

A Theory of Epistemic Feedback Loops (Or: How Not To Get Evidence)*

Nick Hughes
University of Oxford
nickhowellhughes@gmail.com

§0. Introduction

Epistemologists spend a great deal of time thinking about how we should respond to our evidence. They spend far less time thinking about the ways that evidence can be acquired in the first place.¹ This is an unfortunate oversight. Some ways of acquiring evidence are better than others. Many normative epistemologies cannot accommodate this fact. This article attempts to develop one that can and does.

Section one explains why evidence acquisition matters and describes a phenomenon – *epistemic feedback loops* – in which it has gone awry, with the result that even a belief based on the evidence is irrational or at least

* This is a draft. Nevertheless, feel free to cite it. Got comments or questions? Email me!

¹ Some heroic exceptions: Hall & Johnson (1998), Greco (2005), Baehr (2011), Salow (2018), Miracchi (2019), Sosa (ms), Dorst (2020), and Flores & Woodard (ms), as well as epistemologists who write about cognitive penetration, some of whom I'll discuss later.

problematic. Section two outlines five desiderata for an adequate theory of feedback loops. Sections three to eight develop a theoretical framework – ‘dispositionalism’ – with which to understand them and show how it satisfies the desiderata. Section nine draws out some consequences for normative epistemology, arguing that a wide range of theories of rational belief must be rejected as they cannot adequately explain feedback loops. Section ten addresses objections.

§I. Epistemic Feedback Loops

To see why evidence acquisition matters, consider these three cases, the first two of which are examples of confirmation bias and the third of which is an example of cognitive penetration:

DOMINEERING CEO: Camille is the powerful and intimidating CEO of a large media company. She wishes that the company was financially stable. This causes her to unconsciously put pressure on the CFO to tell her that the company *is* financially stable. As a result of the CFO’s testimony, she forms the belief that matches her desire.

LIKEABLE LEVI: Levi fears that his Sunday league football teammates don’t like him. This causes him to unconsciously seek out evidence that disconfirms his fear and avoid evidence that might confirm it. Having gathered a wealth of evidence that his teammates like him and none indicating that they don’t, he comes to believe that he is liked after all.

UNHAPPY ZORA: Asher suspects that Zora is unhappy. When he bumps into her at a bar later on, this suspicion causes him to have a perceptual experience as of her face wearing an unhappy expression. As a result he comes to believe that she is unhappy. In fact her expression is neutral.²

Confirmation bias and cognitive penetration are usually theorised separately, but from the point of view of evidence acquisition they have important similarities, making a theory that explains them conjointly desirable.³

What similarities? First, in each of these cases a psychological attitude of an agent (Camille's wish, Levi's fear, Asher's suspicion) causes them to unknowingly manufacture evidence in favour of a proposition *p* (The company's finances are stable, Levi is liked, Zora is unhappy) which causes them to form a belief that *p*. Second, the resulting beliefs are irrational, or at least epistemically problematic.⁴ But the problem is not with way that Camille, Levi, and Asher *respond* to their evidence. Had the CFO's testimony been uncoerced, had Levi's investigation not been biased, and had Asher's perceptual experience not been caused by his suspicion that Zora is unhappy, their beliefs would have been fine. Rather,

² Adapted from Siegel (2012). More on Siegel's take on cases like this later.

³ Confirmation bias involves two separable components: *biased assimilation* and *selective exposure* (Lord *et al.* 1979). Biased assimilation happens when agents respond to their evidence in a way that favours a desired or pre-existing belief. Selective exposure happens when agents acquire (select, gather, etc.) evidence in a way that favours a desired or pre-existing belief. DOMINEERING CEO and LIKEABLE LEVI are cases of selective exposure rather than biased assimilation.

⁴ I'm happy to describe them as irrational, for reasons I will explain later (§10.1). But even readers who prefer to think of them as rational-but-nevertheless-problematic will be able to use the framework I'll develop to explain what's going wrong.

the problem is the way the evidence was acquired in the first place. In each case the evidence has, as I will put it, a *deviant etiology*.⁵

I'll call cases like these *epistemic feedback loops*. This article develops a framework which explains why they are epistemically problematic. Its key observation is simple: the ways that Camille, Levi, and Asher acquire their evidence could easily result in them acquiring misleading evidence. The challenge is to sharpen this observation and situate it within a systematic theory.

Epistemic feedback loops should not be of interest only to epistemologists. They can have serious consequences. Cognitive penetration may reinforce racism (Siegel 2017). Confirmation bias has been cited as a causal factor in the rise of online 'echo chambers' (Quattrocio *et al.* 2016). Feedback loops may even shape world-historical events: Jonathan Leader Maynard (fc) argues that Joseph Stalin's paranoid convictions were a primary cause of the Soviet Terror of 1936-1938. Suspecting internal conspiracies, Stalin ordered apparatchiks to root out perceived saboteurs from within the party. Under pressure, they coerced 'confessions' from the accused which were taken at face-value by the party leadership as evidence of wrongdoing. Executions swiftly followed.

⁵ Why stipulate that Camille, Asher, and Levi are *unconscious* of the mechanisms which cause them to manufacture evidence? Because if they were aware that their evidence has a deviant etiology it wouldn't be evidence in the first place. If, for example, Asher knows that it is his suspicion that Zora is unhappy that has caused him to have a perceptual experience as of her face wearing an unhappy expression, then the fact that her face appears to be wearing an unhappy expression would not be evidence that she is unhappy. See Salow (2018) for discussion.

§II. Desiderata

An adequate theory of feedback loops must satisfy a number of desiderata. First, not every feedback loop is epistemically problematic. Suppose Jake desires (perhaps unconsciously) to be naked. This desire causes him to take his clothes off, which causes him to believe that he's naked. There is nothing epistemically problematic here, even though the case has the same structure as the problematic cases: a prior psychological attitude towards *p* causes the manufacture of evidence for *p* which causes belief that *p*. Our first desideratum:

EXCEPTIONS: The theory must be able to account for the fact that some feedback loops are unproblematic and explain the differences between these and problematic feedback loops.

Second, although they all share the same structure at a certain level of abstraction, there are differences between DOMINEERING CEO, LIKEABLE LEVI, and UNHAPPY ZORA. One difference is between Camille and Asher, on the one hand, and Levi, on the other. Camille and Asher's prior attitude towards *p* leads directly to the *creation* of evidence for *p*. Whilst Levi might also directly create some of his evidence (by fishing for compliments, for instance), he also engages in a process better described as *selecting* evidence, by filtering out unwanted sources. Another difference is in the kind of evidence that is manufactured. Asher manufactures visual-perceptual evidence for *p*. Camille manufactures testimonial evidence for *p*. Levi most likely manufactures a combination of the two. A third difference is that Asher manufactures evidence that is in some sense internal to his mind, whereas Camille and Levi manufacture evidence in the external world. A fourth difference is in the prior attitudes themselves. It is Camille's *wish* that causes her to manufacture evidence, it is Levi's

fear, and it is Asher's *suspicion*. Epistemic feedback loops can take many different forms. An adequate theory must be able to accommodate all the forms they can take. This is our second desideratum:

VARIETY: An adequate theory must be able to accommodate the various forms that feedback loops can take.

Third, the phenomenon is not limited to sophisticated, reflective, agents like Camille, Levi, and Asher. Consider:

WOLSEY: Wolsey the dog loves sausages and hopes to find one. This causes him to perceptually experience a stick as a sausage, which causes him to believe that it's a sausage.

Wolsey's belief is epistemically problematic in much the same way as Asher's (though we might hesitate to call it 'irrational' – more on this shortly). An adequate theory should be able to explain why. Hence, our third desideratum:

INCLUSIVITY: The theory must be able to explain why feedback loops are epistemically problematic even when they arise in unsophisticated agents.

Fourth, loop cases (as I'll call them) are importantly different to bad cases of the kind epistemologists are familiar with – cases in which an agent is envatted, deceived by an evil demon, lied to by an apparently credible person, tricked by a hidden light source into believing that a white table is red, and so on. Although the beliefs of agents in these cases are false, they are rationally unimpeachable, unlike Camille, Levi, and Asher's beliefs. This gives us a fourth desideratum:

NUANCE: The theory must be able to explain the differences between loop cases and familiar bad cases.

Fifth, although beliefs resulting from feedback loops are problematic, they are liable to give rise to ambivalence. Camille, Levi, and Asher aren't comfortably lumped in with agents who fail to *respond* correctly to their evidence. Compare them with X, a patient suffering from Othello Syndrome described in a case study by Pal *et al.* (2012). Amongst other things, X believed that pictures of fruit on his wife's social media page indicated that she was having an affair. Although Camille, Levi, and Asher acquire their evidence in a way that makes their beliefs problematic, they are, unlike X, doing well in at least one way; *given* the evidence they have, it would be surprising if they didn't believe what they do. Indeed, were they to suspend judgement on p or disbelieve that p with this evidence in hand, they would seem to exhibit another kind of rational failing. Our theory should explain why. This gives us a fifth desideratum:

AMBIVALENCE: The theory should explain why beliefs produced by feedback loops are okay on a certain dimension of evaluation.

We have five desiderata: EXCEPTIONS, VARIETY, INCLUSIVITY, NUANCE and AMBIVALENCE. Is there a theory that satisfies all of them? I think there is. In the next sections, I'll present and defend it.

§III. Functional and Dysfunctional Cognition

§3.1. Setting the Stages

My starting assumption is that a core function of cognition is to produce knowledge which the agent can use to guide action. I'll develop a simple theoretical framework to explain feedback loops by reference to cognition working well or badly relative to this purpose with respect to a given proposition p .⁶ The framework builds on ideas from Maria Lasonen-Aarnio (2010, fc1, fc2) and Timothy Williamson (fc1, fc2). I call it 'dispositionalism'. Since it places knowledge front and centre, it can be thought of as part of the knowledge-first programme kicked off by Williamson (2000).

The framework makes use of two distinctions.

