which this happens whenever you suspend judgment in the face of what is in fact a misleading defeater. Of course, you won't actually mistake knowledge for belief in that sort of case, because, at least typically,

you won't actually come to believe the thing in question, and so you won't take anything for belief. But you will nonetheless mistake your good epistemic position for a bad one.)[21]

As I indicated above, self-consciousness, like any way of knowing, involves a species of theoretical rational causation, i.e., a species of causation that's productive of knowledge. Self-conscious rational causation, as we can call it, comes into view in successful exercises of self-consciousness, i.e., acts of coming to know that you're in a state (that you've performed an act) by being in it (by performing it): for example, an act of coming to know that you're in pain by being in pain. As before, we can use such acts to define causation that's actually productive of self-conscious* knowledge (i.e., the kind of knowledge acquired in such acts).[22] And we can then use the latter to define causation that's merely potentially productive of self-conscious* knowledge. The result is that we can understand an act of coming to believe that you're in pain, where that belief is in fact caused by an itch (so that your belief is false, and so not knowledgeable), as involving self-conscious rational causation: in effect, the belief that you're in pain is formed via an exercise of self-consciousness just in case, had that belief been self-consciously rationally caused *by your being in pain* (rather than by your having an itch), then you would thereby have come to *know* that you were in pain.

So self-consciousness is the capacity to know that you're in a state (that you've performed an act) by being in it (by performing it). A successful exercise of this capacity issues in knowledge of the state (act) in question. An unsuccessful exercise, however, will issue in mere belief about the state (act) in question—though that belief will be self-consciously rationally caused. But, importantly, some

[21] On the issues raised in this parenthetical, cf. Sosa 2009: chapter 7 on the relation between animal knowledge and reflective knowledge.

[22] The expression "self-conscious knowledge" most naturally refers to knowledge that you know yourself to have, where you know yourself to have it via an exercise of self-consciousness. Here, however, I mean the expression to refer instead to the second-order knowledge that you have, via self-consciousness, of some *other* mental state. Since I can't think of a better form of expression, I will indicate the latter use of the expression by appending an asterisk (*) to the word "self-conscious."

exercises of self-consciousness will be *partially* successful. This is what happens when you take yourself to know something you don't really know: you'll be wrong in taking yourself to know, but right in taking yourself to believe; and so you'll merely believe that you know, but you'll nonetheless *know* that you believe. The possibility of partially successful exercises of self-consciousness will be crucial in the accounts of both (merely) good self-conscious inferences and, finally, fallacious inferences.

We can now apply this account of self-consciousness to inference. As before, we begin with the epistemically successful case. An epistemically successful self-conscious inference is an act in which (i) you come to know something on the basis of something else you already know, (ii) you come to know that you've come to know the former on the basis of the latter, and (iii) you do so precisely *by* coming to know the former on the basis of the latter. Or, to put it another way: (i) your knowledge of your premises inferentially rationally causes your knowledge of your conclusion and (ii) that rational-causal act itself self-consciously rationally causes your knowledge that your knowledge of your premises has inferentially rationally caused your knowledge of your conclusion. The causation in (ii), again, is *rational* because it's causation that's productive of knowledge: it's productive of the second-order knowledge that you've come to know your conclusion by inferring it from your premises.

Merely good inferences (inferences that are valid or strong, but not sound) are similar, but with two differences. First, because you don't know your premises, you don't come to know your conclusion, but merely to believe it. Second, because you don't come to know your conclusion, you don't come to *know* that you've come to know your conclusion. Instead, you come merely to *believe* that you've come to know your conclusion. More precisely, (i) you come to believe something on the basis of something else you already believe, (ii) you come to believe (mistakenly) that you've come to know the former on the basis of the latter (which you mistakenly think you know), and (iii) you do so by coming to believe the former on the

basis of the latter.[23] Or, to put it the other way: (i) your premise-beliefs inferentially rationally cause your conclusion-belief and (ii) that rational-causal act itself self-consciously rationally causes your belief that your knowledge of your premises has inferentially rationally caused your knowledge of your conclusion. Here the causation in (ii) is rational because it's causation that's *potentially* productive of knowledge: had you known your premises, you would have come to know—via self-consciousness—that you had come to know your conclusion by inferring it from your premises. You also do in fact come to *know* that your premise-beliefs have inferentially rationally caused your conclusion-belief, and hence that you believe your conclusion (on the basis of your premises). So the act of self-consciousness here is, while defective, still *partially* successful.[24]

This account of good self-conscious inferences reflects the idea of Sellars's that I mentioned towards the end of §3: "[e]ven our consciousness of what is going on in our own mind is a conceptual *response* which must be distinguished from that which evokes the re-

[23] Condition (ii) here is closely akin to Broome's (2013, 2014) requirement that your inference seem right to you. Unlike Broome, however, I explain this seeming in epistemic terms, i.e., ultimately, in terms of (inferential) knowledge.

