Reasoning, Defeasibility, and the Taking Condition

Markos Valaris

[Preprtint version; published in *Philosopher's Imprint* vol. 20 no. 28, 2020.]

1. Introduction

Jacinta hears the doorbell ring and, as a result, she comes to believe that Amit is home. There are different ways for this belief to come about. It might be the result of sheer habit, or even a direct causal effect of the stimulation of Jacinta's auditory nerves (Jacinta might have some strange brain wiring). Alternatively, however, it might be that Jacinta *infers* or *reasons* that Amit is home.[1] In that case, Jacinta's belief that Amit is home is not (or not merely) a causal effect of her belief that the doorbell is ringing, but rather it is *rationally based* on it: Jacinta now believes that Amit is home on the *grounds that*, or *for the reason that*, the doorbell is ringing.[2] Furthermore, on natural ways of filling in the story, this is a way for Jacinta to *justifiably believe*, and even *know*, that Amit is home.

Nothing I have said so far should be controversial. But what is it that sets reasoning apart from the other cases? Still painting with an extremely broad brush, we can say the

---

[1] I use 'inference' and 'reasoning' interchangeably in what follows, on the assumption that, even if they are not exact synonyms in English, they refer to the same psychological phenomenon. Dancy (2004, 102–5) disagrees, while acknowledging that he has no explicit account of the distinction. As should become clear in what follows, any disagreement with Dancy here is only terminological: what is at issue is the 'movement of thought' (Dancy's term) in which an agent responds in light of relevant features of her circumstances—where the response may be a belief, intention, or even overt action (see also Dancy (2014)). Such movements of thought are to be distinguished from *arguments*, understood as abstract structures of propositions related to each other by relations of consequence (or probabilification).

[2] Reasoning is closely related to what epistemologists call the 'basing relation', but the two are not obviously co-extensive. For example, my anger at you may be based on my belief that you wronged me (that you wronged me is my reason for being angry at you), but anger does not seem to be the sort of response one *reasons* to. I will have to leave the connection between basing and reasoning open for now.

following: in reasoning, you do not merely *respond to* relevant features of your situation, but rather you respond *in light of* those features. In some sense, Jacinta *treats* the doorbell's ringing as warranting or supporting her believing that Amit is home. One of the challenges for the theory of reasoning has been to spell out what exactly this means.

On one approach to this challenge, reasoning is subject to the 'Taking Condition' (a label due to Boghossian (2014)). On this approach, reasoning to a certain response—paradigmatically, a propositional attitude such as a belief or an intention[3]—requires *taking it that* that response is in some sense warranted or supported by your reasons: [4]

> TAKING CONDITION: Necessarily, if your response *R* in certain circumstances *C* is a
> case of reasoning, then your *R*-ing is at least in part to be explained by your taking it, of
> some feature (or set of features) of *C* or other, that it warrants *R*-ing.

This statement of the Taking Condition is still too vague. For it to mark out a determinate view, we need to make some assumptions about the nature of the 'takings' involved. In what follows (and in contrast with Boghossian himself), I will make some fairly strong assumptions: I will assume that they are representational states with intentional content, and are moreover subject to epistemic evaluation. They are, that is, states that can be correct or incorrect, and also rational or irrational (in an epistemic sense). This is a combination of features paradigmatically possessed by *beliefs*. On this view, therefore, the 'takings' required by the Taking Condition are belief-like

---

[3] Some comments about the terms 'proposition' and 'propositional attitude'. To begin with, it should be clear that my talk of propositions is not meant to imply *linguistic* representations. On my usage, a propositional representation is simply any representation that possesses truth- or accuracy-conditions. Nothing at all is presupposed about the *vehicles* of propositional representations: photographs and maps, for example, can have propositional content (Rescorla 2009). Second, there is a traditional conception of practical reasoning, according to which such reasoning culminates in *action*, rather than (merely) intention. This view may be consistent with the statement in the text, if we think of intentional actions as embodying intentions or judgments of some kind. For present purposes I will set this issue aside.

[4] The statement of the Taking Condition in the text is not Boghossian's own, though it is clearly related to it. Boghossian (2014, 5) discusses the taking condition as a condition on *theoretical* reasoning, but if the condition is motivated at all, it is equally motivated in the practical case as well.

states; accordingly, I will call this the 'doxastic approach' to reasoning. This is the approach I will defend in this paper.[5]

I take it that the doxastic approach is at least *prima facie* attractive. It connects with the intuitive idea that, as a reasoner, you assume responsibility for managing your beliefs and intentions. As a reasoner, you can be held responsible not just for your beliefs and intentions themselves, but also for what considerations you treat as reasons for those beliefs and intentions. The doxastic approach captures this idea in a very straightforward way: treating something as a reason involves taking it to be a reason, in a sense that is open to at least some relevant kinds of normative evaluation.

Despite its initial attractions, however, the doxastic approach has been the subject of much criticism. Some argue that it is too intellectualist. Young children and even some non-human animals seem capable of reasoning; but it seems like a further question whether they are capable of justified belief-like states about what is a reason for what (for a recent version of the criticism, see Buckner (2019)). Perhaps the most famous criticism of all, however, is that the doxastic approach is subject to problems of vicious circularity or regress. For one thing, if we were to reduce treating something as a reason to a mere belief or belief-like state, then wouldn't you have to *reason* from that state in order to work out what to do? If so, then it seems that regress looms (for classic statements of arguments of this sort, see Carroll (1895), Ryle ([1949]2002) and Boghossian (2003)). Moreover, one might worry about how the states required by the doxastic approach could ever be justified, if not *also* by reasoning (e.g., Wedgwood 2006; McHugh and Way 2018).

[5] Proponents of the doxastic view include Leite (2008), Marcus (2012; forthcoming), Neta (2013; 2019), and Valaris (2014; 2016a; 2018). Hlobil (2019) adopts a view on which, instead of a belief, the 'takings' in question involve a *sui generis* attitude of 'attaching inferential force' to collections of propositions. For present purposes, this is sufficiently like a belief that Hlobil's view would also count as doxastic. Tucker (2012) and Broome (2013, 2014) argue that reasoning requires its seeming to you that your premises support your conclusion, but deny that such 'seemings' are the sort of thing that admits of epistemic evaluation. Their views, therefore, contrast with the doxastic approach.

While I believe that these criticisms can be answered, I will not attempt to do so in this paper (except for some brief remarks at the end). What I aim to do, instead, is make a positive case *for* the doxastic approach. More specifically, the central theme of the paper is going to be this: recognizing the pervasive defeasibility of our reasoning provides strong reasons in favor of the Taking Condition, and hence in favor of the doxastic approach. Before getting to the substance of this argument, however, we need to go through some preliminaries.

2. Setting the Stage: Taking, Dispositions and Defeasibility

My aim in this section is to set the stage for the main arguments of the paper, by giving some more background to the debate, identifying some of the main positions and challenges, and clarifying my use of some terminology.

Let me begin by looking more closely at the type of view I will treat as a foil in what follows—that is, views that *reject* the Taking Condition. For our purposes, the first thing to note is that such views do not deny the need to explain what sets reasoning apart from other ways of responding to one's circumstances. More specifically, the acknowledge that reasoning involves the reasoner *in some sense* treating a certain response as warranted by her circumstances. What they deny is that this needs to be understood in terms of any kind of belief or belief-like 'taking'. They aim, rather, to capture this idea in terms of *dispositions*, which are assumed not to involve such states. Accordingly, I will call these views 'dispositionalist' approaches to reasoning.[6]

There are two main parts to my argument against such views, and in favor of the Taking Condition to reasoning.

