CrossMark

Meta-epistemic defeat

J. Adam Carter[1]

Abstract An account of *meta-epistemic defeaters*—distinct from traditional (first-order) epistemic defeaters—is motivated and defended, drawing from case studies involving epistemic error-theory (e.g., Olson 2011, Reasons for belief; cf., Streumer 2012, J Philos 110:1–25) and epistemic relativism (e.g., MacFarlane 2005, Oxford Stud Epistemol 1:197–233; 2011; 2014, Assessment sensitivity: Relative truth and its applications). Mechanisms of traditional epistemic defeat and meta-epistemic defeat are compared and contrasted, and some new puzzles are introduced.

1 Introduction

Most contemporary discussions of epistemic defeat focus on *first-order* epistemic defeaters. A defeater defeats at (merely) the first-order if it downgrades the epistemic status of a belief in a way that is neutral with respect to the *meta-epistemic* status of the belief. The *meta-epistemic* status of a belief is the status it has not in virtue of the *particular* belief it is, but in virtue of its being the *kind of attitude* that it is, whereas, the (first-order) epistemic status of a belief is just the epistemic status it has in virtue of the particular belief it is (e.g., formed in the fashion that it was, on the basis of the reasons for which it was held, etc.).

✉ J. Adam Carter
  j.adam.carter@ed.ac.uk

[1] University of Edinburgh, Edinburgh, UK

Examples of first-order epistemic status, the kind of status that stands subject to traditional (e.g., rebutting and undercutting[1]) first-order defeaters, include knowledge and justification[2]. Examples of second-order epistemic status are things like being *objectively* justified, being *objectively* known, viz., status that a given belief has in virtue of the kind of attitude it is, however, not in virtue of the particular belief it is (e.g., and so not in virtue of its being formed as it was, on the basis of the reasons for which it was held, etc.).

Obviously, first-order epistemic status matters. We care about whether our beliefs are justified, known, etc. But meta-epistemic status matters, too. As Paul Boghossian (2001; 2006) argues[3], when we assess—in epistemology as well as in ordinary circumstances—whether a given belief is justified or known, we generally take it that that there is an objective fact of the matter, one that is potentially knowable[4]. And so we—most of us, anyway[5]—take it that the first-order epistemic status our beliefs enjoy is *objective* epistemic status, viz., objective justification and knowledge. Accordingly, just as we take it to be a kind of epistemic loss if (in the presence of a defeater) a given belief, *p*, is no longer justified or known, it is also natural to suppose we've suffered some sort of epistemic loss if, upon acquiring some new belief *q*, we regard the objectivity of our justification for believing *p* to be imperiled or vanquished[6].

The rich literature on defeaters in epistemology is a testament to the fact that we care about the various mechanisms under which first order status can be downgraded, and potentially regained. But given that meta-epistemic status matters, too, it's curious that the defeasibility of *this* kind of status has been comparatively overlooked. *How* exactly does meta-epistemic status get downgraded, when it does? What are the mechanisms by which this could happen, and how do these relate to the mechanisms by which first-

---

[1] The traditional distinction, between *rebutting* and *undercutting* defeaters is due to Pollock (1986), and this distinction concerns whether the defeater in question defeats by supporting the negation of the target proposition, in the rebutting case, as opposed to indicating that the grounds of the target belief do not sufficiently support its truth, in the undercutting case. Another important distinction, which has been highlighted by (among others) Lackey (1999, 2005), concerns *psychological* (or, mental state defeaters) versus *normative* defeaters, where the latter are beliefs that one (epistemically) should have, regardless of whether one in fact has them. For an overview of these differences, see Carter (2015). For general discussion, and some additional taxonomies of defeat, see Sudduth (2008).

[2] The defeat of the former asymmetrically entails the defeat of the latter. Whereas psychological or mental state defeaters are capable of defeating either justification or knowledge, it's debatable whether 'propositional defeaters' suffice to defeat justification. For discussion on this point, see Sudduth (2008). Cf., Baker-Hytch and Benton (2015) for a recent statement of opposition to regarding knowledge as subject to a defeater-framed condition.

[3] Most influentially, in his 2006 monograph *Fear of Knowledge.* See also Siegel (2013) and Seidel (2014).

[4] Boghossian (2006, p. 76) adds that 'Whenever we confidently judge that some belief is justified on the basis of a given piece of information, we are tacitly assuming that such facts are not only knowable but that they are known. And in doing epistemology, we not only assume that they are knowable, we assume that they are knowable a priori'. For related remarks, see Boghossian (2001, p. 2). For more general defences of the view that epistemologists often (in their first-order projects) typically tacitly assume a minimal kind of meta-epistemic realism, see Carter (2016, Ch. 1). Cf., Cuneo (2007).