[24] There may also be another kind of good inference: one in which you neither know your premises nor believe yourself to know them (for example, because, although you have, and take yourself to have, good evidence that your premise-beliefs are true, you don't take that evidence to be good enough to give you knowledge), but merely believe them (perhaps with justification) and believe (or even know) yourself merely to believe them (with justification). In such a case, when you infer self-consciously, you will come to believe only that you've (rationally or justifiably) come to *believe* your conclusion by inferring it from your premises. Whether there are such inferences depends on whether it is possible to self-consciously believe something without taking yourself to know it. If such a thing is possible, however, my account will explain the possibility of the resulting sort of inference: in inferring self-consciously while taking yourself not to know your premises, you won't come to believe that you've come to know your conclusion by inferring it from your premises, precisely because you don't take yourself to know your premises. You'll instead come to believe only that you've come to believe your conclusion (with whatever level of justification you take to be appropriate given the level of justification you take yourself to have for your premises). If such a thing isn't possible, on the other hand, my account will equally explain the impossibility of the relevant sort of inference. So, while I assume in the text that you can't self-consciously believe something without taking yourself to know it, nothing hangs on this assumption.

sponse" (1967: 280, his emphasis). Here, what evokes the response is an act of coming to believe something on the basis of something else you already believe. The response is itself an attempt—successful in one kind of case, unsuccessful in another—to grasp the nature of that act. Thus, this account of self-conscious inference is motivated by two plausible ideas: first, that, in the best case, coming to know something by inferring it from something else you already know involves coming to know that that's what you've done; second, that the latter knowledge is a response to the former inference, in the sense that the inference is both causally and epistemically prior to the knowledge of it acquired via an exercise of self-consciousness, i.e., of the capacity to know that you've performed an act by performing it.

8 Fallacious inferences

So how does this account of self-conscious inference help us to explain fallacious inferences? In particular, how does it help us to explain how a deviant-causal act can count as an inference?

Well, if, via an exercise of self-consciousness, you can come to know, and thus to believe, that your premise-beliefs have inferentially rationally caused your conclusion-belief when, in fact, they have, then you can also, in the same way, come to believe that your premise-beliefs have inferentially rationally caused your conclusion-belief when, as a matter of fact, they haven't. In particular, you can come to believe that your premise-beliefs have inferentially rationally caused your conclusion-belief when the former have only deviantly caused the latter—when, as I will say, you have merely associated. Precisely this, I submit, is what happens when you perform a fallacious inference. Thus, to perform a fallacious inference is for you (i) to associate and (ii) to come to believe that your premise-beliefs have inferentially rationally caused your conclusion-belief, where (iii) this second-order belief is self-consciously rationally caused by the act of association itself.[25] In other words: to perform

[25]Note that this formulation is compatible with the possibility of the kind of self-conscious inference discussed in note 24 above, in which, because you don't

a fallacious inference is to come to believe that your premise-beliefs have inferentially rationally caused your conclusion-belief when, and precisely because, you have merely associated, where the 'because' here expresses the kind of rational causation that is distinctive of self-consciousness.

This last requirement—that the second-order belief be self-consciously rationally caused—is important. Without it, the following event would count as an inference: (i) your beliefs that q and that r deviantly cause you to believe that p, and (ii) because they've done so, a mad scientist causes you to believe that you've come to know that p by inferring it from $\langle q \rangle$ and $\langle r \rangle$.[26] In this case, the *result* is in some sense the same as in the case of a fallacious inference: your beliefs that q and that r deviantly caused you to believe that p; you now believe, falsely, that you have come to know that p by inferring it from $\langle q \rangle$ and $\langle r \rangle$; and the former act of association is part of the causal explanation of the latter belief. But, crucially, the second-order belief you end up with (namely, that you've come to know that p by inferring it from $\langle q \rangle$ and $\langle r \rangle$) was not formed via self-consciousness; rather, it was caused by the mad scientist. Thus, because the second-order beliefs involved in fallacious inferences need to be self-consciously rationally caused by the acts of association they're about, this sort of case is ruled out.

As in the case of good self-conscious inferences, fallacious inferences involve a second-order act that is a *response* to something going on in the mind. In this case, however, what evokes the response is not an act in which some of your beliefs inferentially rationally cause another. It is, instead, an act in which some of your beliefs *deviantly* cause another. But the response evoked is similar: it is a belief that you've come to know something by inferring it

take yourself to know your premises, you don't come to believe that you *know* your conclusion, but instead come to believe merely that you (rationally or justifiably) *believe* it.