---

[6] The general idea that rational responsiveness to the demands of one's circumstances can be captured in such terms goes back at least to Ryle ([1949] 2002). For more recent defences of such approaches see Winters (1983), Johnston (1988), Brewer (1995), Boghossian (2003), Railton (2006), Wedgwood (2006), Sosa (2007), Mantel (2017), and McHugh and Way (2018). Admittedly, the issues can be confusing, because belief itself is often taken to be a complex dispositional state. I assume that the dividing line is this: the dispositions in question are *not* subject to the sort of epistemic evaluation that beliefs are.

The argument of section 3 takes on a particular (but very popular and widespread) way of developing dispositionalism. To get a sense of the nature of the target of this argument, note that not *all* dispositions to respond to one's circumstances—and not even all dispositions to respond to one's circumstances *in accordance with* what one has most reason to do—are dispositions to *reason*. The point is already illustrated by the example of Jacinta above: if Jacinta's response is due to sheer habit or odd brain-wiring, for example, it should not be counted as a case of *reasoning*. Slightly more controversially, many philosophers argue that at least some beliefs based on *perception* should be construed as non-inferential, and so as not cases of reasoning. Dispositionalists, therefore, face a challenge: *which* of our dispositions to respond to our circumstances are dispositions to reason? [7]

In response, dispositionalists tend to appeal to dispositions to *follow rules* (see, e.g., Boghossian (2002; 2003; 2008); Broome (2013; 2014); McHugh and Way (2018)).[8] The rules in question, in turn, are supposed to be general principles to the effect that, given some set of personal-level mental states (such as beliefs, desires, or intentions), another such state (or perhaps set of states) is either permitted, required, or prohibited. There are disputes both about how exactly to formulate such rules (are they normative statements, or imperatives?), and about what, in general, it is to follow a rule. For present purposes I will set these debates aside. The concern I raise in section 3 is simpler: it is just extremely implausible that human reasoning can be reduced to a set of rules (at least on any conception of rules dispositionalists have on offer).

Familiarly, much—perhaps all—of our reasoning is defeasible, in the following sense. Suppose that in some given circumstances it is rational for you to infer Q from P. Then, if the

---

[7] Some authors—e.g., Skow (2016), Mantel (2017)—seek to provide dispositional accounts of doing things *for reasons*, rather than of *reasoning*. Such views might cast their net more broadly than I do here, including, for example, perceptual beliefs. But the two topics, although related, are not the same.

[8] It is worth noting that, in principle, it is possible to hold a rule-following view *and also* embrace the Taking Condition. In practice, however, most authors embrace rule-following as an alternative to the Taking Condition.

inference in question is defeasible, there will be some *D* (consistent with *P*) such it is not rational for you to infer *Q* from an expanded set of premises that includes *D* in addition to *P*. *D* in such a case is a *defeater* of the inference from *P* to *Q*. For example, it may be rational to infer that Tweety can fly from the claim that Tweety is a bird, while it is not rational to infer this from the claim that Tweety is not just any kind of bird, but specifically an emu. [9] As we will see in section 3, it is the defeasibility of reasoning that makes rule-following approaches unlikely to succeed. [10]

Now, of course, simply showing that such accounts fail is not tantamount to an affirmative defense of the doxastic approach to reasoning. It might be possible, after all, to give a dispositionalist account of reasoning that does not rely on rule-following. I work towards such a defense in the second main part of my argument, in section 4. Once again, defeasibility plays a major role. More specifically, the Taking Condition seems to be essential to *explaining why* reasoning is subject to certain sorts of defeat. (While the two strands of my argument against dispositionalism will develop largely independently of each other, an important—if somewhat speculativeconnection will emerge near the end of section 4.)

---

[9] Can the defeasibility of everyday reasoning be denied? Some authors—e.g., Norton (2003) and Bird (2005)—have argued that patterns of reasoning that have traditionally be thought of as ampliative are, in fact, deductive. These are controversial views, and I cannot discuss them here on the merits. It is, however, worth noting that even deductive reasoning may be subject to some forms of defeat. To be sure, deductive inference is monotonic, in the sense that adding further premises to a deductively valid argument cannot produce an invalid argument. Still, monotonicity does not guarantee *indefeasibility*, in the sense relevant here: indefeasibility is a matter of *rationality*, not *truth-preservation*. For example, the (misleading, let us suppose) testimony of an expert logician may make it *irrational* for you to rely on a certain proof, even though the proof is, in fact, valid.

[10] Dancy (1983; 2004) has long used intuitions about the defeasibility of reasoning to defend his 'particularism', that is, the view that reasons (and specifically *moral* reasons) are not governed by general principles. The issues I raise in section 3 are similar in spirit to Dancy's. As some recent authors have noted, however, Dancy's observations can be accommodated in systems based on rules, at least on a broad enough conception of rules (Schroeder 2011; Horty 2012, 148–55). As we will see below, this accommodation proves cold comfort to dispositionalists.

The details of the two parts of my argument may get complex. However, the basic conception of what we do when we reason that motivates both is straightforward. Reasoning involves responding in light of the rational demands of a given situation. But taking defeasibility seriously requires us to recognize that *working out* the rational demands of a given situation is *itself* a substantive cognitive achievement—even if it is something that we typically do unreflectively, and even automatically. It is the sort of thing you can be right or wrong about, and is itself subject to rational evaluation. But this is tantamount to the Taking Condition: if your response *R* in certain circumstances *C* is a case of reasoning, then it involves your taking that *C* warrants *R*, in a substantive sense that is open to epistemic assessment. I close this paper in section 5 by sketching some implications of this way of thinking about reasoning.

Before going on, let me note that I am using the terms 'rational', 'justified', and 'warranted' interchangeably. Moreover, I take them to apply not just to the attitudes that serve as the starting- and end-points of instances of reasoning, but also to instances of reasoning as such. The idea is straightforward: instances of reasoning can surely be evaluated as better or worse, and this evaluation can affect an agent's overall rationality (see, e.g., Wedgwood (2012) for this usage). Furthermore, just as in the case of the rational appraisal of states like belief and intention, we can make such assessments *prospectively*, or *ex ante*: we can ask what inferences would be rational or irrational for you to make, even if you do not actually make any of them. (This would be the analogue, for the case of reasoning or inference, of the familiar concept of *propositional* justification for beliefs.) Except where otherwise noted, this is the kind of evaluation I will be concerned with in what follows.

Finally, I want to say something about the content of the 'takings' required by the Taking Condition. For the purposes of the present paper, at least, I will work with a fairly modest view on this score. In particular, I will not assume that reasoning requires higher-order beliefs, that is, beliefs about one's own mental states. What it *does* require, rather, is some kind of belief about how one's premises relate to one's conclusions. Focusing on theoretical reasoning for the

7

purposes of illustration, conditional beliefs expressible in first-order terms, such as '*Q*, given *P*', might be sufficient for that purpose. Such conditional beliefs involve a *commitment to* believing *Q* given that *P* is part of one's evidence, but no explicit higher-order content. With these preliminaries out of the way, we can now turn to the argument.

3.   Is reasoning a matter of following rules?

As explained above, dispositionalists typically appeal to dispositions to follow rules in order to account for what sets reasoning apart from other ways of responding to our circumstances. As I will argue in this section, it is very implausible that our reasoning can be reduced to following rules (at least in a way that would offer any comfort to dispositionalists).