The first is between cognition working well or badly *locally* and cognition working well or badly *broadly*. Local evaluations focus on the results cognition produces in the world in which it takes place. Broad evaluations focus on the results it produces in a range of counterfactual worlds.

The second is between cognition working well or badly when the agent *acquires* evidence and cognition working well or badly when the agent *responds* to their evidence. I'll call these the *acquisition stage* and the *response stage* respectively.

I'll say that cognition is *functional* iff it is working well and *dysfunctional* iff it is working badly.

Since the framework aims to explain commonalities between Camille, Levi, and Asher and less sophisticated agents like Wolsey the dog it deliberately avoids employing the anthropocentric language, familiar to

⁶ It's possible to develop a version of the framework on the starting assumption that a core function of cognition is merely to produce *true beliefs* (rather than knowledge), though I will not do that here.

epistemologists, of justifications, excuses, responsibility, reasons, rationality, reasonableness, virtue, oughts, duties, requirements, praise, blame, and so on. Such concepts are only comfortably applied to sophisticated, self-conscious, reflective, adult humans. Later we will see how the framework can be used to shed new light on some of them.

I'll make some simplifying assumptions. Firstly, for every p , an agent either believes that p , suspends on p , or disbelieves that p . Secondly, an agent believes that p iff they don't suspend on p or disbelieve that p , disbelieves that p iff they don't suspend on p or believe that p , and suspends on p iff they don't believe that p or disbelieve that p . Thirdly, an agent disbelieves that p iff they believe that not- p . The first two assumptions are arguably unrealistic. It might be possible to take no attitude at all towards a proposition. It might be possible to simultaneously take multiple attitudes towards a single proposition (for example, under different modes of presentation). Suspension might be an umbrella concept covering multiple distinct attitudes (Lord & Sylvan *fc*). I'll ignore these complications because nothing turns on them. Accommodating them in the framework would involve needlessly bringing in distracting complexities.

§IV. The Acquisition Stage

Starting with the acquisition stage, the two distinctions described above are applied as follows:

- Cognition is *locally functional* with regards to p at the acquisition stage iff the agent does not acquire misleading evidence about p .

- Cognition is *locally dysfunctional* with regards to p at the acquisition stage iff the agent acquires misleading evidence about p .
- Cognition is *broadly functional* with regards to p at the acquisition stage iff the way the agent acquires evidence about p would not normally yield misleading evidence about p .
- Cognition is *broadly dysfunctional* with regards to p at the acquisition stage iff it would not be abnormal for the way the agent acquires evidence about p to yield misleading evidence about p .

This part of the framework makes use of two concepts – misleading evidence and normality – which require further comment. I'll explain them in turn.

§4.1. Misleading Evidence

§4.1.1. Two Conceptions of Misleading Evidence

What is misleading evidence? We can distinguish between (at least) two explications of the concept, both of which are perfectly legitimate, but only one of which is relevant for our purposes.

According to the first explication (the irrelevant one), a piece of evidence e is misleading with respect to p just in case p is false and the probability of p is higher conditional on e than the unconditional probability of p . Formally:

- $(P(p|e) > P(p)) \wedge \sim p$

This explication focuses on individual pieces of evidence – a single proposition known to be true, for instance. The second explication focuses on one's total body of evidence E , comprised of all of the propositions that are part of one's evidence. We are interested in whether E is misleading, not whether e is. Virtually any way of acquiring evidence about p will yield some misleading evidence e bearing on p in some normal circumstances. Seeing that everyone in the room is under two meters tall is *some* evidence e for the false proposition that everyone in the world is under two meters tall ($= p$) – it raises the probability of p , if only very slightly. But nothing abnormal need happen for one to see that everyone in the room is under two meters tall.

§4.1.2. *Delusive Evidence*

When is one's total evidence E misleading with respect to p ? A natural idea is that E is misleading with respect to p iff E leads an agent into believing that p when not- p . Which agent? The answer cannot simply be the agent who possesses E . The presence of images of fruit on his wife's social media page lead X to believe that she was having an affair. It does not follow that his evidence was misleading with respect to p ($= X$'s wife is having an affair); the problem was with him, not with his evidence. A more plausible idea is that E is misleading with respect to p iff a *rational* response to E yields a false belief that p .

This is close to what I have in mind, but it isn't *quite* right for our purposes. It would make sense if we took the function of cognition to be to produce true beliefs. But my starting assumption is that its function is to produce *knowledge*, not merely true beliefs. If so, a non-knowledgeable belief is as bad as a false belief. So it's more natural within the current framework to say that E is misleading with respect to p iff a rational response to E yields a *non-knowledgeable* belief that p (this could be a false belief, of course). In

section §5.2 we will see how this idea can be de-anthropocentrised and made more precise – a rational response is a *broadly functional* response.

This way of conceptualising misleading evidence is unconventional insofar as it delivers the result that, for instance, an agent in a Gettier case has misleading evidence. ‘Misleading evidence’ is usually taken to mean: misleading with respect to the *truth-value* of p . But we will think of it as meaning: misleading with respect to whether one *should believe* that p . If beliefs should be knowledge, then evidence that will lead to non-knowledgeable beliefs is evidence that should be avoided.

Since this is an unconventional way of thinking about misleading evidence, it will be useful to have a name for it, in order to avoid confusion with more traditional truth-centric ways of thinking. So from here on I’ll call it *delusive evidence*:

DELUSIVE EVIDENCE: One’s total evidence E is delusive with respect to p if and only if a broadly functional (i.e. rational) response to E yields a non-knowledgeable belief that p .⁷

In line with this, the conditions on functional and dysfunctional acquisition described above can be restated as follows:

⁷ To be clear, I’m not saying that no other kind of evidence deserves to be called ‘misleading’. [omitted] suggested this case to me. “Suppose there are 10 balls in an urn, nine blue and one red. You know there are 10 balls and that they are all blue or red, but you don’t know the proportion. You make, say, 15 draws with replacement, and each time the red ball comes out. This seems to be insufficient evidence to warrant forming an outright belief about the proportion of blue and red balls, yet it is natural to say that you have misleading evidence about the proportion”. I agree. It’s just that this isn’t a concept of misleading evidence that my framework makes use of. It may well be useful for other purposes.

- Cognition is locally functional with regards to p at the acquisition stage iff the agent does not acquire delusive evidence about p.
- Cognition is locally dysfunctional with regards to p at the acquisition stage iff the agent acquires delusive evidence about p.
- Cognition is broadly functional with regards to p at the acquisition stage iff the way the agent acquires evidence about p would not normally yield delusive evidence about p.
- Cognition is broadly dysfunctional with regards to p at the acquisition stage iff it would not be abnormal for the way the agent acquires evidence about p to yield delusive evidence about p.

§4.2. Normality

The framework also makes use of the concept of normality. What does it mean to say that a way of acquiring evidence would not normally yield delusive evidence, or that it would not be abnormal for it to yield delusive evidence?

§4.2.1. *Normal Circumstances*

‘Normally’ is being used here as a restricted necessity operator. To say that a way of acquiring evidence would not normally yield delusive evidence is to say that in *normal circumstances* it does not yield delusive evidence. To say that it would not be abnormal for a way of acquiring evidence to yield delusive evidence is to say that even in some normal circumstances it yields delusive evidence.

Normality and abnormality are indexed to *ways*. So to say that a way *w* of acquiring evidence would not yield delusive evidence in normal circumstances is to say that *w* would not yield delusive evidence *in circumstances normal for w*.

What are normal circumstances? The question is best answered ostensively, by looking at examples.

§4.2.2. Good/Bad Case Pairs

Consider a good/bad case pair of the familiar kind. In the good case Annie is walking down rue d'Aboukir in Paris, sees that number 44 is a typical Parisian apartment building, and so believes, correctly, that *p* (= 44 is an apartment building). Cognition is locally functional – Annie does not acquire delusive evidence about *p*. Filling in the details in the natural way, it is also broadly functional. The way Annie acquires her evidence – roughly, through a disinterested openness to receiving information from her environment through visual perception⁸ – will yield delusive evidence about *p* only in *abnormal* circumstances, such as those in which 44 is actually, say, a giant ventilation shaft hidden behind an ingenious façade giving it the appearance of being a typical Parisian apartment building.

In the bad case the set-up is exactly the same except that 44 is a giant ventilation shaft hidden behind an ingenious façade giving it the appearance of being a typical Parisian apartment building.⁹ As before, Annie believes that *p*, only this time falsely and hence non-knowledgeably. Cognition is locally dysfunctional, as Annie acquires delusive evidence. But it is still *broadly functional*. Although Annie acquires delusive evidence, the way she acquires it would not normally yield

⁸ Disinterested as in impartial, not *uninterested* as in indifferent.

⁹ As it in fact is. Have a look on Google Earth.

delusive evidence. It only does so because she is in abnormal circumstances.

What goes for Annie in the bad case also goes for more radical bad cases in which the agent is not in touch with their environment, such as those involving envattment or the machinations of evil epistemic demons. Suppose that you believe on the basis of your perceptual experience that LeBron James has just hit a game-winning three-pointer. Alas, you're wrong. You're not sitting courtside at Staples Center, you've been envatted. Even radical cases like this are set up on the assumption that the agent acquires their evidence in a disinterested and open way, so that were they reconnected with the environment they would not acquire delusive evidence. Indeed, it is precisely this feature of the agent's cognition that the evil demon or mad scientist exploits as part of their strategy to get the agent to believe falsehoods.

§4.2.3. Loop Cases

By contrast, the situation is quite different in problematic loop cases. Here cognition is *broadly dysfunctional*. Unlike Annie, Camille, Levi, and Asher are *not* disinterestedly open to receiving information from the environment. Instead they manufacture evidence that harmonises with a prior psychological state. It wouldn't be at all abnormal for the ways they acquire their evidence to yield delusive evidence – most obviously by yielding evidence that will cause them to falsely believe that p. Nothing weird or out-of-the-ordinary would have to happen for this state of affairs to obtain. The reason why is clear: in virtue of being biased towards harmony with a prior psychological state, their ways of acquiring evidence are *indifferent* and *insensitive* to their environments.

