

Safety, Explanation, Iteration

Daniel Greco

Forthcoming in *Philosophical Issues*, a supplement to *Noûs*

Abstract

This paper argues for several related theses. First, the epistemological position that knowledge requires safe belief can be motivated by views in the philosophy of science, according to which good explanations show that their explananda are robust. This motivation goes via the idea—recently defended on both conceptual and empirical grounds—that knowledge attributions play a crucial role in explaining successful action. Second, motivating the safety requirement in this way creates a choice point—depending on how we understand robustness, we’ll end up with different conceptions of safety in epistemology. Lastly, and most controversially, there’s an attractive choice at this point that will *not* vindicate some of the most influential applications of the safety-theoretic framework in epistemology, e.g., Williamson’s (2000) arguments against the KK principle, and luminosity.

1 Safety and Explanation

Much recent epistemology defends the idea that, for a true belief to constitute knowledge, it must be *safe*.¹ While different versions of this idea have been proposed, they all share in common the idea that if a subject S knows that P, then there is some range of relevant

For helpful comments and discussion, thanks to Kevin Dorst, Elizabeth Miller, Bernhard Salow, Jack Spencer, Jason Stanley, Michael Strevens, and audiences at both MIT and an APA symposium on the epistemology of higher-order states.

¹Some of the most influential early works in this regard were Sosa (1999) and Williamson (2000).

situations in which S is safe from error in believing that P—if S falsely believes that P in one of the relevant cases, then S’s belief is not safe. In some sense, it is in *danger* of being false, and so does not constitute knowledge. It is often specified that the relevant cases to a given case C are the ones that are *nearby* or close to C.² I do not build this into my characterization of safety, because, ultimately, I want to suggest that the most attractive versions of safety needn’t involve appealing to anything like a distance metric on possible cases. But that will be the payoff, arrived at only towards the end of the paper. My strategy for getting there will be to argue that an attractive motivation for something very much *like* safety, doesn’t actually get us all the way to the particular versions of that requirement typically appealed to by epistemologists.

Sometimes versions of the safety requirement are motivated by appeal to thought experiments in which they seem to deliver plausible results, especially as contrasted with other modal requirements on knowledge (e.g., sensitivity requirements). My aim in this section is to offer a more theoretical motivation for safety. The first main assumption I’ll make will be that in a good explanation, the explanandum is shown to be, somehow, *robust*. The notion of robustness appealed to by philosophers of science working on explanation is closely related to (though more general than) the notion of safety used by epistemologists, though to my knowledge, the connection between these literatures hasn’t yet been drawn. The second main assumption in the argument will be that knowledge attributions play a crucial role in explaining successful action. I don’t have anything new to say in defense of this idea, though I will point to some relevant recent work on the topic. I’ll argue that, given this second assumption, the idea that knowledge requires safety is plausibly just a special case of the more general idea that good explanations (whether psychological or not) show their explananda to be robust.

²See, e.g., Pritchard (2009).

1.1 Explanation and Robustness

Suppose a marble is dropped in a basin, and after rolling around for awhile, it comes to rest at the bottom of the basin.³ Consider two ways we might go about explaining why it came to rest where it did:

Bad: Calculate the trajectory of the marble, given the starting point, and the gravitational and frictional forces acting on the marble at each moment. Note that at the end of its trajectory, it is at rest at the bottom of the basin.

Good: Appeal to some general principles about gravity and potential energy, to show that, *no matter where* it was dropped from, it would've ended up at rest at the bottom of the basin.

As the labels suggest, the second explanation has a virtue that the first lacks. A natural way of capturing this is in terms of *robustness*. If you understood the first explanation, but not the second, you might think that the marble's ending up at the bottom of the basin was an accident or fluke. But once you grasp the second explanation, you see that it's no accident that the marble came to rest where it did—not only did it *actually* end up there, but it would have done so under a wide range of conditions.

Let's take one more example. Suppose a car's gas pedal is depressed, and the car speeds up. Why did that happen? A satisfying explanation will abstract away from lots of detail—it will mention how gas pedals generally work, but won't discuss the details of, e.g., the car's initial velocity, the type of road, whether the car was on an incline, etc. If the explanation works, then it shows that the explanandum—the car's speeding up—was robust; given the depression of the gas pedal, the car would have sped up across a wide range of conditions.⁴

Granted, some explananda are *not* robust, and in such cases, the best explanation that can be offered will be analogous to the “bad” explanation above. If you hit the

³The example is from (Strevens, 2008, pp.434-5).

⁴See (Woodward, 2000, pp.215-17)

jackpot in a fair lottery, the best we can do to explain this is to show how your winning was possible, and had a certain (low) chance of occurring. We can't show that it was robust, because it wasn't. So while I do claim that it's a *virtue* of an explanation that it shows its explanandum to be robust, I don't claim that *all* explanations involve robustness, or even that the *best* explanations always involve robustness—sometimes, there's none to be had. But this shouldn't be surprising—sometimes the best explanation of an event is the best of a bad lot; when our explanandum is, ultimately, a fluke, it won't admit of *any* good explanation.⁵ We'll return to the idea of robustness momentarily. First, however, we'll have to draw a connection between explanation on the one hand, and knowledge on the other.

1.2 Explaining Action

There's a rich tradition in the philosophy of mind according to which a central role of propositional attitude ascriptions is to explain behavior.⁶ Why did Alice carry an umbrella this morning? Because she believed that it was going to rain, and wanted not to get wet. So far, not so controversial. More controversially, a number of philosophers have argued that knowledge ascriptions in particular play an ineliminable role in explaining behavior.⁷ Suppose we explain why Alice was at the Colosseum last evening by appeal to her desire to see Roman antiquities, and her knowledge that there is a metro station located conveniently right next to the Colosseum. In this case, it's not so clear that we need to appeal to her *knowledge* in order to explain her success—one might have thought it's enough to appeal to her *beliefs* (e.g., her beliefs about the location of the Metro), together with non-mental environmental conditions (e.g., facts about the actual location of the Metro) in order to explain her arriving at the Colosseum.

⁵See (Strevens, 2008, pp.171-2) for some discussion of cases like this. Whether it's right to say that your winning a fair lottery has no explanation at all, rather than no explanation *of a certain sort* or no explanation *with certain virtues* won't much matter for my purposes.

⁶For some influential examples, see Lewis (1974), Dennett (1981), Stalnaker (1984), Millikan (1984), and Dretske (1988).