[5] The epistemic relativist demurs here. See, for example, Kusch (2010).

[6] What this involves will be clarified in Sect. 4.

order epistemic status gets downgraded? More generally: how does meta-epistemic defeat work[7]?

This paper will attempt to examine this general question through two case studies involving two very different positions in meta-epistemology: error theory and relativism. Here is the plan. Section 2 discusses what I will call *error-theoretic meta-epistemic defeaters*. Section 3 discusses *relativist meta-epistemic defeaters*. Though there are some interesting differences between how these two kinds of meta-epistemic defeaters work, there are also important commonalities. In Sect. 4, I'll summarise these commonalities and taxonomise how meta-epistemic defeaters interface with familiar distinctions in the literature on (first-order) defeaters, raising some new puzzles in the process.

2 Error-theoretic meta-epistemic defeaters

Philosophers who are reluctant to agree that claims about knowledge and justification attain objective epistemic status generally agree on a much more basic point: that there *are* epistemic reasons for beliefs. For example, epistemic relativists such as Rorty (1980) countenance epistemic reasons. Rorty just thinks that the epistemic authority that epistemic reasons aspire to is confined to the social contexts giving rise to them, a point that absolutists like Boghossian deny. But *neither* denies that there *are* epistemic reasons.

Enter here the epistemic error theorist. Think of epistemic error-theory as analogous to moral error-theory[8]. The moral error theorist (e.g., Mackie 1977) draws our attention to the queerness of moral properties, which would be "qualities or relations of a very strange sort, utterly different from anything else in the universe" (1977, p. 38). We have, Mackie thinks, no reason to think moral properties are ever actually instantiated in the world, and thus, that there are no facts consisting in the instantiation of such properties. All statements that attribute moral properties (e.g., including attributions of moral reasons) are false.

The meta-epistemic analogue to this position—epistemic error-theory (e.g, Olson 2011) regards epistemic properties as in the same boat as moral properties. Thus, as Nishi Shah (2011, p. 96) puts it:

> ... Just as no persons have the property of being a witch, so [according to the epistemic error theorist] there are no facts or states of affairs that have the property of being a reason for belief. Any judgment that attributes to something the property of being a reason for belief is, therefore, false.

Epistemic error-theory is of course an odd view; if it's true, it means that our practice of attributing epistemic reasons is fundamentally mistaken[9]. As Bart Streumer (2012)

---

[7] Of course, one might initially reject a presupposition of this question by suggesting there is no such interesting epistemological phenomenon. I motivate the case that there is in Sects. 2 and 3 .

[8] For a comprehensive discussion of meta-epistemological parallels to meta-ethics, see Cuneo (2007, Ch. 3).

[9] Note that epistemic error theory is, in principle, a view that concerns epistemic properties *tout court,* and not just epistemic reasons attributions. As it so happens, most discussions of epistemic error-theory have

has argued, it might not even be possible to *believe* this view, regardless of whether it is true[10]. And, as Crispin Wright (1994) suggests, the truth of epistemic error theory would seem to call into doubt the sensicality of continuing to engage seriously with epistemological questions[11]. But let's set all of this aside.

There are currently arguments available, the most developed of which are due to Jonas Olson (e.g., 2011, 2014) which claim that epistemic error theory is in fact true. The arguments are sophisticated, though for our purposes, the details won't matter[12]. Let's now consider the following case.

> *Lemonade:* A student, *A*, sees what appears to be a glass of lemonade sitting on his meta-ethics professor's desk, and, accordingly, forms the belief that there is. Call this belief LEMONADE. After some further discussion with the professor and two classmates, *A* then acquires three further beliefs, in addition to LEMONADE, which are: (i) MOUNTAIN DEW, (ii) HALLUCINATION and (iii) ERROR THEORY. MOUNTAIN DEW is the belief that the meta-ethics professor drinks only Mountain Dew (which is different from, but looks like, lemonade). HALLUCINATION is the belief that *A* has recently ingested crystal DNT[13], which causes inaccurate but highly believable hallucinations. ERROR THEORY is the belief that Jonas Olson's arguments for epistemic error theory (as outlined during the class by the professor drinking the yellow looking drink) are persuasive, and thus, that there are no reasons for belief.