§4.2.4. Ways of Acquiring Evidence

I've described two ways of acquiring evidence. The first is through a disinterested openness to one's environment. The second is by manufacturing evidence that harmonises with a prior psychological state. This individuation of ways of acquiring evidence is obviously very coarse-grained. More fine-grained individuations are possible. We could, for instance, distinguish between ways of acquiring evidence that yield visual-perceptual evidence and ways that yield testimonial evidence. We could also distinguish between ways of acquiring evidence that fall under the umbrella description of being disinterestedly open to one's environment. Equally clearly, this is not a complete taxonomy of ways of acquiring evidence. Nevertheless, we have enough here for the purposes of explaining feedback loops, so there is no need to go more fine-grained.

§4.2.5. *What Are Normal Circumstances?*

Can we say something general about normal and abnormal circumstances, beyond pointing to intuitive examples? This turns out to be tricky. It may be that normality cannot be defined in terms of more basic concepts. Nevertheless, some things can be said. What counts as normal for our purposes is not a matter of statistical frequency. It would not be abnormal for this week's winning lottery numbers to be [04, 10, 18, 19, 45, 36] even though these numbers will win only very infrequently.¹⁰ It would also be a mistake to take normal circumstances for performing a task (in our case, acquiring evidence) to be circumstances that are *favourable* to the successful execution of that task (e.g. acquiring non-delusive evidence).¹¹ One reason for this is because doing so would give us the implausible result that cognition is broadly functional *no matter how* one acquires one's evidence. Another is because some favourable circumstances are intuitively highly abnormal. For instance, favourable circumstances for

¹⁰ C.f. Smith (2010, 2016)

¹¹ C.f. Beddor & Pavese (2020).

acquiring evidence in the way that Levi does include those in which a guardian angel changes the world to fit with his evidence.

Bob Beddor and Carlotta Pavese (2020) suggest that normal circumstances for performing a task are, roughly, circumstances in which it would be *fair to evaluate* the performance.¹² This seems to fit with what we've seen so far. Intuitively, it is not fair to negatively evaluate Annie's way of acquiring evidence based on how it performs in the bad case. She was unlucky. By contrast, it *is* fair to negatively evaluate Camille, Levi, and Asher's ways of acquiring evidence based on how they perform in cases in which p is false.

Martin Smith (2010, 2016) argues that normal circumstances are those that *require no special explanation*. Adapting this idea for our purposes, we might say that normal circumstances for way w of acquiring evidence to have some outcome o (e.g. yielding delusive evidence) are circumstances in which w resulting in o doesn't require a special explanation, and that abnormal circumstances for w resulting in o are those in which w resulting in o does require a special explanation. Again, this seems to fit with what we've seen. Were Annie to end up with delusive evidence, given the way she acquires it, this would cry out for an explanation – for instance, that she was taken in by an ingenious façade. By contrast, no special explanation is called for in order for us to understand what's happened when Camille, Levi, and Asher end up with delusive evidence, given how they acquire it.¹³

¹² Beddor and Pavese wisely refrain from proposing this as a non-circular analysis of normal circumstances; it may be that judgements about whether it is fair to evaluate a performance in such-and-such circumstances are guided by tacit judgements about whether those circumstances are normal or abnormal for performances of the relevant kind.

¹³ Again, we need not view this as an attempt to offer a non-circular analysis.

I won't take a stance on these proposals here. Nor will I offer my own. Instead I propose to treat normality as primitive and let judgements about cases be our guide.^{14,15}

§5. The Response Stage

§5.1. Local Response

Now for the response stage. In keeping with the assumption that a core function of cognition is to produce knowledge, cognition is locally functional with regards to p at the response stage iff:

1. The agent knows that p , or:
2. The agent knows that not- p , or:
3. The agent suspends on p and the agent is not in a position to know that p and not in a position to know that not- p .

Cognition is locally dysfunctional with regards to p at the response stage iff:

1. The agent believes that p without knowing that p , or:

¹⁴ For other ways of thinking about normality, see Millikan (1984), Nickel (2016), and Graham (2017), amongst others.

¹⁵ An objection I've heard is that even the very best ways of acquiring evidence yield delusive evidence in *some* normal circumstances, and so that the claim that cognition is broadly functional at the acquisition stage iff the way the agent acquires evidence yields delusive evidence *only* in abnormal circumstances is too strong. I'll address this objection in §10.4

2. The agent disbelieves that p without knowing that not-p, or:
3. The agent suspends on p and the agent is in a position to know that p or is in a position to know that not-p.

§5.1.1. *Being In a Position to Know*

This part of the framework makes use of the concept of ‘being in a position to know’. When is an agent in a position to know that p? The question to ask is: ‘If the agent believed that p at *t*, would they know that p at *t*?’. If the answer is ‘yes’ they are in a position to know that p at *t*. If the answer is ‘no’, they are not.¹⁶ For illustration, take the proposition ‘Angela Merkel’s father was not over 5 meters tall’. Presumably you believe it, and rightly so: that belief is knowledge. By contrast, take the proposition ‘There were 2991 Tigers in India at noon on July 7th 2020’. Presumably you suspend judgement on it, and rightly so: you’re not in a position to know if it’s true.

Cognition would have been locally dysfunctional if you’d believed that there were 2991 Tigers in India at noon on July 7th 2020, since you are not in a position to know whether it is true or false. It would also have been locally dysfunctional if you’d suspended judgement on whether Angela Merkel’s father was over 5 meters tall, since you are in a position to know that he wasn’t. A consequence is that suspension is not a safe harbour which is always epistemically okay, as it is sometimes thought. I think this is the right result. If the function of cognition is to produce knowledge

¹⁶ The conditional ‘if S believed that p, they would know that p’ is a useful heuristic for the purposes of latching on to the intended interpretation of ‘in a position to know’, but it is fallible. We can contrive situations in which S is, intuitively, in a position to know that p, but were S to believe that p, p would be rendered false and hence not knowable.

which the agent can use to guide action, it is not working well if it is resistant to knowledge.¹⁷

§5.2. Broad Response

As the biconditionals above indicate, whether cognition is locally functional or dysfunctional at the response stage is a matter of the results it produces in the world in which cognition takes place. Whether it is *broadly* functional or dysfunctional at the response stage depends, as it does at the acquisition stage, on the results it produces in a range of counterfactual worlds. We focus on the results it produces when the evidence is acquired in a *non-deviant* way.

Cognition is broadly functional with regards to p at the response stage iff:

1. The agent believes that p and the way the agent responds to the evidence for p manifests a disposition to believe that p iff they are in a position to know that p, or:
2. The agent disbelieves that p and the way the agent responds to the evidence for p manifests a disposition to disbelieve that p iff they are in a position to know that not-p, or:
3. The agent suspends on p and the way the agent responds to the evidence for p manifests a disposition to suspend on p iff they are

¹⁷ C.f. Williamson (fc2). A worry I've heard is that the claim that cognition is dysfunctional if the agent suspends on p whilst being in a position to know that p is unrealistically demanding. There is an enormous number of trivial things one is in a position to know, but it would be foolish to spend all of one's time forming beliefs about them (c.f. Nelson 2010, Littlejohn 2013). I'll address this concern in §10.5

not in a position to know that p and not in a position to know that not-p.

Cognition is broadly dysfunctional with regards to p at the response stage iff:

1. The agent believes that p and the way the agent responds to the evidence for p does not manifest a disposition to believe that p iff they are in a position to know that p, or:
2. The agent disbelieves that p and the way the agent responds to the evidence for p does not manifest a disposition to disbelieve that p iff they are in a position to know that not-p, or:
3. The agent suspends on p and the way the agent responds to the evidence for p does not manifest a disposition to suspend on p iff they are not in a position to know that p and not in a position to know that not-p.

§5.2.1. Knowledge-Conducive/Resistant Dispositions

The idea that doxastic states should be evaluated by looking at the dispositions they manifest relative to knowledge is developed in detail by Lasonen-Aarnio (2010, fc1, fc2). Following Lasonen-Aarnio, I'll call the functional dispositions 'knowledge-conducive'. I'll call the dysfunctional dispositions 'knowledge-resistant'.

What does it mean to say that an agent manifests a disposition to believe that p iff they are in a position to know that p? A plausible answer takes its lead from our definition of broadly functional acquisition, in which evidence acquisition is broadly functional just in case the way the agent acquires evidence does not yield delusive evidence in normal

circumstances. Following on from that idea, we can say that an agent manifests a disposition to believe that p iff they are in a position to know that p just in case *in normal circumstances* the way they respond to their evidence results in them believing that p iff they are in a position to know that p .¹⁸ Again, it will be helpful to look at some examples, starting with variations on the good and bad cases.

§5.2.2. *Knowledge-Conducive Dispositions In the Bad Case*

So, think again about Annie in the bad case, walking down rue d'Aboukir. She believes that 44 is an apartment building ($= p$). In fact it is an ingenious façade hiding a ventilation shaft. Although cognition is locally dysfunctional since her belief is not knowledge, on the natural way of filling in the details of the case it is broadly functional: the way Annie responds to her evidence would have resulted in her knowing that p in normal circumstances. The only reason she believes without knowing is because she is in abnormal circumstances. Hence, she manifests a knowledge-conducive disposition. The same goes in more radical bad cases. You believe on the basis of your perceptual experience that LeBron James has just made a game-winning three-pointer, but you're wrong – you've been envatted. You believe that p without knowing that p , but the way you respond to your evidence would have resulted in you knowing that p in normal circumstances. Hence, you manifest a knowledge-conducive disposition.