Let us, for now, stick to theoretical reasoning. Rule-following theorists typically take logical inference rules (such as *modus ponens*) as their model. On the face of it, however, logical inferences comprise only a small part of our inferential behavior. What happens when we go beyond logic? Consider again the case of Jacinta, who believes that Amit is home upon hearing the doorbell ring. What rule of inference has she instantiated here? Certainly, Jacinta need have no disposition to reason in accordance with the insane rule, 'every time the doorbell rings, believe that Amit is home'. But what better options are there? The challenge for rule-following theorists, therefore, is simply this: to identify a set of rules that are plausible candidates for the rules we follow in our reasoning, when we step beyond logic.

To see how rule-following theorists would address this challenge, we need to look at what they have to say about the *sorts* of rules we should be looking for. Different views differ with regard to the conditions they impose on the rules of reasoning. Let us consider the options.

To begin with, it should be clear that dispositionalists cannot plausibly *refuse* to impose principled restrictions on what counts as a rule of reasoning. If just any repeatable pattern could count as a rule of reasoning, then the appeal to specifically *rule-following* dispositions (as opposed

to just any dispositions whatsoever) seems empty: for any disposition to respond in a certain way that involves propositional attitudes we could trivially specify a 'rule' you are disposed to follow.[11] But then, what should those restrictions be?

On one kind of view, the rules of reasoning are *formal*, in roughly the same sense that the rules of a formal proof-theory are. The clearest proponents of this kind of view are Quilty-Dunn and Mandelbaum (2018). Quilty-Dunn and Mandelbaum (2018, 536-8) explicitly require inferential rules to apply to structured representations, and to do so specifically in virtue of their *constituent structure* rather than their content.

This type of view is not the only one currently on offer. Indeed, rather more popular is the view that the rules of reasoning have a *constitutive* status: they are constitutive of our concepts (Boghossian 2003), the attitudes involved in it (Wedgwood 1999), or flow from the 'constitutive aim' of reasoning (McHugh and Way 2018), or perhaps from 'rationally required projects' (Enoch and Schechter 2008). This constitutive conception is somewhat harder to assess, as it depends on some (often implicit) story regarding the nature (or perhaps 'constitutive aim') of things like concepts and cognitive capacities. Nevertheless, as we will see, the details will not turn out to matter much: the problems facing rule-following approaches seem to generalize to *any* principled attempt to identify a set of rules as the rules of reasoning.

Let us now consider the options in turn. Begin with the formalist approach. On the face of it, this approach is unable to account for reasoning like Jacinta's. Clearly, simply looking at the constituent structure of 'the doorbell is ringing' and 'Amit is home' is not going to help: nothing in the constituent structure of these two sentences suggests that the one follows from the other.

---

[11] McHugh and Way (2018, 173–74) seem to interpret Broome as imposing no conditions on what the rules of reasoning might be. For the reason described in the text, I doubt that this can be the right way to interpret his view: such a view would make any reference to rules an idle wheel, in an attempt to demarcate reasoning. However, interpreting Broome on this point is tricky, as an anonymous referee impressed upon me, and I will not try to push this point further here.

In response, friends of formality might argue that Jacinta's reasoning should be treated as enthymematic: it relies on further implicit premises, and once these are identified we will be able to locate the relevant formal rules. This is the strategy Quilty-Dunn and Mandelbaum deploy to handle 'semantic entailment' (ibid., 541-542).[12] In that case, the challenge is to explain the inference from the belief that 'this apple is red' to 'this apple is colored'. Their suggestion is that, to the extent that such transitions are inferential at all, they are mediated by a further belief, with the content that 'if x is red x is colored'. Once this belief is included as a further premise, the inference of course is logically valid.

I do not think such a strategy can help, however. There is no doubt, of course, that Jacinta's inference depends on what else she happens to believe at the time. It may, for example, depend on Jacinta's knowledge that Amit is out dropping off the kids at tennis practice, that he left without his house keys, that he is likely to come home straight after, and so on. But what reason is there to think that this dependence can be captured in formal rules?

Consider again the example of Tweety the emu. The information that Tweety is a bird is supposed to enable you to infer that Tweety can fly. So, we might ascribe to you a rule that from 'x is a bird' you can infer that 'x can fly'. But then, how are we to capture the fact that the inference is defeated by the information that Tweety is an emu? Since an emu is a type of bird, this information should not lead you to *abandon* the belief that Tweety is a bird. If so, however, the formalist approach predicts that you should *still* infer that Tweety can fly, despite his being an emu. And, obviously, the point generalizes.

---

[12] In a brief section on inductive reasoning, Quilty-Dunn and Mandelbaum (2018, 543) suggest a different approach, based on work by Goodman et al. (2015). But there is a problem with this appeal, which Quilty-Dunn and Mandelbaum do not note. The core innovation Goodman et al. claim for their approach is that they supplement a 'language of thought' framework with the tools to perform approximate Bayesian inference. This may be fine, but it is important to recognize that it represents a major departure from the formalist approach: Bayesian inference does not rely on the constituent structure of representations, but on facts about probabilities.

So far as I can see, the best hope for rule-followers at this point would be to appeal to some version of defeasible or 'default' logic (see, for example Horty (2012)). Such logics codify defeasible inference by supplementing classical logical rules with 'default' rules, which admit of exceptions. On this approach Jacinta might indeed be represented as using a rule that licenses an inference from the doorbell's ringing to Amit's being home, but this would be a default rule, admitting of exceptions. Such a conception of the rules of reasoning, however, stretches the idea that reasoning is a matter of following formal rules to breaking point.

For one thing, as everyone acknowledges, Jacinta's inference depends on situation-specific background beliefs and knowledge. As a result, it is possible that the 'rule' she employs is such that it is only ever used once, by a single agent, in the entire history of the universe. It seems reasonable to worry that such 'rules' are rules in name only. Furthermore, consider the question of where the default rules of default logic are supposed to come from. The natural answer is that these rules simply codify some of our substantive knowledge of the situation at hand. For example, the rule 'from the belief that the doorbell is ringing, infer that Amit is home', and its associated 'weight' or 'priority', simply encodes Jacinta's belief that Amit's being home is a more plausible explanation of the doorbell's ringing than, say, fundraisers for LDS having come to call. It is hard to see why the choice to represent this knowledge by means of a default 'rule' should have any deep significance for the theory of reasoning.

So far, I have been using as my target the view that the rules of reasoning must be formal. But, in fact, very similar considerations apply to other approaches as well. As we saw, many rule-following theorists characterize the rules of reasoning as constitutive of our cognitive capacities, concepts, or projects. But what constitutive rules figure in the kind of defeasible inference exemplified by Jacinta? The defeasible pattern 'from the belief that the doorbell is ringing, infer that Amit is home' is not, as such, plausibly constitutive of any of Jacinta's cognitive capacities, concepts, or rationally required projects. What it reflects, rather, is her

relevant empirical beliefs, such as the belief that Amit is dropping off the kids at tennis practice and the like. It is hard to see what role constitutive rules could play here.