[26] I owe the example to conversation with Errol Lord and Jonathan Way. I'm grateful to them for helping me to see that the second-order beliefs involved in fallacious inferences need to be grounded in (and not merely caused by) the acts of association that cause them—which means, on the present view, that those beliefs need to be formed via exercises of self-consciousness.

from something else you already knew.[27] The difference is that the response is now doubly mistaken. It is mistaken not only in that it represents you as knowing your conclusion, but also (and more fundamentally) in that it represents your premise-beliefs as inferentially rationally causing your conclusion-belief. In other words: a merely good self-conscious inference misrepresents its subject's premise- and conclusion-beliefs, whereas a fallacious inference misrepresents the causal connection between them. Even in the case of a fallacious inference, however, your second-order belief will *correctly* represent your conclusion-belief as having been *caused* by your premise-beliefs. In fact, it will still amount to *knowledge* that your premise-beliefs caused your conclusion-belief. It simply involves a misidentification of the *kind* of causation involved.[28]

9 Conclusion

On the view I have proposed, an epistemically successful self-conscious inference is, in effect, a joint actualization of two distinct capacities: (i) the capacity to infer (i.e., to come to know something on the basis of something else you already know) and (ii) the capacity to come to know that you've performed an act by performing it (i.e., self-consciousness). A merely good inference involves the joint actualization of these same two capacities, but the actualization of each of them is now defective: the actualization of the capacity to infer is defective because you don't know your premises, and the actualization of self-consciousness is defective because, believing, mistakenly, that you know your premises, you come to believe, again

[27] Or, again, that you've come justifiably to believe something by inferring it from something else you already justifiably believed. Cf. note 24 above.

[28] The second-order beliefs involved in the kinds of inferences discussed in note 24 will also be doubly mistaken, for they will represent the subject as justifiably believing the conclusion (which is false, given that the premises don't support the conclusion) and they will represent the subject's premise-beliefs as having inferentially rationally caused her conclusion-belief (which is also false). But they will still correctly (indeed, knowledgeably) represent her conclusion-belief as having been caused by her premise-beliefs. So some self-conscious* knowledge is acquired even here.

mistakenly, that you've come to know your conclusion by inferring it from your premises.

Paradoxically, however, a fallacious inference does not involve an actualization of the capacity to infer—not even a defective one. Instead, it involves only a defective actualization of self-consciousness. What happens is that an act that is not an inference at all, but instead a mere association, self-consciously rationally causes you to come to believe that you've performed an epistemically successful inference. In other words, self-consciousness mistakes a mere association for an epistemically successful inference.

There is a sense, then, in which a fallacious inference is not an inference at all. It is instead something else—a mere association—that its subject mistakes for an inference, via a defective exercise of self-consciousness.[29] But it obviously makes a kind of sense to count fallacious inferences, so described, as genuine inferences—if only by a sort of courtesy. For the subject of a fallacious inference, so described, will necessarily *think* she's performed an inference (indeed, an epistemically successful one). A fallacious inference should count as an inference, then, for essentially the same reason that a bad philosopher—one who defends mistaken philosophical views—should count as a philosopher and not a sophist: in both of these cases, the subject's conception of herself, and of what she's doing, makes all the difference.*

[29] It's a consequence of this view that there are no non-self-conscious fallacious inferences—and thus that non-rational animals, animals that are not self-conscious, cannot commit fallacies—even if they *can* perform good inferences (say, because they can come to know and believe things, in some sense of 'know' and 'believe', on the basis of other things they already know or believe, in the same sense of 'know' or 'believe'). I grant, of course, that we often use the word 'inference' to pick out transitions between beliefs that, on the present view, aren't really inferences at all, but are mere associations. So this aspect of my view is (I think mildly) revisionist. The revision does, I think, raise some interesting and difficult questions, but the issues are complicated enough that I think it best to leave discussion of them for another occasion. I do think, however, that the revision is defensible.

*Earlier versions of this essay were presented at the University of Chicago, Central Washington University, the University of Warwick, the University of Illinois at Chicago, and the University of Leipzig. I would like to thank the audiences at these talks for valuable feedback. I would also or especially like to thank the following people for helpful discussion of earlier versions of the essay: Amos Browne, Agnes

References

Audi, Robert. 1993. *The Structure of Justification*. Cambridge: Cambridge University Press.

Black, Max (ed.) 1965. *Philosophy in America*. Ithaca: Cornell University Press.

Boghossian, Paul. 2014. What is inference? *Philosophical Studies* 169: 1–18.

Booth, Anthony Robert and Darrell P. Rowbottom (eds.) 2014. *Intuitions*. Oxford: Oxford University Press.

Boyle, Matthew. 2015. Die Spontaneität des Verstandes bei Kant und einigen Neokantianern. *Deutsche Zeitschrift für Philosophie* 63(4): 705–726.