§5.2.3. *Knowledge-Resistant Dispositions In the Good and Bad Case*

Now imagine that in the bad case Annie instead refuses to believe her eyes and inexplicably suspends judgement on whether 44 is an apartment building ($= p$). Cognition is locally functional, since she's not in a position

¹⁸ C.f. Lasonen-Aarnio (2010, fc1, fc2)

to know that p and not in a position to know that not- p . However, it is broadly *dysfunctional*: in normal circumstances the disposition Annie manifests would lead to her suspending even when she is in a position to know. Hence, she is *knowledge-resistant*. The same goes for you if you suspend judgement on whether LeBron just made a game-winning three-pointer.

Next, consider the good case again, in which 44 is an apartment building ($= p$). Annie believes that p and this belief is knowledge. Now suppose that a prankster resident erroneously tells her that it is a façade hiding a ventilation shaft. Annie refuses to believe him, despite having no reason to think that he is lying. Again, cognition is broadly dysfunctional, as the way Annie responds to her evidence for p does not manifest a disposition to believe that p iff she is in a position to know that p : it would not be abnormal for Annie's dogmatism to cause her to believe that p even when she doesn't know that p .¹⁹ She is knowledge-resistant.

§5.2.4. Knowledge-Conducive/Resistant Dispositions In Loop Cases

We saw that Camille, Levi, and Asher's cognition is broadly *dysfunctional* at the acquisition stage. How do they fare at the response stage? Here cognition is broadly *functional* when they believe that p . Recall, Camille has been told by the CFO that the company is financially stable, Levi has received lots of compliments and no criticism, and Asher has perceptually experienced Zora's face as of wearing an unhappy expression. It's no surprise that they believe what they do, given this evidence. As noted earlier, the problem is not with how they respond to their evidence, but how it was acquired. Whilst there are normal circumstances in which they believe that p without knowing that p , owing to the deviant etiology of their evidence, the way they respond to their evidence manifests a

¹⁹ C.f. Lasonen-Aarnio (fc2)

disposition to believe that p iff they are in a position to know that p *when the evidence is acquired in a non-deviant way*: were the evidence *not* to have a deviant etiology, the way they respond to it would normally result in them knowing that p . By contrast, cognition would be broadly *dysfunctional* at the response stage were they to suspend judgement on p . Like Annie when she inexplicably suspends judgement on whether 44 is an apartment building, they would manifest a disposition to suspend even when they are in a position to know. Hence, they would be knowledge-resistant.

§VI. Applying the Framework

§6.1. Good/Bad Case Pairs

This framework sheds light on epistemic feedback loops. Before we see exactly what it says about them, it will be useful to summarise what it says about good/bad case pairs in order to highlight the similarities and differences between these and loop cases.

The good and bad cases are represented in table 1. The belief rows tell us whether cognition is functional or dysfunctional when the agent believes that p , the suspension rows tell us whether it is functional or dysfunctional when the agent suspends on p , the disbelief rows tell us whether it is functional or dysfunctional when the agent disbelieves that p . A green cell with an ' F ' means cognition is functional. A red cell with a ' DF ' means it is dysfunctional.

Table 1	Doxastic attitude	Local acquisition	Broad acquisition	Local response	Broad response
Good case	Belief	F	F	F	F
	Suspension			DF	DF
	Disbelief			DF	DF
Bad case	Belief	DF	F	DF	F
	Suspension			F	DF
	Disbelief			DF	DF

In [omitted] I argue that the bad case is an *epistemic dilemma* – a situation in which one is doomed to do wrong whichever doxastic attitude one adopts towards p . The table reflects this idea. In the bad case there is no way for cognition to be functional both locally and broadly at the response stage.²⁰ If the agent suspends on p , then whilst cognition is locally functional, it is broadly dysfunctional. If the agent believes that p , then whilst cognition is broadly functional, it is locally dysfunctional. If the agent disbelieves that p , cognition is both locally and broadly

²⁰ Williamson (fc2) echoes this point.

dysfunctional. As we'll see, feedback loops are also epistemic dilemmas, though in a different way.

§6.2. Loop Cases

How should we fill in a table like the one above for loop cases? It depends. Is it possible for Camille, Levi, and Asher to come to *know* that *p*, given the deviant etiology of their beliefs? Intuitively, it is not – or at least, so it seems to me. Normality-theoretic approaches to knowledge (e.g., Goodman & Salow 2018, Beddor & Pavese 2020) deliver this verdict. According to these views, an agent knows that *p* *only if* the way the agent forms their belief results in a true belief in all normal circumstances. Since there are normal circumstances in which Camille, Levi, and Asher falsely believe that *p*, it follows that they don't know that *p*.

However, according to minimalist similarity-theoretic approaches to knowledge, which take safety from error in similar worlds to be necessary *and sufficient* for knowing (Lasonen-Aarnio 2010, Byrne 2018, Hirvela *fc*), it seems to be *in principle* possible for an agent in a loop cases to come to know that *p*, as there may be no nearby world in which they falsely believes that *p*. Perhaps, for instance, Levi is such a fantastic guy that things would have to be very different for him to believe that he is liked when he isn't.

The dispositionalist framework is able to capture the distinctive epistemic properties of feedback loops on either assumption. For this reason, I'll remain neutral on the knowledge question and simply show how things look given the respective assumptions.

§6.2.1. No Knowledge

DOMINEERING CEO, LIKEABLE LEVI, and UNHAPPY ZORA don't specify whether p is true (i.e. whether the company is financially stable, Levi is liked, and Zora is unhappy). Obviously in versions of the cases in which p is false Camille, Levi, and Asher can't come to know that p . But how do things look on the assumption that they can't come to know that p even if p is true (perhaps because they falsely believe that p in some normal circumstances)?

In that case, cognition is:

- *Locally dysfunctional at the acquisition stage*, as they acquire delusive evidence about p .
- *Locally dysfunctional at the response stage when they believe that p* , as they don't know that p .
- *Broadly dysfunctional at the acquisition stage*, as the way they acquire evidence yields delusive evidence about p in some normal circumstances.
- *Broadly functional at the response stage when they believe that p* , as they manifest a knowledge-conducive disposition.

Table 2 represents the situation and compares it with the bad case:

Table 2	Doxastic attitude	Local acquisition	Broad acquisition	Local response	Broad response
Feedback case without knowledge	Belief	DF	DF	DF	F
	Suspension			F	DF
	Disbelief			DF	DF
Bad case	Belief	DF	F	DF	F
	Suspension			F	DF
	Disbelief			DF	DF

Notice the contrast. None of Annie, Camille, Levi, or Asher know that p . However, the framework identifies an additional problem with Camille, Levi, and Asher that Annie doesn't have – cognition is broadly dysfunctional at the acquisition stage.

§6.2.2. Knowledge

How do things look on the assumption that knowledge is possible in loop cases? Of course, if p is false then Camille, Levi, and Asher can't come to know that p . In those versions of the cases the situation will be the same

as above. But what about versions of the cases in which p is *true* and they each successfully come to know that p ?

In these cases cognition is:

- *Locally functional at the acquisition stage*, as they do not acquire delusive evidence about p
- *Locally functional at the response stage when they believe that p* , as they believe knowledgeably.
- *Broadly dysfunctional at the acquisition stage*, as the way they acquire their evidence for p yields delusive evidence about p in some normal circumstances.
- *Broadly functional at the response stage when they believe that p* , as the way they respond to the evidence manifests a knowledge-conducive disposition.

Table 3 represents the situation and compares it with the good case:

Table 3	Doxastic attitude	Local acquisition	Broad acquisition	Local response	Broad response
Feedback case with knowledge	Belief	F	DF	F	F
	Suspension			DF	DF
	Disbelief			DF	DF
Good case	Belief	F	F	F	F
	Suspension			DF	DF
	Disbelief			DF	DF

Again, notice the contrast. Although Annie, Camille, Levi, and Asher all know that p , the framework identifies a problem with Camille, Levi, and Asher that Annie doesn't have – cognition is broadly dysfunctional at the acquisition stage.

§6.3. Dilemmas

On either assumption, a consequence is that feedback loops are also epistemic dilemmas – this time at the level of broad cognition. Since

cognition is broadly dysfunctional at the acquisition stage whatever doxastic attitude the agent adopts, it is dysfunctional *overall*, even though it is broadly functional at the response stage when the agent believes that p.

There would be no such dilemma if Camille, Levi, and Asher had not engaged in broadly dysfunctional evidence acquisition in the first place, and it is plausible that there is a norm prohibiting broadly dysfunctional acquisition. Call it the ‘Acquisition Norm’:

ACQUISITION NORM: One must not engage in broadly dysfunctional evidence acquisition.

Hence, these are dilemmas *perplexity secundum quid* – they only arise because agents in loop cases have already violated the Acquisition Norm. However, given that they now *have* their evidence they must respond to it – the evidence cannot be ‘undone’, so to speak.²¹ In this respect feedback loops are different to situations in which an agent dysfunctionally believes that p, knows that if p then q, and is on the cusp of inferring that q. This agent has an ‘escape route’ – the thing for them to do isn’t to believe that q, but to give up their belief that p.

§VII. Rationality & Oughts

In §1 I described Camille, Levi, and Asher’s beliefs as ‘irrational’. Is there a natural way of mapping the concept of rationality on to the

²¹ One might think that these are not dilemmas because believing that p is *more rational* than suspending on p, even if cognition is not functioning well overall. I’ll address this issue in §10.2

dispositionalist framework, which does not make use of it, but rather makes use of the concepts of functional and dysfunctional cognition? I think there is. A natural proposal is that a doxastic attitude is rational iff cognition is broadly functional at *both* the acquisition and response stage, and irrational otherwise. This gives us the result that the beliefs of agents in familiar good and bad cases, like Annie, are rational, as are those of agents envatted and tricked by evil epistemic demons, but that the beliefs of agents in loop cases are irrational, even if they are knowledgeable (assuming that they *can* be knowledgeable).²² It also gives us the result, which may seem undesirable to some, that feedback loops are rational dilemmas.