Friends of the constitutive approach typically proceed by listing examples of the sorts of rules they have in mind, rather than a clear-cut criterion. This makes it hard to assess how they would respond to the present challenge. To illustrate the difficulties they face, however, let us consider just one of the examples given in the literature, namely, the rule 'from believing that someone said *p*, move to believing *p*' (McHugh and Way 2018, 183). I suppose it is not implausible that a disposition described in these terms could be considered 'constitutive' of our testimonial practices. The trouble is that a rule like this is certainly not sufficient to account for *all* the ways in which we reason with testimony. In particular, what should we say about cases where our disposition to believe another's testimony is defeated? Our responses in such cases might well count as reasoning, too. However, they are clearly not instances of the very same rule; so what rule governs *these* inferences? On the rule-following picture, for each possible testimonial defeater there has to be a corresponding rule—and, on the type of view we are considering, that rule must itself be granted a constitutive status of some sort or other. This seems hard to accept. And, of course, this point is not restricted to testimony: defeaters of *every* sort will need their own constitutive rules.

Some semantic inferentialists might protest that, on their view, it is indeed the case that the totality of our inferential dispositions—including the facts about when they are defeated—*is* constitutive of the contents of our intentional states (e.g., Brandom 1998, 2000). I cannot discuss such views on their merits. What matters for present purposes is just that this type of inferentialism does not support the rule-following views I am considering here, and in fact might be inconsistent with them. This is because semantic inferentialism seems to *presuppose*, rather than explain, the distinction between those of our dispositions that count as 'inferential' and those that are either merely associative or brutely causal. In carrying out the inferentialist program, we are supposed to *start* by identifying the inferences an agent (or a community) is disposed to make

or endorse, and then use them to *determine* the agent's (or community's) concepts. The rule-following views in question, by contrast, proceed in the opposite way: we are *first* supposed to identify the relevant concepts (or other cognitive capacities), and then use facts about their constitution to identify the rules of reasoning.

Now, in addition to default rules, systems of default logic also include rules that prescribe how the default rules are to be combined with and weighed against each other to issue in a recommended response. It is not clear whether such rules—we can call them 'meta-rules'—are meant to be employed by the agents whom the system models, or rather by the theorist who models those agents. But suppose we waive this point, and grant that these meta-rules *are* general rules that govern reasoning. Would this address the problems I have been raising?[13]

This move does have something to recommend it. For one thing, the meta-rules in question are formal in the desired sense: they do not care about the *content* of the representations they apply to, but only about their constituent structure (see, for example, the definition of proper scenarios as 'fixed points' in Horty (2012, 222-223)). For another, it may not be implausible to argue that they are constitutive of our capacity to reason. I do not think, however, that appealing to them helps the dispositionalists' cause. To see this, consider the nature and point of these rules. These are rules that are supposed to operate *on* systems of default rules and factual information, in order to select which defaults to apply. If that is the cognitive activity codified by the rules of reasoning, however, then rule-following does not provide a genuine alternative to the Taking Condition—just a story about how we come to satisfy it.

This is not an *ad hoc* response. To see why, it will help to sketch a picture of reasoning to set against the rule-following one. On this picture, reasoners do not confront deliberation pre-armed with a stock of rules of reasoning. Rather, an essential part of deliberation is *working out*

---

[13] Schroeder (2011), in his response to Dancy's (e.g., 1983; 2004) particularism, appeals to such high-level generalizations about correct deliberation, alluding to his own account in terms of 'weighing' reasons (Schroeder 2007, 137–41). Schroeder does not engage the debate concerning the Taking Condition, so nothing I say here is intended as a criticism of his position.

what, given your relevant background knowledge and beliefs, your current evidence supports. Jacinta does not start out equipped with a rule of inference of the form 'from the fact that the doorbell is ringing, infer that Amit is home'. Instead, drawing on her background knowledge and beliefs, she *assesses* that the doorbell's ringing is good evidence for Amit's being home. This assessment may be very quick and almost automatic on her part; nevertheless, it represents a substantive epistemic achievement, and may be evaluated as such. It is how Jacinta gets to meet the Taking Condition. Moreover, it is this activity that, seen from the point of view of the picture just sketched, the meta-rules of default logic codify: they codify Jacinta's working out what her evidence supports, based on her background knowledge and beliefs (including the part codified in the form of default rules).[14]

Let us take stock. I have not, so far, argued directly against dispositionalist approaches to reasoning. What I have done, instead, is argue that a particularly popular way of carrying out the dispositionalist program—namely, the appeal to rule-following—faces a serious challenge: much of our reasoning does not seem to be reducible to rule-following. This argument can only go so far, of course: perhaps dispositionalists can abandon rule-following altogether. The argument of the next section, by contrast, will take on dispositionalism directly.

4. The challenge of explaining defeat

My aim in this section is to show that defeasibility also poses an *explanatory* challenge for dispositionalism: such theories have difficulty explaining why reasoning is subject to at least

---

[14] Although I cannot argue for this conclusively, I suspect that a similar point would apply to *any* highly abstract sort of rule that purports to subsume defeasible reasoning. Consider, for example, the rule of Bayesian conditionalization. (This is a rule that applies to *credences* rather than full beliefs, but we can let this pass, perhaps on the assumption that belief just is credence above some—perhaps highly context-dependent—threshold.) Like the meta-rules of default logic, it would not be implausible for a proponent of conditionalization to take it to be constitutive of our reasoning faculty. Still, this will not help dispositionalism, because applying conditionalization depends precisely on *deriving the probability of your conclusion on your evidence*. *Prima facie*, at least, this seems to be just a way of satisfying the Taking Condition.

some forms of defeat. [15] My primary focus will again be on the case of theoretical reasoning, but the practical case will be relevant as well.

Consider the following case. Seeing that the lights are on in Nadia's house, Fred infers that Nadia is home. As a result, Fred now believes that Nadia is at home; or, if we adopt a more fine-grained way of speaking, Fred is now *more* confident than before that Nadia is at home. Assume also that Fred has no other evidence directly relevant to the matter at hand. Fred's inference then seems justified. As a result, his belief that Nadia is at home is also justified.

Now, Fred's inference can be defeated in a number of different ways. For one thing, Fred might get evidence against its conclusion. For instance, turning around he might see Nadia standing right behind him. Once Fred has this new evidence, it is of course irrational for him to go on believing that Nadia is in the house. But the new evidence does not necessarily cast doubt on the connection between the lights' being on and Nadia's being home: Fred may well be correct in still thinking that the lights' being on was good (though not conclusive) grounds from which to infer that Nadia was home. Defeaters that work in this way—as evidence directly against the conclusion of an inference—are 'rebutting' defeaters.

Suppose, however, that instead of evidence that Nadia is not in the house Fred got evidence of the following kind: Nadia, who usually lives alone, has family visiting, so that for a few days there are multiple people coming and going at her house. This is not direct evidence that Nadia is not at home. Simply learning that Nadia has family staying with her should not lead Fred to think it any less likely than before that Nadia is at home. It is, however, evidence that someone else, other than Nadia, may be responsible for the lights being on in Nadia's house.

---

[15] Neta (2019) draws on a somewhat similar set of considerations to argue against dispositional accounts of the basing relation. The position that Neta defends is stronger than the one I defend in this paper. On Neta's view, the content of the 'takings' required by the Taking Condition must involve a component referring *de re* to the particular instance of the basing relation at issue. For all I say in this paper, this stronger position may be correct; there is no space here to distinguish between different versions of the Taking Condition.

And this, in turn, means that the lights' being on is no longer good (or, *as* good) evidence that Nadia is in the house. As a result, it functions as a defeater of Fred's inference. If Fred gets such evidence, and assuming that he has no independent evidence regarding Nadia's whereabouts, it would be irrational for him not to become less confident that Nadia is at home than he would otherwise.[16] Defeaters of this sort, which function as evidence against the connection between the premises and the conclusion of an inference, are known as 'undercutting' or 'undermining' defeaters. This is the kind of defeat I will focus on in what follows.