Stipulate that *A*'s original belief, LEMONADE, is epistemically unimpeachable at the time of its formation. *A*, let's assume, is justified in believing LEMONADE[14], prior to

---

Footnote 9 continued

had as their focus epistemic reasons, and so for the present purposes I'm articulating the view in a way that reflects this focus.

[10] However, as Streumer (2012) argues, it is this peculiar feature of epistemic error theory that as Streumer sees it insulates the view from familiar kinds of objections.

[11] See, for example, Wright (1994, p. 9). Cf., Miller (2014, Sect. 3) For a related argument to the effect that epistemic error theory is self-defeating, see Cuneo (2007, p. 117).

[12] In short, Olson (2011) rejects epistemic reasons understood as *categorical* reasons for belief, reasons which transcend individuals' desires, roles, etc. For some representative discussions of the categorical character of epistemic reasons, as such, see for example Feldman (2000, p. 682) and Railton (1997, p. 53). Olson rejects categorical and thus epistemic reasons for belief because, if there were any, they would be reasons that are implied by what Olson calls *transcendent* norms (as opposed to 'immanent norms', the latter of which imply merely *non*-categorical, or hypothetical, reasons). But Olson thinks that the arguments for rejecting transcendent norms in ethics also support the rejection of transcendent norms in epistemology. Olson describes the difference between transcendent and immanent norms as follows:

> "[Transcendent norms'] reason-giving force transcends agents' desires, ends, activities, or institutions. Immanent norms, by contrast, are those whose reason-giving force depends on agents' desires or ends, their engagement in certain rule-governed activities, or their occupation of certain roles; the reason-giving force of immanent norms does not transcend desires or ends, or rule-governed activities or roles, which is why immanent norms imply only non-categorical reasons."

[13] N-Dimethyltryptamine.

[14] We could just as easily run a version of the case in which *A knows* LEMONADE, by supposing that it's true that the professor's class contains lemonade, and thus that the first two defeaters (i.e., mountain dew

acquiring the further beliefs A acquires. By acquiring the belief MOUNTAIN DEW, A acquires (in the sense of Pollock 1986) a *rebutting* defeater for believing LEMON-ADE, which defeats A's justification for LEMONADE by indicating that LEMONADE is false. By acquiring the belief HALLUCINATION, A acquires an *undercutting* defeater for LEMONADE. HALLUCINATION indicates that the process (perception) that issued A's LEMONADE belief is unreliable in the present circumstances, given the ingestion of crystal DNT, and so therefore not sufficiently indicative of the truth of LEMONADE. But what about ERROR THEORY? What implication does A's acquiring ERROR THEORY have for the status of A's original justified belief that what is sitting on the professor's desk is a glass of lemonade?

One line of response to this question is that A's acquisition of the ERROR THEORY belief is entirely orthogonal to the epistemic status of A's LEMONADE belief. Call this *orthogonalism*. If orthogonalism is true, then this of course undermines the thought that A's acquiring ERROR THEORY might (in some way) defeat in any interesting way the epistemic status of A's belief LEMONADE.

The reasoning in support of orthogonalism goes as follows. (i) A's acquisition of ERROR THEORY can potentially defeat the epistemic status of A's belief LEMONADE *only if* A's acquisition of ERROR THEORY is relevant in some way to the truth of A's belief that LEMONADE. But, whether (ii) A's belief that ERROR THEORY is true is irrelevant to the truth of A's belief that LEMONADE; (iii) Therefore, it's not the case that A's acquisition of ERROR THEORY can potentially defeat the epistemic status of A's belief that LEMONADE.

Let's simply grant the orthogonalist's premise (i)—i.e., that A's acquisition of ERROR THEORY can potentially defeat the epistemic status of A's belief LEMON-ADE only if A's acquisition of ERROR THEORY is relevant in some way to the truth of A's belief that LEMONADE. Even if this much is conceded, we needn't accept the conclusion of the orthognalist's reasoning. This is because we should reject premise (ii). In short, this is because premise (ii) overgeneralizes in such a way that it has implausible implications concerning *undercutting* defeat.

Consider that undercutting defeaters are epistemically significant because they indicate that the grounds one has for one's belief aren't sufficiently indicative of the truth of the target proposition. After all, HALLUCINATION is epistemically significant for A *vis-à-vis* LEMONADE because it indicates that A's grounds for believing lemonade aren't sufficiently indicative of the truth of LEMONADE. But, A's acquisition of the belief ERROR THEORY *also* (like HALLUCINATION) indicates that A's grounds aren't sufficiently indicative of the truth of lemonade. Specifically, *A's* acquisition of ERROR THEORY indicates that A's grounds aren't sufficiently indicative of the truth of LEMON-ADE by indicating that A simply lacks the grounds which A originally thought A had which indicated the truth of LEMONADE.