⁷See, e.g., Williamson (2000), Gibbons (2001), and Nagel (2013).

While there's been a good deal of literature on this question, it has tended to focus on the nature of psychological explanation in particular—may the mental states that figure in psychological explanations be broad—i.e., incorporate information about the relationship between a subject and her environment—or must they be narrow?⁸ These are interesting and important questions, but my aim here is to sketch a route to the idea that knowledge as such plays an important role in explanations that doesn't turn on considerations specific to *psychological* explanation as such.

Often, it's not just behavior, but *success* that we take as an explanandum. We ask how Alice managed to end up, not just anywhere, but *where she wanted to be*. And in many of these cases, true beliefs are, effectively, necessary for success; barring outlandish scenarios, Alice won't end up at the Colosseum without some true beliefs about how to get there.

Let's suppose Alice successfully makes her way to the Colosseum. However, let's also suppose that her true beliefs about how to get there are *mere* true beliefs—in particular, suppose they are unsafe, and so (given a safety requirement) do not constitute knowledge. E.g., suppose Alice comes to believe, out of sheer optimism, that there's a train to the Colosseum that stops at the nearest metro station to her hotel, and as it turns out she's right. Two observations:

1. Her belief about how to get to the Colosseum is not knowledge.
2. We don't seem to have a satisfying explanation of why she ended up where she wanted to be—we have to regard her success as a fluke.

I suggest that, if there is a safety requirement on knowledge, (1) and (2) are related—unsafe beliefs will typically not provide good explanations of successful actions. When a subject's belief is unsafe, there is some relevant case in which her belief is false. Call that case “E”, for error. If the subject's success depends upon her holding a true belief,

⁸See Pettit (1993) for the argument that psychological explanation requires appealing to narrow mental states, and Williamson (2000), especially chapters 1-3, for arguments against such a restriction.

then in E, she will not be successful. In the example above, “E” might refer to a case in which there is no train to the Colosseum that stops at the nearest station to Alice’s hotel. And in such a case, she will not be successful—she will not reach the Colosseum. That strongly suggests, however, that her reaching the Colosseum is not robust—after all, there is a relevant case in which it fails to occur.

Can we tighten up this line of thought? Suppose a subject S truly believes that P, and successfully acts upon that belief. If safety theorists are right, then there will be some range of cases R_1 such that if S is to know that P, S cannot be in error in any of the cases in R_1 . If S’s success is to be robust—given our assumptions, robustness is required for her success to admit of a good explanation—then there will be some range of cases R_2 such that S succeeds throughout R_2 . If R_1 is a subset of R_2 , then we have our conclusion—successful actions cannot be explained by unsafe beliefs, because in such cases, success will not be robust. Success based on unsafe beliefs is always a fluke.

Even if R_1 is not a subset of R_2 , as long as *many* of the cases in R_1 are also in R_2 , there will be a strong correlation between a belief’s being unsafe, and successful action based on that belief being inexplicable. Ultimately, as long as there is a good deal of overlap between the range of cases in which error must be avoided in order for a belief to be safe, and the range of cases in which an event must occur for it to be robust, there will be a strong connection between safety on the one hand, and the explicability of success on the other.⁹

So far this should all be congenial to the safety theorist. I’ve tried to show how a popular package of views in epistemology—the views that knowledge both requires safety, and plays a special role in explaining successful action—can be motivated by appeal to general considerations about robustness as an explanatory virtue.

⁹Depending on just how strong the safety theorist thinks the connection between safety and knowledge is, this can lead to a weaker or stronger connection between knowledge on the one hand, and the explicability of success on the other. If safety is necessary *and* sufficient for knowledge, then the connection will be very tight. And such views have been defended—see, e.g., Aarnio (2010). But even views on which safety is merely necessary for knowledge can appeal to the above considerations to argue that there is an important connection between knowledge and the explicability of success.

2 Two Views of Robustness

Just how should we understand the sort of robustness required for good explanations? In this section I'll identify a choice point. If we understand robustness one way, it will support (via the considerations discussed in the previous section) the version of safety typically defended by safety theorists in epistemology. If we understand robustness in a different way, however, it will support a different version of safety, with some very different consequences in epistemology.

I mentioned earlier that safety theorists in epistemology tend to define safety in terms of similarity or distance—a belief is safe in case C just in case it isn't false in any cases *similar* or *nearby* to C. We can define a notion of nearby case robustness (NCR) in a parallel way:

NCR: An event E is robust in case C just in case E occurs in all cases *similar* or *nearby* to C.

How would this idea apply to examples of explanatory robustness discussed in the previous section? In the case of the marble dropped in the basin, an advocate of NCR could say that the more general, abstract explanation is good because it shows that, not only did the marble *actually* end up at rest at the bottom of the basin, but it would have done so in all cases nearby to the actual case—e.g., even if it had been dropped from a slightly different position. Similarly, in the case of the decelerating car, an advocate of NCR could say that a good explanation will show why, even in conditions similar but not identical to the actual conditions, the car still would have sped up had its gas pedal been depressed.

NCR gives us a case-relative notion of robustness—for an event to be robust *at a case C*, it must occur at all cases *close to C*. It may seem obvious that any adequate explication of robustness will be case-relative in something like this sense. But this isn't so. We might instead appeal to context-sensitivity to do much of the same work, ending

up with a context-sensitive notion of robustness (CSR):

CSR: An event E is robust given standards of context X ¹⁰ just in case E occurs in all relevant (by the standards of X) cases.¹¹

The idea that what counts as a good explanation might be sensitive to context in various ways, is a familiar one.¹² According to CSR, one way in which the goodness of explanation is context sensitive is that, while events must hold in some range of cases in order to be robust (i.e., in order for them to admit of good explanations), *which* range of cases they must hold in will depend on our explanatory context—perhaps, it will depend on e.g., our interests as inquirers, or our background knowledge, or what we’re willing to take for granted. Just how these sorts of context dependence might come into play in discussions of robustness should become clearer as we compare and contrast NCR and CSR.

Concerning the two examples already introduced, an advocate of CSR can offer very similar diagnoses to an advocate of NCR. E.g., she can say that in the case of the dropped marble, the contextually relevant range of cases includes cases where the marble is dropped from different locations to the one at which it was actually dropped (though not locations *so* different that it wouldn’t fall in the basin), and that a good explanation will show why it would’ve come to rest at the bottom of the basin throughout that range of cases. Similarly, with the accelerating car, she can say that the contextually relevant range of cases includes cases where road conditions were different in various ways (though not *so* different that the gas pedal would be ineffective).