Could one deny that evidence of this sort undermines the rationality of our inferences? Some authors have suggested that, at least in some cases or in some sense, it may still be rational to reason in the face of adverse 'higher-order' evidence (Christensen 2007; Lasonen-Aarnio 2014). Higher-order defeat (or would-be defeat) is not always clearly distinguished from undercutting defeat, but for present purposes it is a good idea to keep them apart. Roughly speaking, the debate over higher-order defeat concerns evidence that the agent's epistemic performance was *defective* in some relevant way. One might wonder, for example, about the rational impact of Fred's learning that he has unwittingly ingested a drug likely to make him misjudge even the simplest evidential relations: can he still rationally infer that Nadia is at home? For present purposes, I believe we can set such cases aside. Notice that such higher-order evidence is not, intuitively, evidence against *either* your conclusion *or* the evidential connection between it and your premises. For example, evidence that Fred has ingested a drug impairing his reasoning abilities seems to be neither evidence that Nadia is not at home nor that her lights' being on would not be a good basis from which to infer that she is at home.[17] Thus, even if it

---

[16] Tucker (2012, 338) suggests that an *unjustified* belief that *P* supports *Q* can be sufficient for rationally inferring *Q* from *P*. But he gives no direct argument for this position, and our intuitive verdicts in cases of undercutting defeat would seem to speak against it.

[17] This feature of higher-order defeat is noted by Christensen (2010), van Wietmarschen (2013) and Silva (2017). Other authors in the literature do not seem to agree on this point—Horowitz (2014, 719), for example, suggests that evidence of temporary cognitive malfunction is evidence that your conclusions are

turns out that Fred *could* be rational in making the inference even in the face of such adverse higher-order evidence, this need have no bearing on the other types of defeat introduced earlier.

Having distinguished between these different kinds of defeat, I can come to the core of my argument. The fact that reasoning is subject to undercutting defeat is strong evidence for the Taking Condition, and thus against dispositionalist approaches to reasoning.

To see this, consider the following question: *why* is reasoning subject to undercutting defeat? *Rebutting* defeat seems straightforward to understand. If I reason from P to Q in the face of evidence that Q is not, in fact, true, then it seems clear that I am involved in some sort of irrationality: believing against my evidence, perhaps, or even outright inconsistency (if I also believe not-Q). Undercutting defeat, by contrast, seems trickier. If D is an undercutting defeater for the inference from P to Q, then, by hypothesis, D is not evidence against Q. It is, rather, evidence against the *evidential connection* between P and Q. Indeed, we can imagine that, on balance, your *total* evidence still strongly supports Q. So, why exactly should it be irrational, in the circumstances, for you to infer Q from P?

Proponents of the doxastic approach can appeal to the Taking Condition to explain this, in the following way:

> Reasoning from P to Q requires taking it that P supports Q. It seems plausible, therefore, that if you cannot *rationally* take it that P supports Q, you cannot *rationally* infer Q from P. Thus evidence against the claim that P supports Q undercuts your ability to rationally infer Q from P. [18]

---

not supported by your evidence. If Horowitz is right, however, then the intuition that we can still reason rationally in the face of adverse higher-order evidence should fade.

[18] Schroeder (2007, 137-139; 2011, 337-338) offers an explanation of undercutting in terms of the weight of reasons. The idea is that undercutting defeaters are reasons to 'place less weight' upon other reasons in deliberation. This account is fine so far as it goes, but it falls short in one important respect: it rests on the unexplained idea of 'placing weight' on something in deliberation. The suggestion here is similar in spirit, but avoids this problem: undercutting defeaters are simply reasons *not* to believe evidential claims (which,

If we reject the doxastic approach and the Taking Condition, however, then it looks like you can be rational in inferring Q from P *without* taking it that (and so without *rationally* taking it that) P supports Q at all. Thus, this straightforward route for explaining the influence of the defeating evidence is no longer available. So how is undercutting defeat to be understood, on the dispositionalist picture?[19]

One response I have encountered goes as follows.[20] An undercutting defeater D, on this picture, does not simply undercut *your justification for taking it* that P supports Q; it *also* undercuts the evidential connection between P and Q itself. Given D, in other words, P no longer gives you even *ex ante* reason for Q. This is why it is not rational for you to infer P from Q, in the face of undercutting defeater D.[21] Clearly, if we were to accept this explanation of undercutting defeat, that would be a boon to dispositionalists, as it does not depend on the truth of the Taking Condition. But I do not think that this explanation is acceptable.

---

as I suggested earlier, can be understood as epistemic conditionals such as 'if the lights are on, then Nadia is home').

[19] An anonymous referee suggests that perhaps dispositionalists can simply take undercutting defeat as a basic fact about rationality. This, I suppose, is always an option. However, it seems to me that the point made in this paragraph should still be cause for some embarrassment for dispositionalists. If you have no opinion on a certain matter, and in fact are not required to have an opinion on it, then it seems like any evidence pertaining to it should be normatively irrelevant to you. If this is so, the challenge is not simply that dispositionalism may struggle to explain undercutting defeat; even more strongly, dispositionalism actively makes undercutting defeat *surprising*. I do not take this by itself to be dispositive reason against dispositionalism. I do, however, take it to show that shirking the explanatory burden is not quite so easy for dispositionalists.

[20] I thank an anonymous referee for pressing me to consider this response. A very similar response was also suggested to me by [REDACTED].

[21] Something like this account seems to be endorsed by Wedgwood (2012). According to Wedgwood, when an inference is defeated this is because 'the believer is somehow committed, by his or her other beliefs, to having only a low degree of conditional confidence in the conclusion of that [inference], conditional on the relevant premises' (2012, 280). Wedgwood makes it clear that this is intended to apply both to rebutting and undercutting defeat.

To begin with, it is not obvious that the existence of good reasons and the rationality of inferences always march in step in this way. More specifically, it seems at least *prima facie* possible for inferences to be defeated by *misleading* undercutting evidence. In such a case, it might be that your inference from P to Q is defeated, even though P really *is* good reason to believe Q.[22]

The general point is familiar primarily from discussions of *practical*, rather than theoretical, rationality: a fire in the hotel's basement that Zahra is entirely unaware of may constitute a reason for her to get out quickly; but no rational inferential route to that course of action seems available to her. If such mismatches are possible, then we can also construct cases where otherwise rational inferences are defeated by misleading undercutting defeaters. Suppose that Zahra, still in her hotel room, hears the fire alarm go off. Normally, needing to get out of a multi-story building quickly makes it rational for Zahra to use the elevator. But Zahra also believes that, in case of fire, taking the elevator is unsafe. This is an undercutting defeater of her usual style of reasoning. As it happens, however—and although Zahra is unaware of this—in the present case taking the elevator *would* be safe, and in fact safer than any other course of action, such as taking the stairs or staying in place. It is at least arguable that in these circumstances hearing the fire alarm *does* give Zahra reason to take the elevator. (Those inclined to say that the fire in the basement was a reason for Zahra to get out in a hurry, at least, would seem to have no grounds to reject this.) Still, it seems clear that it would be *irrational* for Zahra to infer, from the fact that the fire alarm is sounding, that she should take the elevator. If this description of the case is correct, then it is a counterexample to the explanation of undercutting defeat in question here.