On the plausible assumption that one ought to be rational and ought not to be irrational, we can also derive deontic conclusions. Agents like Annie in good and bad cases ought to believe that *p*, as ought agents who are envatted or tricked by evil epistemic demons. Agents in loop cases ought not to believe that *p* and face deontic dilemmas.

§VIII. The Desiderata

How does dispositionalism fare on the five desiderata described in §2?

§8.1. EXCEPTIONS

Let's start with EXCEPTIONS. Consider Jake again. He desires to be naked. This desire causes him to take his clothes off, which causes him to believe

²² What about Wolsey? We might be reluctant to call his beliefs 'rational' or 'irrational'. If so, we can simply note that his cognition is broadly dysfunctional at the acquisition stage and broadly functional at the response stage.

that he is naked. Jake acquires evidence for p by *making* p the case. There are no normal circumstances in which this way of acquiring evidence yields delusive evidence. As a result, although this case has the same superficial structure as problematic loop cases, the dispositionalist framework judges it quite differently. Unlike with Camille, Levi, and Asher, Jake's cognition is broadly functional at the acquisition stage. Since his response to the evidence is also broadly functional, the theory tells us that Jake's belief is rational, unlike Asher, Camille, and Levi's. This is the right result. Hence, the EXCEPTIONS desideratum is satisfied.

§8.2. VARIETY

What about VARIETY? Dispositionalism says that cognition is broadly dysfunctional at the acquisition stage iff the *way* the agent acquires evidence yields delusive evidence in some normal circumstances. This diagnosis is indifferent to whether the evidence is created (as in UNHAPPY ZORA and DOMINEERING CEO) or selected (as in LIKEABLE LEVI); whether the prior attitude towards p is suspicion, desire, fear, or something else; whether the evidence is testimonial, visual-perceptual, or of some other form; whether the evidence is internal or external to the mind; and whether the case is one of confirmation bias or cognitive penetration. Despite these differences, in each case dispositionalism identifies a common problem: the way in which the evidence was acquired yields delusive evidence in some normal circumstances. Hence, the VARIETY desideratum is satisfied.

§8.3. INCLUSIVITY

What about INCLUSIVITY? Recall, Wolsey's belief is just as epistemically problematic as Asher's. Since dispositionalism makes no appeal to properties found only in sophisticated agents, it is capable of explaining

what's going wrong with Wolsey. The way in which he acquires evidence yields delusive evidence in some normal circumstances. Hence, the INCLUSIVITY desideratum is satisfied.

§8.4. NUANCE

What about NUANCE? We've already seen that there is a difference between Annie in the bad case and our feedback agents. Even though Annie's cognition is locally dysfunctional at the acquisition stage, it is broadly functional – Annie would not acquire delusive evidence in normal circumstances. This explains why it is rational for her to believe that *p*, but not rational for Asher, Camille, and Levi to believe that *p*. What goes for Annie also goes for more far-reaching sceptical scenarios such as those involving envattment and the machinations of evil epistemic demons – here too cognition is broadly functional at the stage of evidence acquisition. Hence, dispositionalism explains the differences between bad cases and loop cases. It also recognises similarities between them – cognition is broadly functional at the response stage with each of Annie, Asher, Camille, and Levi when they believe that *p*. Hence, the NUANCE desideratum is satisfied.

§8.5. AMBIVALENCE

What about AMBIVALENCE? As noted, in one way cognition is functional with Asher, Camille, and Levi when they believe that *p* – it is broadly functional at the response stage. It would *not* be functional if they were to suspend judgement on *p* or disbelieve that *p*. If we were to focus only on the response stage and ignore the way in which their evidence was acquired, Asher, Camille, and Levi would seem to be paragons of rationality. By contrast, patient X is broadly *dysfunctional* at the response stage. His evidence, we may suppose, is not acquired in a deviant way.

Rather, he fails to *respond* to it correctly – images of fruit on your wife’s social media page are not evidence that she is having an affair. So dispositionalism identifies a significant difference between Camille, Levi, and Asher, on the one hand, and patient X on the other. Hence, it satisfies the AMBIVALENCE desideratum.

§IX. Consequences for Normative Epistemology

This completes my account of feedback loops. In this section I’ll draw out some consequences for normative epistemology. In section 10 I’ll reply to objections.

Many theories of rational belief are unable to explain epistemic feedback loops and so should be rejected. Here I’ll give two examples from the recent literature, before extracting a general lesson. In each case the problem is the same. The theories are designed to deal with the response stage of cognition but ignore the acquisition stage. Patching them up by appealing to the dispositionalist framework results in theories that are *ad hoc* and disunified.

§9.1. Lord’s Reasons-First View

Errol Lord (2018) argues that one rationally believes that p iff one believes that p for a sufficient objective reason r to do so and manifests know-how to use r as a reason to believe that p .²³ He maintains that only facts are objective reasons. In order to accommodate the observation that it is rational for Annie to believe that 44 is an apartment building in the bad

²³ Kieseewetter (2017) holds a similar view.

case, despite it not being a fact that it is an apartment building, he maintains that the fact that it *perceptually appears* to Annie that p is a sufficient objective reason for her to believe that p .

However, just as it perceptually appears to Annie that 44 is an apartment building, so too it perceptually appears to Asher that Zora is unhappy. Yet whereas Annie's belief is rational, Asher's is irrational. As far as I can see, nothing in Lord's account provides him with the resources to explain this difference. The upshot is that his theory cannot handle feedback loops. So it should be rejected.²⁴

You might think that Lord can help himself to the idea of cognition being broadly functional or dysfunctional at the acquisition stage in order to deal with the problem. The resulting view would be that one rationally believes that p iff one believes that p for a sufficient objective reason r to do so and manifests know-how to use r as a reason to believe that p *and one acquires r in a broadly functional way*. But this view is an *ad hoc* and disunified mutant hybrid. It uses one theoretical framework to deal with the acquisition stage – employing the concept of the manifestation of ways of acquiring evidence in normal circumstances – and an entirely different framework to deal with the response stage – employing the concept of believing for a sufficient objective reason r and manifesting know-how to use r as a reason to believe that p . We should not accept a cobbled together view like this when a more unified one is available.

§9.2. Dutant & Littlejohn's Probable Knowledge View

Dutant and Littlejohn (fc) argue that it is rational for one to believe that p iff it is sufficiently probable on one's evidence that one knows that p . Is it

²⁴ Miracchi (2019) also argues that Lord's view cannot handle cases where the agent acquires their evidence in a bad way, though her approach is different to mine.

sufficiently probable on Annie's evidence in the bad case that she knows that p ? Dutant and Littlejohn don't say (they focus on how their view handles lottery and preface cases). Presumably it is. If it wasn't, the view would be a non-starter – any plausible theory of rationality must deliver the verdict that Annie's belief is rational. Is it probable on Camille, Levi, and Asher's evidence that *they* know that p ? Again, the answer is presumably 'yes' – at least according to Dutant and Littlejohn's view. They make no distinction between functional and dysfunctional evidence acquisition and Camille, Levi, and Asher's beliefs would be rational were their evidence not to have a deviant etiology. But this means that their view faces the same problem as Lord's: it cannot explain why Annie's belief in the bad case is rational but Camille, Levi, and Asher's beliefs are irrational.

Dutant and Littlejohn might try to help themselves to the idea of cognition being locally or broadly functional or dysfunctional at the acquisition stage in order to deal with the problem. But the result would once again be an *ad hoc* and disunified mutant hybrid. The more unified view is preferable.

§9.3. A Lesson

A broader lesson can be taken from the failure of these views to accommodate feedback loops. The theory of rationality needs to account for the fact that some ways of acquiring evidence are bad and others are okay. The best way to do so, I suggest, is to look at the outcomes that different ways of acquiring evidence produce in normal circumstances. Since a more unified theory is, *ceteris paribus*, a better theory, this gives us a reason to adopt the same approach to evaluating ways of responding to evidence – looking at the outcomes that different ways of responding to evidence produce in normal circumstances. So theories that do not take

this approach should be rejected – and there are plenty of them. They include (at least) Conee and Feldman’s (2004) version of Evidentialism; many forms of Coherentism (e.g. Lehrer 1990); Phenomenal Conservatism (Huemer 2001); Dogmatism (Pryor 2000);²⁵ Process Reliabilism (Goldman 1979); and several Knowledge-First theories (Bird 2007, Reynolds 2011, Ichikawa 2014).

§X. Questions and Answers

The rest of the paper addresses some objections to my view. If you’re already convinced, you can stop reading here. If not, carry on.

§10.1. Are these beliefs really irrational?

I’ve proposed that a doxastic attitude is rational iff cognition is broadly functional at *both* the acquisition and response stage, and irrational otherwise. This proposal might rankle some epistemologists. It is not uncommon to find endorsements of one or more subset of the following claims:

- Only factors accessible to consciousness make a difference to whether it is rational for an agent to believe that p.²⁶

²⁵ I’m not the first to point out that cases like UNHAPPY ZORA cause problems for Phenomenal Conservatism and Dogmatism. See Lyons (2011), Siegel (2013, 2017), and McGrath (2013b).

²⁶ The most comprehensive defence of this claim I know of is Smithies (2019).

- What is outside one's control is outside the sphere of rationality.²⁷
- Facts about what it is rational for them to believe must be capable of guiding agents in belief formation.²⁸
- An irrational belief is a belief for which the agent can be held responsible.²⁹
- If an agent is irrational in believing that p, they are blameworthy and criticisable for believing that p.³⁰

Now consider UNHAPPY ZORA. The process that renders Asher's evidence acquisition dysfunctional is consciously inaccessible to him and out of his control. Arguably, he cannot be guided in belief formation by the (would-be) fact that it is irrational for him to believe that p. It also seems inappropriate to criticise or blame him for his belief, or to hold him responsible for it – in part precisely *because* the dysfunctional process is not consciously accessible to him or within his control. So anyone who goes in for one or more of the claims listed above will presumably reject my claim that Asher is irrational when he believes that Zora is unhappy.