Broome, John. 2013. *Rationality Through Reasoning*. Oxford: Wiley-Blackwell.

— 2014. Comments on Boghossian. *Philosophical Studies* 169: 19–25.

Burge, Tyler. 1996. Our entitlement to self-knowledge. *Proceedings of the Aristotelian Society* 96: 91–116.

Cassam, Quassim. 2007. Ways of knowing. *Proceedings of the Aristotelian Society* 107: 339–358.

Chudnoff, Elijah. 2014. The rational roles of intuition. In Booth and Rowbottom 2014: 9–35.

Ford, Anton. 2011. Action and generality. In Ford, Hornsby, and Stoutland 2011: 76–104.

Callard, Thomas Crowther, Rebeka Ferreira, David Finkelstein, Jeremy Fix, Anton Ford, Reza Hadisi, Jonas Held, Ulf Hlobil, Errol Lord, Eric Marcus, Santiago Mejia, Josh Mendelsohn, Ram Neta, Edgar Phillips, Julia Tanney, Matt Teichman, Markos Valaris, Jonathan Way, Ralph Wedgwood, Andy Werner, Malte Willer, and several anonymous referees, including two for *Synthese*.

Ford, Anton, Jennifer Hornsby, and Frederick Stoutland (eds.) 2011. *Essays on Anscombe's* Intention. Cambridge, MA: Harvard University Press.

Fumerton, Richard. 2004. Inferential internalism and the presuppositions of skeptical arguments. In Schantz 2004: 157–168.

Grice, H. P. 2001. *Aspects of Reason*. Oxford: Clarendon Press.

Hlobil, Ulf. MS. In defense of the taking condition.
— 2014. Against Boghossian, Wright, and Broome on inference. *Philosophical Studies* 167: 419–429.

Hossack, Keith. 2007. *The Metaphysics of Knowledge*. Oxford: Oxford University Press.

Koziolek, Nicholas. 2015. *Representing Reason*. PhD thesis, The University of Chicago.

Lavin, Douglas. 2011. Problems of intellectualism: Raz on reason and its objects. *Jurisprudence* 2(2): 367–378.

Leite, Adam. 2008. Believing one's reasons are good. *Synthese* 161: 419–441.

Longuenesse, Béatrice. 2017. *I, Me, Mine: Back to Kant, and Back Again*. Oxford: Oxford University Press.

Marcus, Eric. 2012. *Rational Causation*. Cambridge, MA: Harvard University Press.

McGlynn, Aidan. 2014. *Knowledge First?* New York: Palgrave Macmillan.

McHugh, Conor and Jonathan Way. MS. What is reasoning? Forthcoming in *Mind*. DOI: 10.1093/mind/fzw068.
— 2016. Against the taking condition. *Philosophical Issues* 26, *Knowledge and Mind*: 314–331.

Neta, Ram. MS. The basing relation.
— 2013. What is an inference? *Philosophical Issues* 23: 388–407.

Pauer-Studer, Herlinde. 2014. Rational requirements and reasoning. *Economics and Philosophy* 30(3): 513–528.

Rödl, Sebastian. 2007. *Self-Consciousness*. Cambridge, MA: Harvard University Press.

Schantz, Richard (ed.) 2004. *The Externalist Challenge*. Berlin: Walter de Gruyter.

Sellars, Wilfrid. 1967. Some remarks on Kant's theory of experience. In Sicha 2002: 269–282.

Setiya, Kieran. 2013. Epistemic agency: some doubts. *Philosophical Issues* 23: 179–198.

Sicha, Jeffrey F. (ed.) 2002. *Kant's Transcendental Metaphysics: Sellars' Cassirer Lecture Notes and Other Essays*. Atascadero, California: Ridgeview Publishing Company.

Sosa, Ernest. 2009. *Reflective Knowledge: Apt Belief and Reflective Knowledge*, *Volume II*. Oxford: Oxford University Press.

Strawson, Galen. 2003. Mental ballistics or the involuntariness of spontaneity. *Proceedings of the Aristotelian Society* 103(1): 227–256.

Stroud, Barry. 1979. Inference, belief, and understanding. *Mind* 88(350): 179–196.

Thomson, Judith Jarvis. 1965. Reasons and reasoning. In Black 1965: 282–303.

Tucker, Chris. 2010. When transmission fails. *Philosophical Review* 119: 497–529.

— 2012. Movin' on up: higher-level requirements and inferential justification. *Philosophical Studies* 157: 323–340.

Valaris, Markos. 2014. Reasoning and regress. *Mind* 123(489): 101–127.

Wedgwood, Ralph. 2006. The normative force of reasoning. *Noûs* 40(4): 660–686.

Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.

Wright, Crispin. 2014. Comment on Paul Boghossian, "What is inference?" *Philosophical Studies* 169: 27–37.