Philosophers of science who write about robustness have not addressed this choice point, and what they have said seems, to me, to admit of both NCR and CSR-friendly

¹⁰I use “X” for context, rather than “C”, since “C” is already being used to refer to cases.

¹¹I don’t claim that NCR and CSR are the only two ways one might say more about what it takes for an event to be robust, but they are the only ones I’ll discuss. Also, it’s certainly possible to combine NCR and CSR—one could endorse NCR, but also hold that context determines just *how* near cases must be in order for them to be relevant to attributions of robustness. I’ll come back to this point later in the paper.

¹²See, for example, van Fraassen (1980) and Achinstein (1983).

interpretations. Here are some typical examples:

A connection between properties or complexes of properties is robust if it holds under a wide range of circumstances, actual and counterfactual. (Strevens, 2008, p.433)

A generalization is invariant if it is stable or robust in the sense that it would continue to hold under a relevant class of changes. (Woodward, 2000, p.197)

The NCR advocate can understand “wide range of circumstances” and “relevant class of changes” as referring to a range of cases you get by starting with a given case, and venturing out some distance in modal space, where the relevant notion of distance does *not* depend in any way on our explanatory context. The CSR advocate, by contrast, can understand those phrases as referring to a range of relevant cases somehow determined by our interests or concerns as inquirers.

Without saying a great deal about both (1) which cases count as near to a given case, according to the NCR advocate, and (2) which cases count as relevant in a given context, according to the CSR advocate, it will be very hard to identify examples of putative robustness (or non-robustness) in which one of NCR and CSR but not the other seems to deliver adequate results.¹³ But that’s not the only way to approach the choice between NCR and CSR.

My strategy in the remainder of this paper will be as follows. First, I’ll identify some important structural differences between NCR and CSR, particularly as they handle the idea of *iterations* of robustness. Next, I’ll explore how those different conceptions of robustness would lead to different conceptions of safety requirements and iterations of knowledge in epistemology. Lastly, I’ll address the choice between NCR and CSR, and the associated versions of safety.

¹³Moreover, there might be principled obstacles to saying enough about (1) to force the NCR advocate to take positions on particular cases. E.g., Williamson (2000) defends the version of safety that (I’ve argued) can be motivated by NCR, but he thinks there are principled reasons why we can’t say much in non-epistemological terms about what it takes for a case to be near to a given case.

3 Iteration

Suppose some event is robust, in the explanatorily relevant sense. What does it take for the event to be robustly robust? As we'll see in this section, NCR and CSR give very different answers to this question, in ways that will have ramifications in epistemology.

According to NCR, if an event is robust in a case C , it occurs in all cases close to C . So if the event is to be robustly robust, it must not only occur, but robustly occur in all cases close to C . Given NCR, that means it must not only occur in all cases close to C , but also in all cases close to cases close to C . If there are no nearby cases in which E fails to occur, but there are *nearly* nearby cases in which it fails to occur, then E will be robust, but not robustly robust. The version of safety motivated by NCR allows for parallel failures of iteration of safety—if there are no nearby cases in which a belief is false, but there are nearly nearby such cases, then the belief will be safe, but not safely safe. Ultimately, this possibility of iteration failure is at the heart of Williamson's anti-luminosity argument, and his more specific arguments against the KK principle (2000, chapters 4, 5). One can know, without knowing that one knows, because one's belief can be safe, without being safely safe.

What about CSR? How does it handle iteration? According to CSR, if an event is robust by the standards of a context X just in case it occurs in all relevant $_X$ cases. So an event is robustly robust (by the standards of X) just in case it robustly occurs in all relevant $_X$ cases. Applying the definition of CSR, this means that an event is robustly robust just in case in all relevant $_X$ cases, in all relevant $_X$ cases, it occurs. But it's not so clear how to understand this requirement. In the case of NCR, iterating "nearby" clearly makes a difference—a case can be nearly nearby, without being nearby. But in the case of CSR, it's not so clear that iterating "relevant" does make a difference. After all, according to CSR, it needn't be the case that each case determines its own set of relevant cases, so that we get a different, stricter requirement by holding that robustly

robust events must hold in all cases that are relevant to cases that are relevant, rather than just in all relevant cases. To appreciate the point, it will help to see how it applies in the context of one of our previous examples.

Recall the marble dropped in the basin. Suppose we don't know exactly what location the marble was dropped from, only that it was dropped somewhere above the basin. Given CSR, it's natural to think that our context will determine the set of cases in which the marble is dropped somewhere above the basin as the relevant one—for an explanation to count as a good one, by the standards of our context, it must show why the marble would've come to rest at the bottom of the basin, no matter where (among the contextually relevant range of locations) it was dropped from. That's what it would take to show our explanandum to be robust. But it's not as if there must be some *further* relativity of relevant cases to where the marble was *actually* dropped from—whether it was actually dropped right above the center of the basin, or closer to one of the edges, the relevant class of cases will be the same. So we won't get a different, stricter requirement for robust robustness by holding that, for the marble's ending up at the bottom of the basin to be robustly robust, it must not only end up there in all relevant cases, but also in all cases relevant to cases that are relevant—because our context determines a single set of relevant cases that doesn't vary with the actual location from which the marble was dropped, the additional clause is at worst meaningless, and at best trivial.¹⁴

¹⁴That's not to say there's no way to make sense of iterated attributions of robustness given CSR—there are various questions that we might, charitably and opportunistically, interpret questions about iterated robustness as getting at. E.g., suppose there's some natural sense in which the marble could've failed to be dropped at all (maybe we're playing a game in which it's up to a player's choice whether or not to drop a marble in a basin). We might express this by saying that the marble's ending up at rest in the bottom of the basin is robust (because it would've done so no matter where it was dropped from) but not robustly robust (because it could've easily failed to be dropped at all). While the advocate of CSR has some work to do in explaining how she can make sense of this sort of claim, I don't see any serious obstacles. In particular, there are independent reasons to hold that, different occurrences of one and the same context-sensitive term might receive different interpretations, even in a single sentence. See Stanley and Williamson (1995). So when we say an event is robustly robust, the two occurrences of the context-sensitive term “robust” are given two different interpretations—the outer “robust” picks out a different range of relevant cases than the inner one. This lets us accommodate the idea that robustly robust events are *more* robust than merely robust events. In the case of the marble's ending up at the bottom of the basin, robust robustness might amount to not only being robust under changes in dropping

I mentioned before that NCR's stricter requirements for iterated robustness have important parallels in epistemology. What sort of epistemological picture is suggested by CSR, and the very different approach to iterating robustness that it entails? First, what sort of conception of safety is suggested by CSR? If a belief is safe just in case it is robustly true, then CSR suggests a picture where safety amounts to truth in all *contextually relevant* cases, rather than all *nearby* cases. And for reasons very much along the same lines as those discussed in the previous paragraph, this leads to a very different picture of what it takes for beliefs to be safely safe than the one familiar in recent epistemology. Again, to see how, it will help to work through an example.