What this means is that if orthogonalism is correct, then, it becomes at best mysterious why undercutting defeaters are epistemically significant. So if A's coming to believe that Jonas Olson's arguments for epistemic error theory are plausible

---

and hallucination) are misleading. However, it's simpler to make the point in terms of justification, which doesn't require that we make either of those defeaters misleading.

can't defeat in any way the epistemic status of *A*'s belief that the glass the professor is drinking from is filled with lemonade, it's *not* because orthogonalism is true.

As I indicated in the previous section, I think there is an interesting kind of epistemic defeat at the *meta*-epistemic level. And that is indeed what I think is going on in the case described above. However, before engaging with this, I want to consider one further diagnosis of the *Lemonade* case according to which *A*'s acquisition of the ERROR THEORY belief *does* count as a defeater for *A*'s justification for LEMONADE, but *not* by defeating the meta-epistemic status of *A*'s justification for LEMONADE. Rather, on the diagnosis I will now consider, the epistemic status that is defeated is just regular first-order (not-second-order) status: by acquiring ERROR THEORY, *A* simply loses (on this envisioned diagnosis) *A*'s justification for believing that LEMONADE.

In order to get this assessment of the case in view, we need to introduce a different kind of epistemic defeater, called a *no-reason defeater.* According to Michael Bergmann (1997), a *no-reason defeater* for a belief, *p*, is a reason to believe that it's no longer reasonable to believe *p* given that (a) one has no reason for believing *p* and (b) the belief that *p* is the sort of belief that it's reasonable to hold only if one has evidence for *p* (Bergmann 1997, pp. 102–103)[15]. For example, suppose I believe that the President is in Boston. But then, on reflection, I can't locate any reason I have for believing this, while maintaining also that this is the kind of proposition that's reasonable to hold only if I have evidence for it. I thereby acquire on Bergmann's view a no-reason defeater for my (would-be) justification for my belief that the President is in Boston.

Consider now the following question: Is *A*'s acquisition of the belief ERROR-THEORY plausibly a no-reason defeater for *A*'s justification for believing LEMONADE? There's an initially plausible story for why it is. And if this story is right, then the notion of meta-epistemic defeat needn't enter the picture here.

The 'no-reason' diagnosis of the *Lemonade* case goes as follows: in *Lemonade, A,* in acquiring the belief ERROR THEORY, is in much the same position as I am when on reflection I can't locate any reason for believing the President is in Boston. In both cases, it looks as though one comes to believe[16] that one lacks a reason for believing the target proposition, even though the target proposition would be reasonable to believe only if there is some evidence for it. In short, on this line of thinking, ERROR THEORY is a 'no reason' defeater because believing it is tantamount to rejecting that you have a reason for believing the professor's glass is full of lemonade.

The 'no-reason' diagnosis is, granted, an elegant way to think about the case. However, it is on closer consideration unworkable. The reason has to do with *defeater defeaters*, the acquisition of which neutralizes the original defeater and in the course of

---

[15] For further discussion, see Sudduth (2008, Sect. 6a). Plantinga (1993) also discusses a variety of defeat with the profile of what Bergmann calls a no-reason defeater. Thanks to Andrew Moon for drawing my attention to this point.

[16] For instance, me, by failing to locate any reason for my belief that the President is on Boston, *A* by coming to believe that *A* lacks any reason to believe there is lemonade in the professor's glass because there are not reasons for belief.

doing so causes the target belief to regain its previous epistemic status prior to having that status defeated.[17]

A defeater-defeater can be illustrated by a simple example. Let's simply run a twist on our case used to illustrate a no-reason defeater. In short, I believe that the President is in Boston, but on reflection, I can't locate any reason I have for believing this. Let's now add to the story. Suppose that what accounts for why I actually hold the belief that the President is in Boston is that I overheard some people mentioning this. Because the circumstances under which I encoded this information were unremarkable (e.g., overheard on a train) I do not remember the reason for my belief when I attempt to locate it.

On the no-reason line, my inability to locate this reason defeats my justification for believing that the President is in Boston. As the recent epistemic situationist literature has indicated[18], our cognitive abilities are to a significant extent influenced by environmental factors, including weather-induced negative mood[19]. Suppose I fly to Svalbard, and my mood deteriorates, improving my memory recall just enough so that I can now locate my reason for believing that the President is in Boston—viz., that I overheard it that morning on a train. Having now accessed this reason which was previously inaccessible to me, I acquire a *defeater-defeater*; it defeats my no-reason defeater. I no longer am such that I cannot locate a reason for my belief. The (previous) defeater is accordingly neutralized.