Elsewhere [omitted] I've argued that each of the claims on this list is false, resting on outdated ideas about mind and agency. There is no space to rehearse those arguments here. Instead I'll note that *if* these claims *are* indeed false then a different way of conceptualising rationality is needed. One of its main tasks, it seems to me, should be to distinguish between

²⁷ Wedgwood (2012) is a very clear statement of this view.

²⁸ This is a common refrain. Pollock (1987) is the contemporary *locus classicus*.

²⁹ See Bonjour (1985), amongst many others.

³⁰ This is a popular thesis amongst those who endorse deontological conceptions of epistemic justification. See Alston (1988) for early criticisms.

flaws in cognition that are attributable to an uncooperative environment and flaws that are attributable to the internal workings of the cognitive system itself. My way of mapping rationality and irrationality onto functional and dysfunctional cognition respects this distinction. Agents in loop cases are irrational because when they believe without knowing it is due to a flaw that is internal to the cognitive process, despite being external to consciousness. Agents in bad cases are rational because when they believe without knowing it is not due to a flaw internal to cognition, but rather because of an uncooperative environment. A theory of rationality that judges agents in loop cases and agents in bad cases to be rationally on a par (perhaps because it concerns itself only with the response stage) misses this important distinction.

However, even readers who see things differently to me, perhaps because they are impressed by purported connections between rationality and control, guidance, praise and blame, responsibility, access, and the like, can make use of my theory. Just replace ‘irrational’ with ‘problematic’ or ‘sub-optimal’, or whatever your preferred description is.

§10.2. Are feedback loops really dilemmas?

I’ve argued that feedback loops are rational dilemmas – situations in which there is no rationally acceptable doxastic attitude for the agent to adopt. Some readers might take issue with this.³¹ When agents in loop cases believe that *p* they are broadly dysfunctional at the acquisition stage but broadly functional at the response stage. But if they were to suspend on *p* or disbelieve that *p* they would be broadly dysfunctional at *both* stages. So isn’t it better for them to believe that *p* than to suspend or disbelieve? And if it is, are these really dilemmas?

³¹ Some readers might take issue with the idea that epistemic dilemmas are even possible in the first place. See [omitted] for an extended defence of the claim that they are.

I agree that an agent who is dysfunctional at both the acquisition and the response stage is in some sense rationally worse off than an agent who is dysfunctional at the acquisition stage but functional at the response stage. However, it would be a mistake to think that just because φ -ing is less rational than ψ -ing, ψ -ing must be rational. Compare: it's less bad to punch someone than it is to shoot them. But that doesn't mean it's okay to punch them. If we say that it is rational *tout court* for agents in loop cases to believe that p , we thereby license dysfunctional evidence acquisition and once again lose the ability to explain how they are different from agents in the regular bad case. These are undesirable outcomes. Hence, I think it is better to treat loop cases as dilemmas.

§10.3. Evidence

I've assumed that each of Camille, Levi, and Asher actually acquire evidence for p . But I haven't said anything much about what evidence *is*, and there is a way of thinking about it according to which Asher's situation is different from Camille and Levi's in this respect. Suppose that $E=K$; that is, that one's total evidence consists of one's total knowledge (Williamson 2000). Camille and Levi know those things that cause them to believe that p – Camille knows that the CFO asserted that p and Levi knows the things that cause him to believe that he is liked (e.g. that he has been complimented and hasn't been criticised). But Asher does not know that Zora's face is wearing an unhappy expression. So, given $E=K$, the proposition that her face is wearing an unhappy expression is *not* part of his evidence. Of course, there are other things in the vicinity that he is presumably in a position to know, such as that it appears to him that Zora's face is wearing an unhappy expression, but he may not form the belief that it appears to him as such, and his canine counterpart, Wolsey, may not be *able* to form such a belief. We should take seriously, then, the

possibility that Asher doesn't acquire any evidence *at all* bearing on p when he bumps into Zora at the bar. This looks like a problem for the framework. If Asher doesn't acquire evidence in the first place, surely believing that p cannot be a broadly functional response to his (non)-evidence?

In reply, an adaptation of the dispositionalist framework can explain UNHAPPY ZORA even without using the concept of evidence. By hypothesis, Asher undergoes a perceptual experience representing Zora's face as wearing an unhappy expression and this perceptual experience causes him to believe that p. We can adapt the framework to fit this description. Replace the stage of 'acquiring evidence' with the stage of 'having a perceptual experience' and the stage of 'responding to the evidence' with the stage of 'causing a belief that p'. Asher's cognition is *broadly dysfunctional* at the stage of having a perceptual experience and *broadly functional* at the stage of causing a belief that p. It is dysfunctional at the first stage in virtue of its deviant etiology: the perceptual experience is not the product of Asher being disinterestedly open to his environment – a belief caused by a perceptual experience with this etiology will fail to be knowledge in a range of normal circumstances. It is functional at the second stage because Asher manifests a knowledge-conducive disposition in believing that p in response to the perceptual experience; the disposition Asher manifests would normally yield knowledge.

This description of the adapted framework is fairly minimal, of course. It doesn't tell us anything about the nature of perceptual experiences, and it leaves open how we should more precisely characterise 'causes one to believe'. But it isn't necessary to address these questions here. What's important for our purposes is the modal properties of Asher's way of coming to believe that p, for it is these that explain why his belief is epistemically problematic. These properties are captured even by this minimal description.

§10.4. *All* normal circumstances?

Dispositionalism says that cognition is broadly functional at the acquisition stage iff the way the agent acquires evidence about *p* does not yield delusive evidence about *p* in *any* normal circumstances. A complaint I've heard is that this is too strong; even the very best ways of acquiring evidence *sometimes* yield delusive evidence in normal circumstances. Here are two examples that seem to illustrate the point.

First example: suppose that Camille is such a naturally intimidating presence that people normally tell her what they think she wants to hear, even when she isn't putting pressure on them. In that case, even good ways of acquiring evidence, such as simply asking people questions *without* pressuring them to give a particular answer, will result in her acquiring delusive evidence in circumstances that are normal *for her*.³²

Second example: suppose that a scientist randomly samples a population in order to discover the prevalence of property *F*. The test indicates that *F* is common (>90% prevalence). In fact, it is rare (<10% prevalence). As sometimes happens, the scientist was unlucky and got a non-representative sample. Yet this outcome, the argument goes, need not be the result of *abnormality*. Nothing weird or out of the ordinary need happen for it to arise – it's *just* bad luck.³³

I'll take these in turn. The first case can be handled by noting that within my framework normal circumstances are not those that are normal *for a given individual*, but rather those that are normal *for a way of acquiring evidence*. It may be normal *for Camille* to get delusive evidence from

³² Thanks to [omitted] for suggesting this case.

³³ Thanks to [omitted] for suggesting this case

disinterested question-asking. But that's because she inhabits a world which is *abnormal* from the point of view of acquiring evidence by disinterestedly asking questions. Normal circumstances for this way of acquiring evidence do not include those in which people tell you whatever they think you want to hear.

I have two replies to the second case, one steadfast and one conciliatory. The steadfast reply is this. It is not, it seems to me, rational for the scientist to form an *outright* belief on the prevalence of F in the population on the basis of random sampling. At most, it is rational for her to form a high credence about the prevalence of F. If so, there are *no* circumstances, let alone normal circumstances, in which her way of acquiring evidence yields *delusive* evidence – that is, evidence which is such that a rational response to it yields a non-knowledgeable *belief* (rather than a high credence in a falsehood). The evidence might well be misleading in some other sense of 'misleading', but that is beside the point.³⁴

The conciliatory reply is this. Even if we accept that the very best ways of acquiring evidence, including those which rationalise outright belief, sometimes yield delusive evidence in normal circumstances, the framework can be adapted to reflect this. Some ways of acquiring evidence are better than others. Being disinterestedly open to receiving information from one's environment is a *better* way of acquiring evidence than is manufacturing evidence that harmonises with a prior psychological state. Similarly, randomly sampling a population is a better way of acquiring evidence about the prevalence of property F than is sampling a particular subset of the population which is likely to contain a disproportionate number of individuals with F. (You shouldn't try to figure out the

³⁴ Those who think that belief metaphysically reduces to high credence, or that it is rational to outright believe that p iff it is rational to have a high credence in p, won't like this reply, of course. I don't think either of these things.

prevalence of arthritis in the general population by sampling only people over 65). Taking inspiration from Lasonen-Aarnio (fc2), we can say that cognition is broadly dysfunctional at the acquisition stage iff evidence about p was acquired by way w^1 and there was a better feasible way w^2 of acquiring evidence about p , where w^2 is better than w^1 iff w^2 yields delusive evidence in *fewer* normal circumstances than w^1 .³⁵ Since there are better ways to acquire evidence than by manufacturing it in harmony with a prior psychological state, we will still get the result that agents in loop cases are broadly dysfunctional.

§10.5. What are we in a position to know?

Dispositionalism says you should believe that p if you're in a position to know that p . An objection I've heard is that this is too demanding. The objection arises in two ways.

§10.5.1. Inquiry

Firstly, there is a lot you are in a position to know if you *inquire*.³⁶ This may be as simple as turning your head to see what made that sound or walking from the office to the kitchen to find out what's going on in there. But you might not care what made the sound or what's going on in the kitchen. If so, it seems wrong to say that you are cognitively dysfunctional if you don't do these things – thereby leaving knowledge on the table. Moreover, there are many different things you could inquire about. Instead of finding

³⁵ The *feasible* proviso is important. Presumably there are ways of acquiring evidence which are very good but not practicable most of the time. Think, for example, of testing the *entire population* for property F , rather than merely a random sample. If we go this way with our thinking about evidence acquisition then ultimately we'll want a story about the conditions under which ways of acquiring evidence count as feasible or infeasible. Unfortunately, there is no space to pursue this here.