Suppose we take for granted the following: Suzy has an excellent memory, especially when it comes to world geography. She's asked what the capital of Vanuatu is, thinks for a moment, and correctly answers "Port Vila." We're now wondering whether her belief is safe, and amounts to knowledge. It's natural, given the CSR-inspired version of safety, to think that the relevant cases in our context will be cases where Suzy's memory is functioning normally, and so will all be cases in which her belief is true. So her belief will count as safe. What will it take for it to be safely safe? Will it be harder than merely being first-order safe? For there to be some further, stricter requirement to meet for her belief to be safely safe, it would have to be the case that while there are no relevant cases where her belief is false, there *are* cases relevant to relevant cases in which it is false. But as before, it may be that our context determines just one set of relevant cases, and that this set doesn't itself vary from case to case. If that's our situation, then a belief's being safely safe won't require it to meet any stricter requirement than merely being safe.

This contextualist version of safety, in which iterations of safety can follow trivially conditions, but *remaining* robust under changes in dropping conditions even when there are *also* changes in the motivations of the players (concerning, e.g., whether to drop at all). It's worth pointing out that interpreting cases like this as involves failures of second-order but not first-order robustness is somewhat tricky for the advocate of NCR too. After all, if we really accept that the marble could've easily failed to be dropped at all, it's not so clear that the advocate of NCR is entitled to the claim that it's ending up at the bottom of the basin is first-order robust.

from first-order safety, fits in nicely with extant contextualist accounts of knowledge. As I argue in Greco (2014), a broadly Lewisian contextualist framework can be used to defend the KK principle—the thesis that if a subject knows that P, then she knows that she knows that P.¹⁵ And the reasons why it can be defended in this framework are structurally very similar to the reasons why iterations of robustness and safety can come for free in the contextualist picture we’ve been exploring—on the view I defend in my (2014), knowing requires avoiding error in a contextually determined set of cases, but which cases those are doesn’t itself systematically vary with the case, but only with the context. And as we’ve seen, when the set of relevant cases (whether relevant for attributions of robustness, safety, or knowledge) doesn’t itself vary from case to case, but only from context to context, second-order robustness/safety/knowledge isn’t more demanding than its first-order cousin.

Let’s take a step back. So far I’ve argued that the view that knowledge plays a special role in explaining successful action can be used to provide—via the idea that robustness is an explanatory virtue—a novel motivation for a safety requirement in epistemology. I’ve also identified a choice point—we can understand an event’s being robust as its holding in all *nearby* cases (NCR), or in all *contextually relevant* cases (CSR). Depending on which choice we make at this point, if we go on to endorse a version of safety motivated by the corresponding conception of robustness, we get very different epistemological consequences concerning the obstacles (or lack thereof) to iterations of safety (and, ultimately, knowledge). I haven’t yet, however, provided any reasons to favor one choice or the other. I turn to that task in the next section.

¹⁵In that paper I claimed that Lewis’s (1999) account itself already entailed KK, following Williamson (2001), who credits Lloyd Humberstone with pointing out that Lewis’ theory vindicates some strong principles of epistemic logic, including KK. But since then I’ve been convinced by Holliday (2015) and Salow (Forthcoming) that Lewis’ own account does not support KK, even if closely related views do.

4 In Favor of CSR

In this section, I'll offer two main reasons to favor CSR over NCR. I don't take them to be dispositive, and I won't attempt to consider everything that advocates of NCR might say in reply. My aim isn't to *settle* the question of which view is preferable, so much as to make it clear that this is a debate that should be had. In the epistemological literature, safety is defined in terms of nearness as a matter of course, and consequences of this picture of safety are treated as relatively secure premises that can be relied on in further argument.¹⁶ But if there is a well-motivated alternative conception of safety, suitably related to a well-motivated conception of robustness, then this practice is premature.

4.1 Explanations Involving Iterated Safety

In the previous section we saw that NCR and CSR have different consequences concerning the obstacles to iterations of safety, but we didn't yet see any reason to think that iterations of safety play an important explanatory role. In this subsection I want to draw attention to a class of explanations that pose no difficulty for CSR, but which must be rejected, or at least significantly reinterpreted, if we opt for NCR.

Consider the contrast between the following two cases:¹⁷

PUBLIC ANNOUNCEMENT: A professor tells her class that they will play the following game. Without communicating to one another in any way, each student in the class will write down the name of a US state on a piece of a paper. If all students write the same state name, with the exception of the name of the state the class is taking place in, the students will each receive \$10. If any two students write down different state names, or if they all write down the name of the state the class is taking place in, no prize money will

¹⁶Even, e.g., when they lead to the result that one can know that P, while it is arbitrarily improbable, given one's evidence, that one knows that P. See Williamson (2014).

¹⁷The cases are from Greco (2015), but they are inspired by similar cases in Heal (1978).

be awarded. Before handing out the pieces of paper, the professor tells the class that she grew up in Maine (which is not the state the class is taking place in), and that it is lovely in the fall.

PRIVATE INFORMATION: Just like the previous case, except instead of publicly announcing that she grew up in Maine, the professor whispers the following to each student privately as she hands out the pieces of paper: “while I’m not telling anybody else this, I’d like you to know that I grew up in Maine, and it is lovely in the fall.”

Suppose the students all write down “Maine”, and win the prize. In PUBLIC ANNOUNCEMENT, this would be unsurprising, and eminently explicable. By contrast, in PRIVATE INFORMATION, this would be quite surprising, and would call out for some further explanation. What’s the difference? It’s tempting to explain the contrast in terms of safety—or equivalently, given a safety-theoretic conception of knowledge, in terms of knowledge. In PRIVATE INFORMATION, if each student writes down “Maine”, they must be taking Maine to be a more likely candidate for coordination than any other state. This would make sense if they believed that Maine had been made salient in some way—had been drawn to the students’ attention. And while that is in fact true in the case—Maine has been drawn to each student’s attention—they are not in a position to know it, since each student knows only that Maine has been drawn to *her own* attention, but not to the attention of the others. So if the reason they successfully coordinate in PRIVATE INFORMATION is that each student believes that the other students are more likely to write down “Maine” than any other state, then the “explanation” of their success is that they all made lucky guesses that happened to pay off. That is, their success depending on their taking something for granted—that Maine had been drawn to the other students’ attention, and was therefore a likely candidate for coordination—that they could’ve easily been wrong about. This is why their success would be so surprising;

it would be based on unsafe beliefs.