Let's return now to *Lemonade*. If ERROR THEORY were a no-reason defeater for LEMONADE, then, as such, it could *itself* be neutralized by the kinds of defeater-defeaters that characteristically defeat no-reason defeaters. But, as I want to suggest, ERROR THEORY can't *in principle* be defeated this way, i.e., *via* one's coming to believe of some reason that that is one's for believing the target proposition. Therefore, the conclusion we should draw is that *A*'s acquisition of the ERROR THEORY belief cannot be a no-reason defeater.

To appreciate this point, the first step will be to run a 'no-reason defeater' variation on the original *Lemonade* case.

> *Lemonade*:* A student, *A*, sees what appears to be a glass of lemonade sitting on his meta-ethics professor's desk. Furthermore, *A* overhears other students talking about the lemonade on the professor's desk, and even more, the professor announces that she is drinking lemonade. *A* accordingly forms the belief that there is lemonade on his professor's desk. Call this belief LEMONADE. *A* then (due to a moment of confusion and blurred vision, due to anxiety) tries but fails to locate a reason for believing LEMONADE but cannot find one. The spell of anxiety subsides. *A* then reflects and is able to locate the various reasons he has for believing that there is a glass of lemonade on his professor's desk.

In *Lemonade**, *A*'s no-reason defeater for his justification for believing LEMONADE is defeated when, after the spell of anxiety subsides, *A* can locate several reasons *A* has for

---

[17] For discussion of defeater-defeaters, see for example Lackey (2003).

[18] For some representative discussions of epistemic situationism, see Alfano (2012, 2014; forthcoming).

[19] As experimental studies reported by Forgas et al. (2009) show, memory recall is better in conditions in which bad weather generates a negative mood, than in good weather.

believing LEMONADE—viz., that *A* saw what appeared to be a glass of lemonade sitting on his professor's desk, that *A* overheard other students talking about this, and that the professor herself announced that she is drinking lemonade. Any one of these reasons *A* locates suffices to defeat the no-reason defeater *A* acquired for LEMONADE when, during the anxiety spell, *A* tried but failed to locate a reason for believing LEMONADE. *A* thus *regains* his justification for believing LEMONADE.

But—and this is why *A*'s acquisition of the ERROR THEORY belief in *Lemonade* cannot be a no-reason defeater—*these kinds of defeater-defeaters cannot defeat* A's ERROR THEORY belief. In order to bring this point into sharp relief, consider a final twist on *Lemonade*:

> *Lemonade**:* A student, *A*, sees what appears to be a glass of lemonade sitting on his meta-ethics professor's desk. Furthermore, *A* overhears other students talking about the lemonade on the professor's desk, and even more, the professor announces that she is drinking lemonade. Over the course of the meta-ethics lecture, *A* comes to believe that that Jonas Olson's arguments for epistemic error theory (as outlined during the class by the professor drinking the yellow looking drink) are persuasive, and thus, that there are no reasons for belief. *A* can on reflection locate what previously she regarded as her reasons for believing LEMONADE, viz., the perceptual evidence and the student and professor testimony.

Notice that, in *Lemonade**, when *A*'s anxiety subsides and *A* subsequently recalls the perceptual and testimonial evidence *A* had acquired for believing LEMONADE, he was by reflecting on this able to defeat the no-reason defeater acquired during the anxiety attack. But reflecting on these things does *not* have the same kind of epistemic significance in *Lemonade***. Having acquired (in *Lemonade***) the belief that Jonas Olson's arguments for epistemic error theory are persuasive, reflecting on such things as that *A*'s professor told him she was drinking lemonade does not by *A*'s own lights furnish *A* with a reason for believing there is lemonade in the glass.

Here, a quick parallel to scepticism is useful. Ordinarily, a hand-like appearance is evidence that there is a hand. However, as Crispin Wright (2007, pp. 26–27) puts it, if you replace the presumption that there is an external world with reason to believe you are instead a handless brain in a vat, whose experiences are being controlled by a highly sophisticated computer program, then the experience of seeing what appears to be a hand supports *not* the claim that there is a hand, but instead the claim that:

> The computer is right now implementing a phase of its program which requires me to suffer the illusion of having a hand and holding it up in front of my face (2007, p. 27).