³⁶ See Friedman (fc) for a discussion of zetetic (i.e. inquiry) norms.

out what's happening in the kitchen you could head to the garden. But you can't engage in *all* possible avenues of inquiry available to you at any given moment. So it looks like you will be *guaranteed* to fail to believe things you are in a position to know. Call these two points, taken together, the *proliferation of inquiry objection*.

§10.5.2. Inference

Secondly, even putting aside inquiry, there is a great deal you are in a position to know simply by working out the consequences of your existing knowledge. But it seems wrong to say that you should spend all your time acquiring this knowledge. A dramatic illustration of the point concerns logical, mathematical, and other a priori truths. There are infinitely many of these which are entailed by your evidence. But it is absurd to suggest that you should spend the rest of your days figuring out as many of them as possible. You have better things to do. A less dramatic example involves the implications of your empirical knowledge. Knowing that p – that your sister is in London, say – you are in a position to know lots of other things. For example, that your sister is not in Paris, not in Istanbul, not in New York, and so on. Again, it is absurd to say that you are cognitively dysfunctional if you don't spend your time drawing out these implications. Call this the *pointless inferences objection*.

§10.5.3. Replies

I'll take these objections in turn. I propose to deal with the proliferation of inquiry objection by interpreting 'in a position to know' *restrictively*. According to an expansive interpretation, one is in a position to know that p provided that one could take steps which would lead to one knowing that p, even if one *presently* has no evidence bearing on p. Clearly, the expansive interpretation is not suitable for our purposes. Nothing is going wrong in cognition if you suspend judgement on whether the kettle has

boiled because you're in the office rather than the kitchen, even though you are, in a sense, in a position to know whether it has boiled (you could go and check). This is not to say that there is nothing wrong with passivity. Often we *need* to actively seek out evidence bearing on questions that are important to us.³⁷ Rather, it is merely to say that norms on active evidence- and knowledge-seeking aren't part of the dispositionalist framework. I'll explain why in the next subsection.

How exactly should we restrict the interpretation of 'in a position to know'? The natural way is to say that cognition is dysfunctional at *t* only if you do not believe that *p*, yet you are in a position to know that *p* *given your evidence at t*. This means that even if you could very easily acquire evidence putting you in a position to know that *p* you are still not in a position to know that *p* on my interpretation, and hence cognition is not dysfunctional when you suspend on *p*.

What about the pointless inferences objection? Restricting 'in a position to know' in this way won't solve that problem – all that inferential would-be knowledge is *entailed* by your evidence at *t*.

In reply, as the Angela-Merkel's-Father and Tigers-in-India examples from earlier were intended to suggest, I am not interested in what sometimes gets called "occurrent belief" (which is better called *conscious judgement*) but rather in "dispositional belief". Knowing that your sister is in London, you might not have entertained the proposition that she is not in Paris, yet you still believe it – dispositionally. Dispositionally, you *already* believe many of the consequences of your current knowledge. You need not consciously work them through and consciously assent to them. When we ask if cognition is functional, we should be looking at the agent's dispositional beliefs, not their conscious judgements.

³⁷ C.f. Friedman (fc).

The focus on dispositional belief and restriction to what you are in a position to know given your evidence are steps in the right direction, but a problem remains. Many logical, mathematical, and other a priori truths entailed by any evidence at all are far beyond the reach of most normal human beings. It seems wrong to say that you are cognitively dysfunctional when you suspend judgement on whether the Reimann Hypothesis is correct, even though the answer is entailed by your evidence. This observation motivates another restriction. I will say that unless one is *able to reliably infer p* from one's evidence, one is not in a position to know that p, and hence cognition is not dysfunctional if one suspends on p. What is the relevant reading of 'able'? We should let judgements about cases be our guide. If S is intuitively doing okay in suspending on p, then she was not in a position to know that p. If S is intuitively not doing okay in suspending on p, then she was in a position to know that p. Intuitively you are doing okay when you suspend on whether the Reimann Hypothesis is correct. You are *not* doing okay if you suspend on whether Angela Merkel's father was over 5 meters tall. Of course, there are bound to be borderline cases in which intuitions are murky. But precision should not be attempted where it is unavailable.

§10.6. Zetetic Epistemology

As we saw in the last subsection, we do more than just passively receive evidence from our environments. We often need to actively seek it out in order to answer pressing questions. The same goes for non-human agents such as Wolsey. It is natural to think that someone who makes no effort to seek out evidence bearing on a question they have a pressing need to answer is in some sense irrational. But the dispositionalist framework says nothing negative about this person. Is that a problem?

I don't think so. What one should seek out evidence for is highly sensitive to practical considerations – one's needs, desires, and priorities, etc. I don't know of any formula or general principle that establishes what one should seek out evidence for and when.³⁸ This is a topic for future research on evidence acquisition, and zetetic epistemology more broadly. But it is not necessary to tackle it here. The dispositionalist framework can be viewed as providing some *minimal* necessary conditions for good evidence acquisition. In particular, it proposes the Acquisition Norm ('one must not engage in broadly dysfunctional evidence acquisition') as a fundamental zetetic norm. Even if it is incomplete, the framework has enough detail to explain feedback loops, which is my goal here.

§10.7. What about existing theories of cognitive penetration?

There is a fairly large literature on the epistemology of cognitive penetration. Might some of this work furnish us with a better theory of feedback loops than mine?

§10.7.1. *The Generalisation Problem*

I don't think so. Such theories focus, naturally enough, on cases like UNHAPPY ZORA. But the phenomenon of epistemic feedback is broader

³⁸ Flores and Woodard (ms) propose:

GATHER-EVIDENCE-TO-UPGRADE: For any person S, and proposition p, if S has any committal doxastic attitude towards p and there is evidence *e* such that S has reason to expect gathering *e* to lead to an epistemic upgrade in their attitude towards p, then S ought to gather *e*.

But this seems to me to be too anthropocentric. Does Wolsey ever have reason to expect that gathering *e* will lead to an epistemic upgrade? I doubt it. He lacks the concepts required to grasp the thought.

than that – cases of confirmation bias like DOMINEERING CEO and LIKEABLE LEVI do not involve cognitive penetration. A problem that many attempts to account for cognitive penetration would face if we treated them as general theories of epistemic feedback loops (though of course they are not offered as such) is that they fail to generalise in a way that explains the broader phenomenon.

For example, consider Matthew McGrath's (2013a, 2013b) 'quasi-inferential' account of cognitive penetration. McGrath argues that beliefs like Asher's are unjustified because the higher-level penetrated perceptual experience (Asher's experience as of Zora's face wearing an unhappy expression) on which the belief that *p* is based is *unjustifiably* quasi-inferred from lower-level perceptual experiences such as shape, size, colour, texture, etc. This process, McGrath argues, is analogous to leaping to a conclusion from insufficient evidence.

Were we to treat this as a general theory of what's going wrong in epistemic feedback loops, it would face the generalisation problem. It doesn't explain what is going wrong in loop cases involving confirmation bias, and so it wouldn't satisfy the VARIETY desiderata. Camille does not believe that the company is financially stable as a result of a quasi-inference from lower- to higher-level perceptual experiences, or, for that matter, as a result of *any* kind of inference akin to leaping to conclusions. She believes it because the CFO told her so (albeit under pressure). Similarly, Levi doesn't believe that he is liked because of a dodgy inference. He believes it because he's had lots of compliments and no criticism. Whatever its merits as a theory of cognitive penetration, McGrath's account would be inadequate as a theory of feedback loops.³⁹

³⁹ The following strikes me as a plausible principle: If A and B are different instantiations of the same underlying phenomenon, and theory 1 successfully explains both A and B, whereas theory 2 could at most explain B, then theory 2 is either redundant (if it is compatible with

Quite a few theories of cognitive penetration would face the same problem as McGrath's – failing to generalise to explain the broader phenomenon. It would, it seems to me, be a problem for Chudnoff's (2020) 'Presentational Conservatism', Brogaard's (2013) 'Sensible Dogmatism', Teng's (2016) 'Imagining Account' and Vahid's (2014) account appealing to the sensorimotor theory of perception, amongst others.

Not every extant theory of cognitive penetration would face the generalisation problem – some of them look like they could potentially explain the broader phenomenon. But they have other problems.

§10.7.2. *Generic Reliabilism*

Jack Lyons (2011, 2016) argues that a generic reliabilism can best explain why beliefs formed on the basis of penetrated experiences are epistemically problematic – the mechanisms that produce them are unreliable. It is not hard to see how this approach might generalise to cases like DOMINEERING CEO and LIKEABLE LEVI. Now, I am not unsympathetic to reliabilism *per se*. Indeed, my view bears some resemblance to Alvin Goldman's (1986) 'normal-worlds reliabilism'. But a *generic* reliabilism seems to me to be clearly inadequate for explaining feedback loops. It satisfies EXCEPTIONS, VARIETY, and INCLUSIVITY. But as Stewart Cohen (1984) pointed out long ago, it cannot explain why beliefs in the regular bad case like Annie's are rational. As a result, it fails to satisfy NUANCE. It also fails to satisfy AMBIVALENCE: the ways that Camille, Levi, and Asher form their beliefs are just as unreliable (in the generic sense) as X's way of

theory 1) or false (if it is incompatible). Either way, theory 2 should be rejected. I think McGrath's view is theory 2. So I'd reject it as a theory of cognitive penetration. The same goes for other theories that face the generalisation problem.

forming his belief. So it doesn't recognise how agents in loop cases are doing well in at least one sense.

§10.7.3. *The Rationality of Perception*

Susanna Siegel (2012, 2017) has put forward an influential account of cognitive penetration. Orthodoxy has it that whilst beliefs can be rational or irrational, perceptual experiences are outside the sphere of rational evaluation. Siegel rejects this orthodoxy, arguing that perceptual experiences and the processes that give rise to them can *themselves* be rational or irrational. Perceptual experiences like Asher's are, she argues, irrational in virtue of their deviant etiologies. Since one cannot arrive at a justified conclusion from an unjustified (i.e. irrationally held) premise, beliefs based on such experiences are themselves irrational.