This suggests that, if only PRIVATE INFORMATION were modified such that the students know that Maine has been drawn to the students' attention—so that they couldn't easily have been wrong about that—then we'd get a case in which their successful coordination would be unsurprising, and explicable. But that suggestion would be too quick. Consider the following case:

MORE PRIVATE INFORMATION: Just like PRIVATE INFORMATION, except this is what the professor whispers: "I'm privately telling everybody in the class that I grew up Maine and that it's lovely in the fall. However, you're the only one who I'm telling *that* I'm telling everyone. Each other student thinks that she's the only one who knows that I grew up Maine."

We already established that in PRIVATE INFORMATION, absent some further story, successful coordination would be an inexplicable fluke. But in MORE PRIVATE INFORMATION, each student thinks that the *other* students take themselves to be in a situation like PRIVATE INFORMATION, and so to be unlikely to pick Maine as opposed to any of various other potentially salient states. So if students are unlikely to coordinate on Maine in PRIVATE INFORMATION, they're also unlikely to coordinate on Maine in MORE PRIVATE INFORMATION. It's easy to see how to keep going constructing variants of PRIVATE INFORMATION in which, at least if students are thinking clearly, coordination would be surprising, and would call out for a special explanation: just let the next case in the sequence be one where all the students think the other students think they're in the previous case in the sequence.¹⁸

¹⁸To spell it out, the first case would be one where nobody is told anything. The second case would be PRIVATE INFORMATION, where while each student has been told that Maine is special, they also think the other students think they're in the first case, where nobody has been told anything. The third case would be MORE PRIVATE INFORMATION, where each student thinks the other students think they're in PRIVATE INFORMATION. The fourth case is somewhat tricky—it would be one where the teacher leads each student to believe that the rest of the students think they're in MORE PRIVATE INFORMATION. The instruction might go like this: "I'm privately telling everybody in the class that I'm privately telling

I claimed earlier in this section that successful coordination in PUBLIC ANNOUNCEMENT would be unsurprising, and eminently explicable. By contrast, in the cases in the hierarchy starting with PRIVATE INFORMATION, successful coordination would be at worst an inexplicable fluke, and would at best call out for a *different* explanation than successful coordination in PUBLIC ANNOUNCEMENT. What explains this contrast? Many writers have thought that the key difference between PUBLIC INFORMATION and any of the cases in the hierarchy starting with PRIVATE INFORMATION and MORE PRIVATE INFORMATION is that in PUBLIC INFORMATION, the students' knowledge doesn't give out at *any* level of iteration. That is, they all know that Maine has been singled out, they all know that they all know this, they all know that they all know that they all know this...and so on for as many iterations of "they all know" as you like. In the jargon the fact that Maine has been singled out is *common knowledge*.¹⁹ Or put in more safety-theoretic terms, none of the students could easily have been wrong about whether Maine had been singled out, nor could they have easily been wrong about whether any students could easily have been wrong about that, or whether any students could easily have been wrong about whether any students could easily have been wrong about whether any students could easily have been wrong, and so on. By contrast, in cases like PRIVATE INFORMATION, there's *some* relevant fact that somebody could've easily been wrong about—maybe whether Maine has been singled out, maybe whether other students could easily have been wrong about whether Maine has been singled out, maybe whether other students could easily have been wrong about whether other students could easily have been wrong about whether Maine has been singled out, etc.

This explanation of the contrast, however, sits much more easily with the CSR conversation in the class that I grew up in Maine and that it's lovely in the fall. So while all the other students think all the other students think I grew up in Maine, each of them thinks she's the only one who knows that everyone knows this. You're the only one who knows that, not only did I grow up in Maine, and not only does everybody know this, but everybody knows *that* everybody knows this." Hopefully it should be clear how the hierarchy could be continued.

¹⁹See Greco (2015) for some discussion of the connection between common knowledge on the one hand, and iteration principles in epistemology on the other.

ception of safety and robustness, rather with the NCR. This is because NCR leads very naturally to the following commitment: any belief that is metaphysically possibly false won't have arbitrarily many iterations of safety. Here's why. If it's metaphysically possible that S falsely believes that P, then even if the possibility in which S falsely believes that P is quite remote, it will be near to a case which is near to a case which is near to a case...[repeat as many times as necessary]...which is near to the actual case. That is to say, even if it is safe, there will be some n such that it is not safely safely safely...[repeat n times]...safe. While this commitment is perhaps not strictly speaking *forced* on the defender of NCR, it's very hard to avoid, especially if NCR will bear the argumentative weight typically placed on it.²⁰

Given this commitment, we can't offer the popular explanation of the contrast—even in PUBLIC INFORMATION, there will be some case—perhaps quite remote, but still nearly nearly nearly...[repeat some large n times]...nearby—in which some student is wrong about whether Maine has been singled out. So there would be nothing to distinguish PUBLIC INFORMATION from some case—perhaps one relatively far along—in the hierarchy starting with PRIVATE INFORMATION. The defender of NCR might accept this commitment, and deny that there really *is* anything so special about PUBLIC INFORMATION, that distinguishes it from *all* cases—even those quite far along—in the hierarchy starting with PRIVATE INFORMATION.²¹ But to the extent that we're sympathetic to the idea that there is an important contrast here, and to the idea that the account above does a good job of capturing it, then we should be sympathetic to CSR, as opposed to NCR.

²⁰For instance, the conception of safety used in Williamson's antiluminosity argument (Williamson, 2000, chap. 4) requires that if some condition C holds in some case w , but it is possible to gradually transition from w to a case in which C does not hold, then a belief in w that C holds will have some finite number of iterations of safety—there will be a nearly nearly nearly...[insert as needed]...nearby case in which the belief that C holds is false. Since this argument is meant to apply extremely broadly, it's natural to interpret Williamson as being committed to the claim that, quite generally, beliefs have only finitely many iterations of safety.