Now, of course, this result might be taken as grounds to *reject* the underlying picture of reasons that allows for such mismatches between what you have reason to do or believe and

---

[22] According to some, defeat by misleading undercutting evidence might be impossible, as justified false beliefs about rationality are impossible (e.g., Smithies (2012), Titelbaum (2015) and Littlejohn (2018)). This is a controversial view (Field 2017); but, as we will see below, even if correct it does not suffice to rescue the suggested account of undercutting defeat.

what inferences are rationally available to you. Perhaps what we should conclude from the fact that Zahra has no rational inferential route to choosing to take the elevator is that the fire alarm does not really give her reason to take the elevator.[23] Such a move would, indeed, avoid the threat of defeat by misleading undercutting evidence. But notice what we would be doing if we were to endorse this move. Rather than drawing on supposedly more basic facts about reasons in order to *explain why* Zahra has no rational inferential route to choosing to take the elevator, we would be doing the reverse: we would be using our understanding of what inferences Zahra can rationally perform in order to determine what she does or does not have reason to do. But then we cannot simply turn around and appeal to facts about her reasons in order to explain what inferences are or are not available to her. In other words, the price we need to pay to rule out misleading undercutting defeaters seems to be the *explanatory independence* of the facts about what reasons you have from the facts about what inferences are rationally available to you. Without this independence, however, the explanation of undercutting defeat in question cannot get off the ground. Either way, then, this explanation is not acceptable.

Returning to our usual focus on the theoretical case, we are still faced with the following question: why should evidence against the evidential connection between P and Q make it irrational for you to infer Q from P, if not because inferring Q from P requires taking it that P supports Q? So far as I can tell, McHugh and Way are the only dispositionalists to explicitly address a challenge in this vicinity. More specifically, they (2018, 191) discuss the apparent incoherence of inferring Q from P while also believing that P does not support Q. Their response is based on the idea that reasoning is a goal-directed activity, which aims at the goal of acquiring 'fitting attitudes'—which, in the case of belief, means *true* beliefs. Thus, they write:

> If P does not support Q, then reasoning from P to Q is not a good way to pursue this
>
> aim. So, reasoning from P to Q while judging that P does not support Q amounts to

---

[23] For a thorough discussion of this issue that ends up endorsing this conclusion see Gibbons (2013).

taking what you acknowledge to be an unreliable means to your end. That looks plainly

irrational. (ibid.)

This response could easily be adapted to our discussion here. Suppose we grant the following: if

D is undercutting evidence for the inference from *P* to *Q* then, once you get evidence *D*, there is

a sense in which are *committed to* (though you do not *believe*) the claim that *P* does not support *Q*.

On this assumption, McHugh and Way's account would go just as above: reasoning from *P* to *Q*

while being committed to the claim that *P* does not support *Q* would, again, appear to be a case

of 'taking what you acknowledge to be an unreliable means to your end', and that, according to

McHugh and Way, is 'plainly irrational'.

The trouble, however, is that this last claim is false. In general, it is perfectly possible for

you to take means *M* to end *E* while believing (or having evidence to believe) that *M* is an

unreliable means to *E*, without being irrational. This can happen if, through no fault of your

own, you *do not recognize that you are taking M as a means to E*. Suppose that I am lost in the woods

on a dark night. Believing that my destination is due north and trusting my compass, I set off in

the direction I believe to be due north. Unbeknownst to me, however, a powerful magnetic field

in the area is messing with my compass. As a result, I happen to be walking due south, as a

means to my end of walking due north—a means that I have plenty of evidence to think is

unreliable, and which in fact I *do* believe is unreliable. Nevertheless, my actions in this case are

not irrational. In order for my actions to be irrational, it seems, I would have to *know* (or at least

believe) *that* I am walking south as a means to reaching my destination in the north.

For McHugh and Way's verdict of irrationality to stick, therefore, it must be the case

that, when reasoning from *P* to *Q*, I recognize that I am reasoning from *P* to *Q* as a means to

acquiring fitting beliefs. But can McHugh and Way require this? On their view, it cannot be the

case that reasoning from *P* to *Q* requires anything like believing that *P* supports *Q*. This, after all,

is just what distinguishes dispositionalist from doxastic accounts of reasoning. The trouble is

that, as we just saw, their understanding of undercutting defeat seems to commit them to the

claim that reasoning from *P* to *Q* requires recognizing that very reasoning as my means to acquiring fitting (that is, true) beliefs. But this just looks like another way of saying that it requires believing that *P* supports *Q*.[24]

Is there, then, some other way in which dispositionalists could respond to the puzzle of undercutting defeat? It is here, I think, where the argument of the previous section against rule-following conceptions of reasoning becomes relevant again.

What dispositionalists need, it seems, is a kind of *acceptance* of an inference, which falls short of belief but which can, like belief, be rationally undermined by contrary evidence. Such a combination of features may seem *ad hoc*. However, it has been argued that, on pain of radical skepticism, our epistemic relation to some basic patterns of inference—variously including basic logical rules, induction, and inference to the best explanation—*must* take this form (e.g., Wright 2004; Enoch and Schechter 2008). The idea is that we have some kind of default or 'unearned' justification to *rely* on such patterns of inference, absent evidence to the contrary (where reliance involves a disposition to infer in accordance with those patterns, but not the belief that the patters are good either in general, or in any particular case). Indeed, I suspect that many dispositionalists would gravitate towards something like this idea of default but defeasible justification as an answer to the puzzle of undercutting defeat.

The argument of section 3, however, strongly suggests that this will not work. Appeal to default justification might help, *if it were possible to reduce defeasible reasoning to a well-demarcated set of patterns*—those that correspond, say, to rules constitutive of our cognitive capacities or of rationally required projects. What the argument of section 3 suggests, however, is that no such reduction is likely forthcoming. And in the absence of such reduction, the appeal to default justification looks unmotivated and implausible. In defeasible inferences, the connection between premises and conclusion relies on highly contingent and local facts about the reasoner's circumstances—facts which it is extremely implausible that we have default justification to either

---

[24] For another response to McHugh and Way's view that is broadly similar in spirit, see Hlobil (2018).

believe or take for granted. Consider again the stock example of Tweety and his flying abilities. As we saw in section 3, it is extremely implausible that our ability to reason correctly in this could be reduced to more basic inferential capacities (perhaps logical ones). Rather, we simply have to draw on your ornithological knowledge to judge that, despite the fact that he is a bird, Tweety cannot fly. This is not the sort of thing that we could plausibly have default justification for.[25]

So, where does all this leave us? I will close this paper with a brief look at what I think are the main lessons for the theory of reasoning.

5. Conclusion

If what I have argued so far is correct, then dispositionalist approaches to reasoning do not look promising. But, it seems fair to ask, why think that the doxastic approach could do better? The answer is straightforward, at least in outline. The problem defeasibility poses for dispositonalism fundamentally comes down to this: defeasible inferences depend on a substantive assessment of what conclusion is supported or warranted in a given circumstance. And this is exactly what the Taking Condition asserts is the case across the board: inferring requires *taking it that* your response is warranted. From the present point of view, therefore, a pressing question for the theory of reasoning is precisely to account for our ability to assess evidential relations (or, taking into account the practical case as well, to assess what is a reason for what).