Likewise, if you replace the presumption that there are reasons for belief with the belief that (as the epistemic error-theorist tells us) there are *not*, then appreciating that you believe that the glass looks like it's full of lemonade, and that a classmate and the professor herself told you it's filled with lemonade, does not support your believing you have a *reason* for believing LEMONADE. Rather, it indicates to you that you have the sort of things that would ordinarily (by others) be called reasons to

believe LEMONADE, but which you (taken in by the professor's exposition of Olson's arguments) now believe are *not* reasons to believe LEMONADE.

In short, no-reason defeaters are defeated by the production of reasons. A's ERROR-THEORY belief however is not defeated by the production of the very kinds of reasons that defeat paradigmatic no-reason defeaters. And so *A*'s error-theory belief is not a no-reason defeater.

The line I want to advance, and which I will sharpen in §4, is that *A*'s acquiring the ERROR THEORY belief in *Lemonade* is best understood as a *meta-epistemic* defeater. As such, it defeats meta-epistemic status, not first-order epistemic status, and it does so on a *reflective* level, and in a way that is subject to an altogether different form of defeat itself. I'll unpack these ideas further. But first, it will be helpful to contrast the kind of error-theoretic defeater outlined in this section with a very different kind of meta-epistemic defeater, which is generated *via* the acquisition of beliefs about relativism and knowledge.

3 Relativist meta-epistemic defeaters

In a series of recent works, MacFarlane (2005, 2010, 2014) has argued that knowledge-ascribing sentences do not get their truth values absolutely, but only relatively[20]. For example, suppose I say that "Rory knows that Paris is the capital of France." Whether this is true, on MacFarlane's relativist semantics for knowledge ascriptions, depends on the epistemic standards that feature in the *assessor's context*, which is the context in which my knowledge ascription is being assessed as true or false. My knowledge-ascription is thus what MacFarlane calls *assessment-sensitive.*

Of course, when I say "Rory knows that Paris is the capital of France," this can be assessed for truth or falsity from *indefinitely* many perspectives. And so what this means is that the sentence I utter when I say "Rory knows that Paris is the capital of France" does not get a truth-value absolutely, only relatively.

MacFarlane's overarching argument for a relativist treatment of knowledge ascriptions is that, in short, it better explains our patterns of using "knows" than do other competitor views—e.g., traditional (insensitive) invariantism[21], attributor contextualism[22], and subject-sensitive invariantism[23].

---

[20] MacFarlane, along with Kölbel (2002, 2003) and others have defended a truth-relativist semantics in various domains of discourse, including (along with knowledge-attributions) predicates of personal taste, indicative conditionals, gradable adjectives, future contingents and epistemic modals. For an overview of some of these projects, see Baghramian and Carter (2015, Sect. 5). For two other defenses of a relativist semantics for knowledge-attributions, see Richard (2004) and Kompa (2002). Kompa's view, however, might be best classified not as relativism but as non-indexical contextualism. For discussion of non-indexical contextualism and its relationship to relativism, see MacFarlane (2009).

[21] Traditional invariantism is the received or standard semantics for knowledge-ascriptions, according to which "knows" is not context-sensitive. For a sample explicit defense of this view, see Williamson (2005).

[22] See for example Cohen (1986), DeRose (1992) and Lewis (1996).

[23] Notable proponents include Hawthorne (2004) and Stanley (2005). For discussion of MacFarlane's arguments against traditional invariantism, contextualism, subject-sensitive invariantism, see Carter (2016, Ch. 7).

Again, the details of MacFarlane's argument won't concern us here. What matters (for our purposes) is that this position is on the table, and that there are established arguments for the view.

Consider now, the following case:

> *Oven*: *A* and *B* have just left for a holiday in Timbuktu. *B* asks *A* if *A* left the oven on. *A* says no, that he distinctly remembers turning it off. "I know the oven is turned off" he says, emphatically. *B* persists to doubt, as *B* did not see *A* turn the oven off. *A* thinks *B*'s doubts are foolish. Because it's a long way to Timbuktu, and the conversation about the oven is stressful, *A* tries to lighten the mood by putting on an audiobook of MacFarlane's (2014) monograph, defending a truth-relativist semantics for a range of different expressions, including for knowledge-attributions. *A* is impressed with MacFarlane's careful arguments, and comes to believe that MacFarlane is right that "knows" is an assessment-sensitive expression, and so that knowledge-attributions do not get absolute truth-values.