²¹Lederman, in both his (2015) and (Ms.), argues that explanations appealing to common knowledge are dispensable, and would likely say the same about the idea that there is some important contrast between PUBLIC INFORMATION and the hierarchy starting with PRIVATE INFORMATION.

As we saw in the previous section, CSR doesn't create the same obstacles to iterated safety that NCR does, so there's nothing to rule out that, by the standards of typical context in which PUBLIC INFORMATION is being discussed, there are no relevant cases in which any of the students is wrong about whether Maine has been singled out, or wrong about what any of the other students believe (or what they believe others believe, or what they believe others believe others believe, and so on). That is, there's nothing to rule out the following simple explanation of the contrast—in PUBLIC INFORMATION, there are no relevant cases in which anybody is wrong about relevant information, while throughout the hierarchy starting with PRIVATE INFORMATION, there are such cases.²²

It may help to take this a bit more slowly. Many writers have explained the contrast between PUBLIC INFORMATION and the PRIVATE INFORMATION hierarchy in terms of common knowledge, which is itself often explained in terms of infinitely iterated knowledge—in PUBLIC INFORMATION, the students all know that they all know that they all know...that Maine has been singled out.²³ Given safety requirements on knowledge, and NCR, it's hard to allow for the possibility of such infinitely iterated knowledge. Even in a case like PUBLIC INFORMATION, where the students seem quite safe from error, because errors are still *possible*, there will be some case in which some student is wrong about whether Maine has been singled out, not near to PUBLIC INFORMATION, but nearly nearly nearly...near, for some number of “nearly”s. That's enough to rule out—via arguments that have been given elsewhere²⁴—that the students all know that they all know that they all know...that Maine has been singled out, for arbitrarily many iterations of “they all know that.” By contrast, if we adopt CSR, there's no parallel argument that, even by whatever standards for relevance in place in discussions of PUBLIC

²²To be clear, this is not offered as a non-question-begging argument against a broadly Williamsonian conception of safety and iterated knowledge. Williamson (2000, ch. 6) is clear that his version of the safety requirement rules out the possibility of common knowledge, and he welcomes this consequence on the grounds that it allows him to avoid some paradoxes involving common knowledge. My aim in this section has been to draw attention to one of the costs of taking on board this package—it requires us to forego explanations that appeal to common knowledge.

²³There are weaker versions of this idea. See Greco (2015) for discussion.

²⁴Primarily, by Timothy Williamson (2000, chapters 4, 5).

INFORMATION, there must be a relevant relevant relevant...relevant case in which some student is in error, for some number of iterations of “relevant.” As we saw in the previous section, iterating “relevant” doesn’t always (or even often) make sense. It may just be that, in a typical context in which PUBLIC INFORMATION is being discussed, there is a single set of possibilities relevant for interpreting knowledge attributions concerning the students, and in none of those possibilities are any of the students wrong about pertinent information concerning what the teacher said, or other students’ states of mind. And that’s enough to block any quick argument that they can’t all know that they all know that they all know...that Maine has been singled out, for arbitrarily many iterations of knowledge.

My aim in this subsection has been to offer some motivation for CSR over NCR by illustrating a style of explanation that it has an easier time making sense of. In the next subsection, I’ll offer a different sort of motivation for CSR. I’ll argue that *some* degree of contextualism about the sort of robustness that makes for good explanation is unavoidable, for reasons having nothing to do with psychological explanation or action explanation in particular. While this minimal sort of contextualism is compatible with NCR—we could be contextualists about the nearness relation—it’s nevertheless enough to render NCR explanatorily idle, and unworthy of acceptance.²⁵

4.2 Minimal Explanatory Contextualism

There has been a great deal of debate among philosophers of science about the extent to which a theory of explanation must include “pragmatic” or “contextual” elements.²⁶ But it’s relatively uncontroversial that broadly contextual factors play *some* role in guiding our judgments about what constitutes a good explanation; the debates concern just how significant that role is, and whether theories of explanation that abstract away from

²⁵Also, there are independent reasons not to like a distance-based conception of safety in a contextualist framework. See Blome-Tillmann (2009).

²⁶See (Woodward, 2014, §6) for an overview.

contextual elements are thereby inadequate. I'll argue, however, that even the minimal role for contextual factors in explanation admitted by virtually all parties in debates about scientific explanation is enough to open the door to CSR, once we think about explanations of successful action in particular.

Suppose a barn is struck by lightning, and catches fire.²⁷ In most contexts, it would be natural to explain the barn's catching fire by reference to the lightning, and only the lightning—other factors, such as the presence of oxygen and the absence of water, would be naturally relegated to the background. But in certain circumstances, an explanation that failed to foreground these factors would seem inadequate. If the audience for the explanation comes from a place where it rains almost every day and structures hit by lightning therefore rarely catch fire, some further explanation might be necessary—we might have to point out that not only was the barn struck by lightning, but it was also dry, since there had been no rain for at least a week. In slightly odder circumstances, omitting to mention the absence of water would be fine, but omitting to mention the presence of oxygen would not. Imagine someone who's grown up on a planet where there is no rain, and no oxygen. For such an audience, a satisfying explanation would have to discuss the presence of oxygen, but could omit to expand on the absence of rain.

These observations are compatible with causal theories of explanation, on which explaining an event essentially involves providing information about its causes.²⁸ Such theories—at least if they hold that the basic causal facts don't depend on context—have a non-pragmatic, non-contextual core; they hold that explanatory relevance is causal relevance.²⁹ Nevertheless, such theories can allow that *which* information about an event's causal history it is appropriate to provide in response to a request for explanation will depend on context, and in particular on which causal factors will be naturally taken for granted by one's audience, and which ones will not.

²⁷The example, and my discussion of it, is indebted to Paul and Hall (2013).

²⁸I have in mind mainly Lewis (1986), though Salmon (1984) defends a related theory.

²⁹In this respect, they contrast with the theories of van Fraassen (1980) and Achinstein (1983), who seem to hold that there is no non-pragmatic, non-contextual core to the concept of explanation.

Even this minimal sort of contextualism about explanation, however, is enough to motivate a more controversial sort of epistemological contextualism, via the links discussed in the earlier part of this essay. To see this, it may help to first connect this minimal sort of contextualism about explanation to contextualism about the sort of robustness required for satisfying explanation.

If we consider examples like the above, but where the target explanandum is not a barn's catching fire but instead some bit of practical success, we'll be led into the CSR conception of safety, and the contextualism about knowledge that it naturally accompanies.