Is it possible to develop this account in such a way that satisfies all of the desiderata on a theory of feedback loops? *Maybe*. EXCEPTIONS doesn't look like a problem – Jake's evidence doesn't have a deviant etiology. Nor does NUANCE – Siegel makes it clear that her view doesn't indict the perceptual experiences of someone like Annie in the bad case. But what about VARIETY, INCLUSIVITY, and AMBIVALENCE? Siegel says nothing that indicates how her view would meet these desiderata – she doesn't say anything about cases of confirmation bias like DOMINEERING CEO and LIKEABLE LEVI; doesn't discuss non-human agents; and doesn't say anything about the ambivalence that feedback loops give rise to. But perhaps her view could be adapted or extended to handle these cases. VARIETY might be dealt with by arguing that the processes by which Camille and Levi acquire their evidence are irrational (indeed, I suggested something similar earlier when I proposed the Acquisition Norm (§6.3)). INCLUSIVITY might be satisfied given a suitably de-anthropomorphised spin on 'rationality'. AMBIVALENCE might be dealt with by observing that although the beliefs of agents in feedback loops are irrational in virtue of

being based on an unjustified premise, the premise is at least the kind of thing that *would* rationalise believing that p, were it to be justified

However, even if all these suggestions could be made to work (a big ‘if’), the account remains, it seems to me, crucially lacking in detail. The problem is that nothing in it explains *why* a deviant etiology makes a belief or experience irrational. Siegel points to analogies between cognitive penetration and cases of circular reasoning, wishful thinking, and leaping to conclusions. But this merely pushes the explanatory question back a step: what makes *these* things epistemically problematic?⁴⁰ By contrast, I have offered an explanation – a deviant etiology makes for irrationality in virtue of the fact that it will cause the agent to have a non-knowledgeable belief in a range of normal circumstances. Siegel doesn’t give us the whole story. I take myself to have given a more complete version of events. Indeed, I suspect that were we to try to enrich her account, it would end up looking a lot like mine.

§XI. References

Alston, W. 1988. “The Deontological Conception of Epistemic Justification” *Philosophical Perspectives* 2: 257-299

Baehr, J. 2011. “Evidentialism, Vice, and Virtue” in Dougherty, T. (ed.) *Evidentialism and its Discontents*. Oxford University Press

Beddor, B. & Pavese, C. 2020. “Modal Virtue Epistemology” *Philosophical and Phenomenological Research* 101 (1): 61-79

⁴⁰ C.f. Ghijzen (2016). Ghijzen also criticises McGrath on these grounds. You won’t be surprised to hear that I think circular reasoning, wishful thinking, and leaping to conclusions are bad because they yield non-knowledgeable beliefs in normal circumstances.

- Bird, A. 2007. "Justified Judging" *Philosophy and Phenomenological Research* 74 (1): 81-101
- Bonjour, L. 1985. *The Structure of Empirical Knowledge*. Harvard University Press.
- Brogaard, B. 2013. "Phenomenal Seemings and Sensible Dogmatism" Tucker, C (ed.) *Seemings and Justification*. Oxford University Press.
- Byrne, A. 2018. *Transparency and Self-Knowledge*. Oxford University Press
- Chudnoff, E. 2020. "Experience and Epistemic Structure: Can Cognitive Penetration Result in Epistemic Downgrade" Nes, A. & Chan, T. (eds.) *Inference and Consciousness*. Routledge
- Cohen, S. 1984. "Justification and Truth" *Philosophical Studies* 46 (3): 279-295
- Conee, E. & Feldman, R. 2004. *Evidentialism: Essays in Epistemology*. Oxford University Press
- Dorst, K. 2020. "Evidence: A Guide for the Uncertain" *Philosophy and Phenomenological Research* 100 (3): 586-632
- Dutant, J. & Littlejohn, C. forthcoming. "Defeaters as Indicators of Ignorance" Simion, M. & Brown, J. (eds.) *Reasons, Justification, and Defeat*. Oxford University Press
- Flores, C. & Woodard, E. manuscript. "Epistemic Vigilance: In Defense of Epistemic Norms on Evidence-Gathering"
- Friedman, J. forthcoming. "The Epistemic and the Zetetic" *The Philosophical Review*

- Ghijsen, H. 2016. "The real problem of cognitive penetration" *Philosophical Studies* 173: 1457-1475
- Goldman, A. 1979. "What Is Justified Belief?" Pappas, G. (ed.) *Justification and Knowledge*. Dordrecht
- Goldman, A. 1986 *Epistemology and Cognition*. Harvard University Press
- Goodman, J. & Salow, B. 2018. "Taking a Chance on KK" *Philosophical Studies* 175 (1): 183-196
- Graham, P. 2017. "Normal Circumstances Reliabilism: Goldman on Reliability and Justified Belief" *Philosophical Topics* 45 (1): 33-61
- Greco, J. 2005. "Justification is not Internal" in *Contemporary Debates in Epistemology*. Blackwell.
- Hall, R. & Johnson, C. 1998. "The Epistemic Duty to Seek More Evidence" *American Philosophical Quarterly* 35 (2): 129-139
- Hirvela, J. forthcoming. "Global Safety – How to Deal With Necessary Truths" *Synthese*
- Huemer, M. 2001. *Skepticism and the Veil of Perception*. Rowman & Littlefield
- Ichikawa, J. "Justification Is Potential Knowledge" *Canadian Journal of Philosophy* 44 (2): 184-206
- Kiesewetter, B. 2017. *The Normativity of Rationality*. Oxford University Press
- Lasonen-Aarnio, M. 2010. "Unreasonable Knowledge" *Philosophical Perspectives*. 24: 1-21

- Lasonen-Aarnio, M. forthcoming-1. "Virtuous Failure and Victims of Deceit" Dutant, J. & Dorsch, F. (eds.) *The New Evil Demon Problem*. Oxford University Press
- Lasonen-Aarnio, M. forthcoming-2. "Dispositional Evaluations and Defeat" *Philosophy and Phenomenological Research*
- Lehrer, K. 1990. *Theory of Knowledge*. Westview Press
- Littlejohn, C. 2013, "The Russellian Retreat" *Proceedings of the Aristotelian Society* 113 (3): 293–320
- Lord, Charles G., Ross, Lee, and Lepper, Mark R., 1979. 'Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence'. *Journal of Personality and Social Psychology*, 37(11): 2098–2109
- Lord, E. 2018. *The Importance of Being Rational*. Oxford University Press
- Lord, E. & Sylvan, K. forthcoming. "Beginning in Wonder: Suspensive Attitudes and Epistemic Dilemmas" Hughes, N. (ed.) *Epistemic Dilemmas*. Oxford University Press
- Lyons, J. 2011. "Circularity, Reliability, and the Cognitive Penetrability of Perception" *Philosophical Issues* 21: 289-311
- Lyons, J. 2016. "Inferentialism and Cognitive Penetration of Perception" *Episteme* 13 (1): 1-28
- Maynard, JL. forthcoming. *Ideology and Mass Killing*. Oxford University Press
- Nelson, M. 2010, "We Have No Positive Epistemic Duties" *Mind* 119 (473): 83–102

- McGrath, M. 2013a. "Siegel and the Impact for Epistemological Internalism" *Philosophical Studies* 162: 723-732
- McGrath, M. 2013b. "Phenomenal Conservatism and Cognitive Penetration: The 'Bad Basis' Counterexamples" Tucker, C. (ed.) *Seemings and Justification*. Oxford University Press
- Millikan, R. 1984. *Language, Thought, and Other Biological Categories*. MIT Press
- Miracchi, L. 2019. "When Evidence Isn't Enough: Suspension, Evidentialism, and Knowledge-first Virtue Epistemology" *Episteme* 16: 413-437
- Nickel, B. *Between Logic and the World: An Integrated Theory of Generics*. Oxford University Press.
- Pal, K. *et al.* 2012. "Othello Syndrome Secondary to Ropinirole: A Case Study" *Case Reports in Psychiatry* 353021
- Pollock, J. 1987. "Epistemic Norms" *Synthese* 71: 61-95
- Pryor, J. 2000. "The Skeptic and the Dogmatist" *Nous* 34 (4): 517-549
- Quattrociocchi, W, Scala, A, & Sunstein, S. 2016 "Echo Chambers on Facebook" Available at <http://dx.doi.org/10.2139/ssrn.2795110>
- Reynolds, S. 2011. "Justification as the Appearance of Knowledge" *Philosophical Studies* 163: 367-383
- Salow, B. 2018. "The Externalist's Guide to Fishing For Compliments" *Mind*, 127 (507): 691-728
- Siegel, S. 2012. "Cognitive Penetrability and Perceptual Justification" *Nous* 46 (2): 201-222

- Siegel, S. 2017. *The Rationality of Perception*. Oxford University Press
- Smith, M. 2010. "What Else Justification Could Be" *Nous* 44 (1): 10-31
- Smith, M. 2016. *Between Probability and Certainty*. Oxford University Press
- Smithies, D. 2019. *The Epistemic Role of Consciousness*. Oxford University Press
- Sosa, E. manuscript. "Epistemic Explanations"
- Teng, L. 2016. "Cognitive Penetration, Imagining, and the Downgrade Thesis" *Philosophical Topics* 44 (2): 405-426
- Vahid, H. 2014. "Cognitive Penetration, the Downgrade Principle, and Extended Cognition" *Philosophical Issues* 24 (1): 439-459
- Wedgwood, R. 2012. "'Rational 'Ought' Implies 'Can'". *Philosophical Issues* 23 (1): 70-92
- Williamson, T. 2000, *Knowledge and its Limits*. Oxford University Press
- Williamson, T. forthcoming-1. "Justifications, excuses, and sceptical scenarios" Dutant, J. & Dorsch, F. (eds.) *The New Evil Demon Problem*. Oxford University Press
- Williamson, T. forthcoming-2. "Epistemological Ambivalence" Hughes, N. (ed.) *Epistemic Dilemmas*. Oxford University Press