In *Oven*, would we expect *A* to continue to purport to know the oven was turned off after acquiring the belief—call it RELATIVISM—that a truth-relativist semantics for knowledge attributions is plausible? Most likely. After all, MacFarlane's semantics predicts that *A* will claim knowledge here. After all, from *A*'s perspective, where the context of use and the context of assessment are the same[24], *A* plausibly satisfies the relevant epistemic standard for knowing. Positioned within the framework of MacFarlane's (2014) most recent presentation of the view, what this means is that *A* can rule out the alternatives relevant at the context from which he is assessing the knowledge claim for truth or falsity. And after listening to the MacFarlane audiobook, *A* remains able to rule out the same alternatives as before. However, there is an important respect in which, at least from *A*'s own perspective, his knowledge claim cannot be put forward with the same kind of *authority* as before[25].

In *Oven*, *A*'s acquisition of the RELATIVISM belief is best understood as a meta-epistemic defeater *vis-à-vis A*'s claim to know that the oven is turned off. It defeats—in a way I will articulate in more detail in the next section—the meta-epistemic status of *A*'s knowledge claim by undermining the reasonableness of *A* to continuing to advancing the claim as objective. It does so much as, in the previous section, *A*'s coming to believe ERROR THEORY, in the *Lemonade* case, undermines the reasonableness of *A*'s regarding the evidence she locates as reasons for believing LEMONADE.

---

[24] See MacFarlane (2014, p. 60).

[25] Recall again Boghossian's (e.g., 2006) point that, ordinarily at least, we take for granted that the status of epistemic assessments is objective, and to the extent that this is right (a point to which we'll return), it means that when *B* denies that *A* knows the oven is turned off, this disagreement isn't faultless. From *A*'s (original) perspective in claiming *A* knows the oven is turned off, *A* regards himself in being committed to regarding *B* as *incorrect* in denying this. However, upon acquiring the belief relativism, *A* no longer regards himself as having this commitment, because (again, influenced by MacFarlane's arguments) *A* regards *A*'s statement that *A* knows the oven is turned off to get a truth-value of 'false' when assessed by *B*. From *A*'s perspective, then, the acquisition of the relativism belief undermines the reasonableness of *A* continuing to put forward *A*'s knowledge claim *as* an objective claim.

First, though, I want to consider two competing diagnoses of the situation. The first draws from Martin Kusch's discussions of metaepistemology, and the second suggests that, if relativism is a defeater at all for *A vis-à-vis A*'s claimed knowledge that the oven is off, it is merely a first-order rebutting defeater.

Let's consider first the argument that draws from Kusch (2010). Take as a starting point that that the acquisition of RELATIVISM for *A* doesn't *undermine* the reasonableness of *A* advancing his knowledge claim as objective unless *A* already regarded, or took for granted, *A*'s knowledge claim as objective in the first place[26]. Now Kusch denies the point Boghossian makes, and which I've argued for elsewhere[27], that ordinarily we take for granted that epistemic evaluations are being advanced as objective. According to Kusch, in ordinary epistemic discourse—i.e., outside of the philosophy classroom where meta-epistemic debates are taking place—individuals' knowledge and justification assessments are made without any commitments one way or another to relativism or absolutism, or any other meta-epistemic position[28]. To the extent that this line is right, we would have cause to wonder how *A*'s acquiring the belief REL-ATIVISM could *undermine* the reasonableness of *A*'s putting forward her knowledge claim as objective when we have no reason (absent positive reasons to think *A* has prior commitments in metaepistemology) to think *A* would be putting forward her knowledge claim as objective in the first place.

It's debatable whether ordinary folk incur meta-epistemic commitments, not by articulating them explicitly (which ordinary folk do not do), but by revealing such commitments in other ways[29]. But, even if Kusch's point is granted, it can be pointed out that at least some individuals, even if not ordinary folk, do intend to put forward knowledge claims as objective. The thought that acquiring the RELATIVISM belief couldn't undermine the reasonableness of putting forward a knowledge claim as objective is, for such individuals at least, by the by.

The second *prima facie* objection to the suggestion that, in *Oven*, *A*'s acquiring the RELATIVISM belief is best understood as a meta-epistemic defeater is that, by acquiring the RELATIVISM belief, *A* merely incurs a traditional rebutting defeater for *A*'s belief that he knows the oven is turned off. A rebutting defeater, recall, defeats first-order epistemic status by indicating that the target belief is false—i.e., by supporting its

---

[26] For example, if *A* was *already* a relativist, or more weakly—did not accept or take for granted that knowledge claims have an objective status—then *A*'s acquisition of the RELATIVISM belief doesn't *undermine* for *A* the reasonableness of *A*'s advancing his knowledge claim as objective in such a way that it could be regarded as *defeating* meta-epistemic status for *A*.