Suppose Jayden is baking muffins. After tasting the first batch and finding himself unsatisfied, he realizes that he forgot to add salt. He decides to bake another batch, this time with salt. He does so, and it turns out perfectly—he succeeds at making delicious muffins. How might we explain this success? There are lots of ways he could've failed, and a satisfying explanation will show how he avoided at least some of them. But depending on our context, it may make sense to background some, and foreground others. If we take for granted that Jayden is baking in a kitchen he's familiar with, it won't seem necessary to discuss his ability to identify the salt; an explanation that makes clear how he knew it was salt that was necessary would be sufficient. But in another context—one in which it's taken for granted that salt was the missing ingredient and that Jayden would've known this—a satisfying explanation might focus instead on Jayden's ability to recognize and retrieve the salt.

This structure should be familiar from the earlier case. When we are explaining an event's occurrence, there will be many causal factors responsible for the event, and context will determine which ones it makes sense to focus on in explanation, and which ones we can relegate to the background. When we are explaining practical success in particular, our explanation may take the form of showing how failure was avoided. But just as there are many causal factors responsible for any given event, there are many

possible ways any given successful action might instead have failed. And just as context will determine which causal factors it makes sense to focus on in explanation, similarly, context will determine which sorts of failure an explanation of success must rule out. Often, these failure possibilities will correspond to ways in which a true belief on which the action was based might instead have been false. E.g., while Jayden truly believes that adding the contents of such-and-such container to the batter will improve the taste of the resulting muffins, there are various ways this belief might have turned out to be false, leading the actions based on it to be unsuccessful. For instance, it might have been that baking soda, rather than salt, was the missing ingredient. Or it might have been that the container held sugar, rather than salt.

This already gets us pretty far along the road to a contextualist conception of safety. Here's why. Suppose we adopt NCR, and we hold that the relevant notion of nearness is context invariant. Then, regardless of our context, either the worlds where sugar is the missing ingredient, or worlds where the salt is mislabeled, will be closer to the actual world.³⁰ If the former, then there can't be any context where a good explanation focuses on Jayden's ability to identify and retrieve salt, but not on his ability to recognize that it's salt that's missing; such an explanation wouldn't reveal his true belief about which container to add from to be genuinely safe, since it wouldn't show how he would have avoided error even if sugar were the missing ingredient. And if the latter, then there can't be any context where a good explanation focuses on Jayden's palette, but not on his ability to identify and retrieve salt, since such an explanation wouldn't show how he'd have avoided error even if the salt were mislabeled. If we have a context invariant notion of nearness that figures in our notion of safety, then there will have to be some context invariant fact about whether a good explanation of Jayden's success will focus on his ability to recognize that it's salt that's missing, or instead focus on

³⁰Or such worlds are equally close to the actual world. Admitting this possibility wouldn't much change the argument—in that case, there would be no contexts in which a good explanation could background either ability.

his ability to recognize and retrieve salt. And since our minimal contextualist starting point, motivated by the barn example, was that we *didn't* want these sorts of facts to be context invariant—we wanted some way to allow that which factors a good explanation will focus on will depend on context—we need a context sensitive notion of safety.

But this means that we can't think of a safe belief in a given case as a belief that is true in some range of cases near to that case, for any context-invariant notion of nearness; if we are to think of safety in terms of nearness, then it must be a context-sensitive notion of nearness. And it's not just that the distance threshold for nearness must be context-sensitive—rather, the *ordering* of cases by nearness must be context-sensitive, if examples like the above are to be accommodated. Call the world of the Jayden example *w*. By the lights of the contexts in which good explanations must foreground Jayden's palette, worlds in which sugar is necessary, rather than salt, must be nearer to *w* than worlds in which the salt is in the container labeled "sugar." By the lights of the other sort of context, this ordering must be reversed—the latter sort of worlds must be nearer to *w* than the former sort.³¹

One potentially surprising consequence of this sort of contextualism about safety is that, depending on what we're willing to take for granted, one and the same successful action might seem explicable, or not. E.g., suppose Jayden has an excellent palette, and is very good at identifying what needs to be added to a recipe in order to improve it. However, he's not always very careful about checking the contents of jars—if jars are mislabeled, he won't notice, and will go ahead and add the wrong ingredient. If we're in a context in which misidentifying the relevant ingredient—thinking it's more baking soda that's necessary, or more sugar—seems like the important or relevant failure possibility, and we're not concerned with other sorts of error, then we'll regard Jayden's success as explicable—as no fluke. Alternatively, if we're in a context in which mislocating or otherwise failing to retrieve that ingredient—mistaking the salt container for the sugar

³¹This sort of contextualism about similarity or nearness orderings on worlds is not unfamiliar. Consider, e.g. the treatment of "Caesar Counterfactuals" in Lewis (1973, pp.66-7).

container, for example—seems like the important or relevant failure possibility, then we’ll regard Jayden’s success as lucky. Moreover, given the links discussed in the first part of this essay, that will naturally go along with thinking of Jayden’s belief as safe knowledge when we’re in the first context, and as an unsafe, mere true belief when we’re in the second. That is, in the first context, the relevant class of cases throughout which Jayden’s belief must be true if it is to count as safe, and knowledge, will include cases in which a different ingredient is missing, but not cases in which the salt jar is mislabeled. In the second context, the relevant class of cases throughout which Jayden’s belief must be true if it is to count as safe knowledge will include only cases in which salt is necessary, but in some of those cases the salt will be in the jar labeled “sugar”.

To be clear, the same sort of structure will show up in cases not involving action. Suppose there’s nothing satisfying to be said about why the barn was dry, but there is something satisfying to be said about why oxygen was present. Then the barn’s catching fire after being hit by lightning will admit of a good explanation by the standards of some explanatory contexts, but not by the standards of others.