At this point, of course, suspicions of circularity or regress doubtlessly arise. Doesn't the claim that reasoning from *P* to *Q* requires a substantive assessment of whether *P* supports *Q*

---

[25] To be clear, this is not an argument against default justification *as such*; it may, for example, have an important role to play in the epistemology of perception, for example. The point is just that it cannot take over from the Taking Condition in explaining undercutting defeat.

amount to the claim that reasoning from *P* to *Q* requires *more reasoning*? If I am correct, however, and dispositionalist theories are not viable, there must be a way to resist this worry.[26]

I believe that we can do this by distinguishing between two rather different things that are often conflated under the umbrella terms 'reasoning', 'inference', or 'deliberation'. On the one hand, we have the phenomenon that has been the topic of this paper—namely, that of drawing a conclusion from a set of premises, or basing a belief or intention on some other mental states. On the other hand, we have the phenomenon of assessing relations of support. These are not obviously the same thing. And if these really are not the same thing, then not only does the threat of circularity vanish, but the dispositionalist approach is revealed as *obscuring* an interesting and pressing epistemological question—namely, *how*, if not by reasoning, do we get to assess relations of support or warrant?

This is somewhat abstract, but it can be made more concrete. Consider a possible worlds model of belief, on which the worlds that are epistemically possible for an agent are subject to a 'plausibility' ordering, derived from the agent's background beliefs. For the purposes of this model, then, the claim that a premise *P* supports a conclusion *Q* (relative to the agent's current doxastic state) can be understood as follows: *Q* is true in all the most plausible *P*-worlds. Given such a model, it seems very natural to identify the agent's *taking it that Q* is supported by *P* with a

---

[26] A different regress worry arises from considering what role, if not that of an extra premise, beliefs about support could play in reasoning (e.g., Carroll 1895; Boghossian 2003). I cannot give a full response to this worry here, but I can outline the sort of response I favor [REDACTED]. The idea is this: if I believe *P*, and also recognize that *Q* follows from *P*, then there are no ways for things to be that are 'real' or 'live' possibilities for me are consistent with *P*'s failing to be true. In a sense, then, my reasoning is *done*, once I have come to see that *Q* follows from *P* (and assuming that I still believe *P*). On this view, therefore, recognizing the connection between *P* and *Q* does not provide me with an extra premise for my reasoning; rather, it constitutes my coming to believe *Q* on the basis of *P*. I thank an anonymous referee for pressing me to address this question. Clearly, more needs to be said, especially about how to extend the approach to the practical case.

conditional belief, whose content can be expressed as '*Q*, given *P*'. [27] The challenge, then, comes down to this: *how*, if not by reasoning, an agent can form conditional beliefs of this sort.

This challenge seems far from insurmountable. This may be easier to see if we paraphrase it thus: how can an agent *elaborate* on a possibility, specified by a certain condition (in our present case, 'the most plausible *P*-worlds'), in order to determine whether a hypothesis is true in that possibility? There can be many ways to give a detailed account of this activity, and no reason to think that all cases must be treated the same way. Sometimes, we may use imaginative mental simulation, as Williamson suggests for the case of counterfactual conditionals (Williamson 2008). In other cases, perhaps, we might simply 'see' whether the most plausible *P*-worlds happen to be *Q*-worlds or not. Such intuitions are, of course, not magic. Perhaps they are cases of mental simulation too, except that in those cases the mental simulation happens automatically, at the sub-personal level, as suggested by the 'mental model' theory of reasoning in psychology (Johnson-Laird 1983; Johnson-Laird and Byrne 1991). In earlier work

---

[27] This type of model is standard in the belief-revision literature (Grove 1988; Gärdenfors 1990). Plausibility is assumed to be connected, transitive, and to obey the so-called 'limit assumption' (Lewis 1973, 19–21). My goal here is not to provide a complete formal model of rational belief. I introduce this model just for illustrative purposes. Still, it is worth noting that this model is capable of modelling rebutting and undercutting defeaters, as follows. A proposition *D* will be a *defeater* of the inference from *P* to *Q* just in case (i) *P* supports *Q* in this sense, while (ii) *P&D* does *not* support *Q*. This will be the case just in case all the most plausible *P*-worlds are *Q*, while among the most plausible *P&D*-worlds there are at least some worlds in which *Q* is false. The distinction between rebutting and undercutting defeaters, now, can be understood as follows. If *D* is a rebutting defeater then *D* supports ~*Q*: i.e., all the most plausible *D*-worlds are ~*Q*-worlds. For example, while the most plausible (from Fred's point of view) worlds in which the lights are on in Nadia's house are such that Nadia is at home, the most plausible worlds in which Nadia is standing right behind Fred are such that Nadia is *not* at home. The same is not true for undercutting defeaters. If *D* is an undercutting but not rebutting defeater, then among the most plausible *D*-worlds there will be some in which *Q* is still true, and so *D* does not support ~*Q*. For example, among the most plausible worlds in which Nadia has relatives staying at her house, there will be some which are such that Nadia is at home.

[REDACTED] This is not the place to discuss any particular suggestion in detail. The point is simply that there is a substantive question here, which a bland appeal to dispositions obscures.

What about practical reasoning? I have had almost nothing to say about specifically practical reasoning in this paper, but evidently the same kind of question arises here as well: are there ways, other than reasoning, to recognize that a given set of considerations supports some course of action? I think the answer to this question should, again, be 'yes'. We are all familiar with the fact that, given an actual or hypothetical scenario, we can have intuitions about what the right (or the wrong) thing to do is. There are ongoing debates, of course, about how reliable such intuitions are, especially in complex ethical scenarios. But the fallibility of our intuitions in such scenarios is beside the point here. All we need is the claim that we often judge non-inferentially about what is a reason for what, in the practical realm. This is enough to give us the beginnings of a response to the threat of circularity for the case of practical reasoning. And, of course, this is not meant to be the final word on this topic; my point, rather, is once again that this is a direction for investigation whose importance is lost sight of, if we respond to threats of circularity or regress by adopting dispositionalism.

References

Bird, Alexander. 2005. "Abductive Knowledge and Holmesian Inference." In *Oxford Studies in Epistemology 1*, edited by Tamar Gendler and John Hawthorne, 1–31. 1. Oxford University Press.

Boghossian, Paul. 2002. "How Are Objective Epistemic Reasons Possible?" In *Reason and Nature*, edited by José Bermudez and Alan Millar, 1–47. New York: Oxford University Press.

———. 2003. "Blind Reasoning." *Aristotelian Society Supplementary Volumes* 177: 225–48.

———. 2008. "Epistemic Rules." *Journal of Philosophy* 105 (9): 472–500.

———. 2014. "What Is Inference?" *Philosophical Studies* 169 (1): 1–18. https://doi.org/10.1007/s11098-012-9903-x.

Brandom, Robert. 1998. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.

———. 2000. *Articulating Reasons: An Introduction to Inferentialism*. Harvard University Press.

Brewer, Bill. 1995. "Mental Causation II: Compulsion by Reason." *Aristotelian Society Supplementary Volumes* 69: 237–53.

Broome, John. 2013. *Rationality through Reasoning*. Chichester: Wiley Blackwell.

———. 2014. "Comments on Boghossian." *Philosophical Studies* 169 (1): 19–25. https://doi.org/10.1007/s11098-012-9894-7.

Buckner, Cameron. 2019. "Rational Inference: The Lowest Bounds." *Philosophy and Phenomenological Research* 98 (3). https://doi.org/10.1111/phpr.12455.

Carroll, Lewis. 1895. "What the Tortoise Said to Achilles." *Mind* 4: 278.

Christensen, David. 2007. "Does Murphy's Law Apply in Epistemology? Self-Doubt and Rational Ideals." In *Oxford Studies in Epistemology 2*, edited by Tamar Gendler and John Hawthorne, 2:3–31. 2. Oxford University Press.