If everything I’ve said in this subsection so far is right, then a contextualist conception of safety and robustness, and a correspondingly contextualist conception of knowledge, is almost unavoidable, at least given the package of views discussed in §1. To be clear, that doesn’t rule out adopting nearness-based conception of safety as well—we might hold that context determines a nearness ordering on worlds, and that a belief counts as safe in a world w by the standards of a context c just in case the subject avoids error in all cases close to w by the standards of c . That is, given CSR, we might adopt NCR as an add-on. Rather than thinking of context as determining a single set of worlds in which a belief must be true if it is to count as safe, we could understand context as doing all that, *and* determining another set of worlds in which a belief must be true if it is to be safely safe, and *another*, still larger set which would guarantee safe safe safety, and so on. So what does it matter whether we adopt only CSR, or CSR *and*

NCR? As discussed in §3, one of the main ways in which NCR has been applied in epistemology has concerned iterations of safety, and iterations of knowledge. We also saw that CSR on its own will not vindicate those applications. But if the arguments of this section are right, then CSR is a kind of baseline that all minimal explanatory contextualists—a relatively uncontroversial position in the philosophy of science—must accept. This robs safety-based arguments against iteration principles for knowledge of much of their force. That is, if the claims in this section are right, then the baseline conception of safety is one that is congenial to iteration principles for knowledge like KK. While we *could* adopt a richer conception of safety that would not be congenial to such principles, such a richer conception would have to be independently motivated. Rather than facing a choice between a contextualist conception of safety that's congenial to KK, and a non-contextualist conception that isn't, we'd face a choice between two contextualist conceptions of safety, the simpler of which is congenial to KK, and the more complicated of which is not. And that leaves the defender of KK in a much better dialectical position than she's typically viewed as being, at least as regards the relationship between knowledge and safety.

5 Conclusions

As this essay has touched on a number of different debates, it may help to take a step back and provide a bird's eye view of the key points. First, I argued that a certain popular combination of views in epistemology—the view that safety is a requirement on knowledge, and the view that knowledge plays a key role in explaining practical success—forms a coherent package with an independently well motivated view in the philosophy of science, to the effect that robustness is an explanatory virtue. This suggested a methodological lesson; we can let our thinking about explanation and robustness more generally guide our views about how to understand safety requirements on knowledge in partic-

ular. So far, this can all be accepted by orthodox safety theorists in epistemology. In *applying* this methodological lesson, however, I made my way to unorthodox results. In particular, skipping ahead to §4.2, the methodological lesson suggested a contextualist moral—a relatively minimal form of contextualism about scientific explanation in general, when transposed into an epistemological key, motivates a more controversial form of contextualism about the sort of safety relevant to knowledge. And, as I argued in §3, if we are contextualists about safety, some of the main epistemological applications of the safety-theoretic framework—ones that crucially depend on understanding safety in terms of (metaphorical) *nearness*—don’t follow. This is not to say that those applications *couldn’t* ultimately be vindicated. But doing so would require argumentative work that hasn’t yet been done; it would require engaging alternative conceptions of safety that don’t support the applications, and showing them to be inferior. And as I argued in §4.1, the contextualist conception of safety I discuss has the advantage of conservativeness; it can vindicate a class of popular explanations about which the nearness-based conception must be revisionist. Ultimately, the upshot is that while safety requirements on knowledge are well-motivated, some of their most influential applications are not.

Bibliography

- Aarnio, Maria Lasonen. 2010. “Unreasonable Knowledge.” *Philosophical Perspectives* 24:1–21.
- Achinstein, Peter, 1983. *The Nature of Explanation*. Oxford University Press.
- Blome-Tillmann, Michael. 2009. “Contextualism, Safety and Epistemic Relevance.” *Philosophical Studies* 143:383–394.
- Dennett, Daniel, 1981. “Intentional systems.” In Haugeland, John, editor, “Mind Design,” Bradford Books.

- Dretske, Fred, 1988. *Explaining Behavior*. MIT Press.
- Gibbons, John. 2001. “Knowledge in Action.” *Philosophy and Phenomenological Research* 62(3):579–600.
- Greco, Daniel. 2014. “Could KK Be OK?” *Journal of Philosophy* 111:169–197.
- . 2015. “Iteration Principles in Epistemology I: Arguments For.” *Philosophy Compass* .
- Heal, Jane. 1978. “Common Knowledge.” *Philosophical Quarterly* 28:116–131.
- Holliday, Wesley H. 2015. “Epistemic Closure and Epistemic Logic I: Relevant Alternatives and Subjunctivism.” *Journal of Philosophical Logic* 44:1–62.
- Lederman, Harvey. 2015. “People with Common Priors Can Agree to Disagree.” *Review of Symbolic Logic* pages 1–35.
- . Ms. “Uncommon Knowledge.” .
- Lewis, David, 1973. *Counterfactuals*. Oxford University Press.
- . 1974. “Radical Interpretation.” *Synthese* 27:331–344.
- , 1986. “Causal Explanation.” In Lewis, David, editor, “Philosophical Papers Vol. II,” pages 214–240. Oxford University Press.
- , 1999. “Elusive Knowledge.” In “Papers in Metaphysics and Epistemology,” Cambridge University Press.
- Millikan, Ruth G., 1984. *Language, Thought and Other Biological Categories*. MIT Press.
- Nagel, Jennifer. 2013. “Knowledge as a Mental State.” *Oxford Studies in Epistemology* 4:275–310.

- Paul, LA, and Ned Hall, 2013. *Causation: A User's Guide*. Oxford.
- Pettit, Frank Jackson and Philip, 1993. "Some Content is Narrow." In Heil, John, and Alfred Mele, editors, "Mental Causation," Oxford University Press.
- Pritchard, Duncan. 2009. "Safety-Based Epistemology." *Journal of Philosophical Research* 34:33–45.
- Salmon, Wesley, 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton University Press.
- Salow, Bernhard. forthcoming. "Lewis on Iterated Knowledge." *Philosophical Studies* pages 1–20.
- Sosa, Ernest. 1999. "How to Defeat Opposition to Moore." *Philosophical Perspectives* pages 141–153.
- Stalnaker, Robert, 1984. *Inquiry*. Cambridge, MA: The MIT Press.
- Stanley, J., and Timothy Williamson. 1995. "Quantifiers and Context Dependence." *Analysis* 55:291–295.
- Strevens, Michael, 2008. *Depth: An Account of Scientific Explanation*. Harvard University Press.
- van Fraassen, Bas, 1980. *The Scientific Image*. Oxford: Oxford University Press.
- Williamson, Timothy, 2000. *Knowledge and its Limits*. Oxford University Press.
- . 2001. "Comments on Michael Williams' Contextualism, Externalism and Epistemic Standards." *Philosophical Studies* 103:25 – 33.
- Williamson, Timothy. 2014. "Very Improbable Knowing." *Erkenntnis* 79:971–999.

Woodward, James. 2000. "Explanation and Invariance in the Special Sciences." *British Journal for the Philosophy of Science* 51:197–254.

———, 2014. "Scientific Explanation." In "Stanford Encyclopedia of Philosophy," <http://plato.stanford.edu/entries/scientific-explanation/>.