———. 2010. "Higher-Order Evidence." *Philosophy and Phenomenological Research* 81 (1): 185–215. https://doi.org/10.1111/j.1933-1592.2010.00366.x.

Dancy, Jonathan. 1983. "Ethical Particularism and Morally Relevant Properties." *Mind* 92 (368): 530–47.

———. 2004. *Ethics Without Principles*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/0199270023.001.0001/acprof-9780199270026-chapter-5.

———. 2014. *From Thought to Action*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198709299.001.0001/acprof-9780198709299-chapter-1.

Enoch, David, and Joshua Schechter. 2008. "How Are Basic Belief-Forming Methods Justified?" *Philosophy and Phenomenological Research* 76 (3): 547–579.

Field, Claire. 2017. "It's OK to Make Mistakes: Against the Fixed Point Thesis." *Episteme*, 1–11.

Gärdenfors, Peter. 1990. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. Cambridge, MA: MIT Press.

Gibbons, John. 2013. *The Norm of Belief*. Oxford: Oxford University Press.

Goodman, Noah, Joshua Tenenbaum, and Tobias Gerstenberg. 2015. "Concepts in a Probabilistic Language of Thought." In *The Conceptual Mind: New Directions in the Study of Concepts*, edited by Eric Margolis and Stephen Laurence, 623–54. Cambridge MA: MIT Press.

Grove, Adam. 1988. "Two Modellings for Theory Change." *Journal of Philosophical Logic* 17 (2): 157–170.

Hlobil, Ulf. 2018. "Goodness-Fixing Isn't Good Enough: A Reply to McHugh and Way." *Mind* Online First.: n.a. https://doi.org/10.1093/mind/fzy035.

———. 2019. "Inferring by Attaching Force." *Australasian Journal of Philosophy* 0 (0): 1–14. https://doi.org/10.1080/00048402.2018.1564060.

Horowitz, Sophie. 2014. "Epistemic Akrasia." *Noûs* 48 (4): 718–44. https://doi.org/10.1111/nous.12026.

Horty, John F. 2012. *Reasons as Defaults*. New York: Oxford University Press.

Johnson-Laird, Philip. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Harvard University Press.

Johnson-Laird, Philip, and Ruth Byrne. 1991. *Deduction*. Hove, UK ; Hillsdale, USA: Psychology Press.

Johnston, Mark. 1988. "Self-Deception and the Nature of Mind." In *Perspectives on Self-Deception*, edited by Brian McLaughlin and Amelie Rorty. Berkeley and Los Angeles: University of California Press.

Lasonen-Aarnio, Maria. 2014. "Higher-Order Evidence and the Limits of Defeat." *Philosophy and Phenomenological Research* 88 (2): 314–45. https://doi.org/10.1111/phpr.12090.

Leite, Adam. 2008. "Believing One's Reasons Are Good." *Synthese* 161 (3): 419–41.

Lewis, David K. 1973. *Counterfactuals*. Blackwell Publishers.

Littlejohn, Clayton. 2018. "Stop Making Sense? On a Puzzle about Rationality." *Philosophy and Phenomenological Research* 96 (2): 257–72. https://doi.org/10.1111/phpr.12271.

Mantel, Susanne. 2017. "Three Cheers for Dispositions: A Dispositional Approach to Acting for a Normative Reason." *Erkenntnis* 82 (3): 561–582.

Marcus, Eric. forthcoming. "Inference as Consciousness of Necessity." *Analytic Philosophy*.

———. 2012. *Rational Causation*. Harvard University Press.

McHugh, Conor, and Jonathan Way. 2018. "What Is Reasoning?" *Mind* 127 (505): 167–96. https://doi.org/10.1093/mind/fzw068.

Neta, Ram. 2013. "What Is an Inference?" *Philosophical Issues* 23 (1): 388–407. https://doi.org/10.1111/phis.12020.

———. 2019. "The Basing Relation." *The Philosophical Review* 128 (2): 179–217.

Norton, John. 2003. "A Material Theory of Induction." *Philosophy of Science* 70: 647–70.

Railton, Peter. 2006. "How to Engage Reason: The Problem of Regress." In *Engaging Reason: Themes from the Moral Philosophy of Joseph Raz*, edited by Ray Wallace, Philip Pettit, Stephen Scheffler, and Michael Smith. Oxford: Oxford University Press.

Rescorla, Michael. 2009. "Predication and Cartographic Representation." *Synthese* 169 (1): 175–200.

Rumfitt, Ian. 2011. "Inference, Deduction, Logic." In *Knowing How: Essays on Knowledge, Mind, and Action*, edited by John Bengson and Marc A. Moffett, 334. Oxford University Press, Usa.

———. 2015. *The Boundary Stones of Thought: An Essay in the Philosophy of Logic*. New York: Oxford University Press.

Ryle, Gilbert. 2002. *The Concept of Mind*. Chicago: University of Chicago Press.

Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford University Press.

———. 2011. "Holism, Weight, and Undercutting." *Noûs* 45 (2): 328–44.

Silva, Paul. 2017. "How Doxastic Justification Helps Us Solve the Puzzle of Misleading Higher-Order Evidence." *Pacific Philosophical Quarterly* 98 (S1): 308–28. https://doi.org/10.1111/papq.12173.

Skow, Bradford. 2016. *Reasons Why*. Oxford: Oxford University Press.

Smithies, Declan. 2012. "Moore's Paradox and the Accessibility of Justification." *Philosophy and Phenomenological Research* 85 (2): 273–300.

Sosa, Ernest. 2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I*. Oxford University Press.

Titelbaum, Michael G. 2015. *Rationality's Fixed Point (or: In Defense of Right Reason)*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198722762.001.0001/acprof-9780198722762-chapter-9.

Tucker, Christopher. 2012. "Movin' on up: Higher-Level Requirements and Inferential Justification." *Philosophical Studies* 157 (3): 323–40. https://doi.org/10.1007/s11098-010-9650-9.

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

———. 2016a. "What Reasoning Might Be." *Synthese*, 2007–24. https://doi.org/10.1007/s11229-016-1034-z.

———. 2016b. "What The Tortoise Has To Say About Diachronic Rationality." *Pacific Philosophical Quarterly*, Early View. https://doi.org/10.1111/papq.12172.

———. 2018. "Reasoning and Deducing." *Mind* Online pre-publication. https://doi.org/10.1093/mind/fzy025.

Wedgwood, Ralph. 1999. "The A Priori Rules of Rationality." *Philosophy and Phenomenological Research* 59 (1): 113–31. https://doi.org/10.2307/2653460.

———. 2006. "The Normative Force of Reasoning." *Noûs* 40 (4): 660–86.

———. 2012. "Justified Inference." *Synthese* 189 (2): 273–95. https://doi.org/10.1007/s11229-011-0012-8.

Wietmarschen, Han van. 2013. "Peer Disagreement, Evidence, and Well-GroundednessPeer Disagreement, Evidence, and Well-GroundednessHan van Wietmarschen." *The Philosophical Review* 122 (3): 395–425. https://doi.org/10.1215/00318108-2087654.

Williamson, Timothy. 2008. *The Philosophy of Philosophy.* 1 edition. Malden, MA: Wiley-Blackwell.

Winters, Barbara. 1983. "Inferring." *Philosophical Studies* 44: 201–20.

Wright, Crispin. 2004. "On Epistemic Entitlement (I): Warrant for Nothing (and Foundations for Free)?" *Aristotelian Society Supplementary Volumes* 79: 168–211.