[27] I defend this point in Carter (2016, Ch. 1).

[28] As Kusch (2010) puts it "[…] according to my own experience of epistemic discussions with untrained students, when pressed on their stance *vis-à-vis* the relativism-absolutism opposition, they find it hard to come up with a straightforward answer. This does not of course suggest that philosophically untrained people are epistemic relativists; what it does indicate instead is that being introduced to, and becoming competent in, the practice of epistemic discourse does not involve deciding between epistemic absolutism and relativism. Most of our epistemic discourse functions in ways that do not bring this meta-epistemic alternative into view. And hence ordinary persons tend not to be committed either way."

[29] Elsewhere (2016, Ch. 6), I've argued against Kusch's line on this point by distinguishing between articulated meta-epistemological commitments which ordinary folk plausibly don't have, and revealed meta-epistemological comments, which I think ordinary folk (no less than epistemologists) do incur in virtue of pragmatic presuppositions.

denial. There's at least one indirect way of making the argument that when *A* comes to believe RELATIVISM, *A* acquires a rebutting defeater for *A*'s would-be knowledge that the oven is off in virtue of acquiring evidence that *A* does not *know* the oven is not off. Here is the reasoning: (i) Knowledge requires truth[30]. (ii) *A*'s evidence for RELATIVISM is evidence that *B*'s assessment that *A* does not know that the oven is off, is true (as this is what MacFarlane's semantics predicts). (iii) Given that knowledge implies truth, evidence for RELATIVISM is thus evidence that (assuming bivalence[31]) it's *not* the case that the oven is off. Thus, (again, given that knowledge implies truth) evidence for RELATIVISM is evidence that *A* does *not* know that the oven is off.

The problem with the above line of reasoning is that it would be viable only if we unpack the first premise—that knowledge implies truth—as the objectivist unpacks it, but not as the relativist unpacks it. The relativist unpacks the slogan that knowledge implies truth differently from the objectivist, so that from "*A* knows that *p*" we cannot infer that "*p* is true for *B*"[32]. An argument aimed at establishing whether *A*'s acquiring the belief RELATIVISM is a rebutting defeater for *A*'s knowledge that the oven is turned off should be in principle *neutral* with respect to whether relativism is true or false; it shouldn't *require* that the view be false (by depending on unpacking the slogan 'knowledge requires truth' in a way that is off limits to the relativist). Thus, the diagnostic story on which *Oven* is best interpreted as one where *A* acquires (by coming to believe RELATIVISM) merely a first-order rebutting defeater is not compelling. It is a story which, in short, begs the question against the relativist.

Yet, after *A* acquires the RELATIVISM belief, *A* really does incur *some* kind of epistemic loss *vis-à-vis A's* purported knowledge that the oven is turned off, provided (to control for Kusch's objection) that we suppose *A* is among those who had previously put forward the claim as objective. *A*'s acquired, on the view I'll now turn to clarifying, a meta-epistemic defeater—the acquisition of the relativism belief *is* significant *vis-à-vis* the epistemic status of *A*'s claim to know the oven is off. But the significance is best appreciated at the second-order, not the first.

4 Meta-epistemic defeat

I've indicated in Sects. 2 and 3 that in cases like *Lemonade* and *Oven*, we have circumstances where an individual acquires a new belief which seems to downgrade the

---

[30] This is near-unanimously accepted. For one recent voice of dissent, at least as regards the ordinary concept of knowledge, see Hazlett (2010).

[31] Note that the relativist can preserve bivalence relative to contexts of assessment, in the sense that a sentence is either true or false but not both relative to a particular context of assessment. For some relevant discussion on this point, see MacFarlane (2014, Ch. 2), concerning standard objections to relativism about truth.

[32] There are close parallels between this way of begging the question against the relativist, and the way in which, as Burnyeat (1976) notes, Plato is often accused by commentators of begging the question against Protagoras in the *Theatetus*, in the famous peritrope or 'self-refutation' argument against Protagoras's global relativism. Consider, in particular, Plato's remarks, at 171ab, that proceed as follows: "[…]it [The Truth that Protagoras wrote] has this exquisite feature: on the subject of his own view, agreeing that everyone judges what is so, he for his part [Protagoras] presumably concedes to be true the opinion of those who judge the opposite to him in that they think that he is mistaken… Accordingly, he would concede that his own view is false, if he agrees that the opinion of those who think him mistaken is true" (*Theatetus*, 171ab).

epistemic status of the target belief in each case, though not in the way that first-order defeaters downgrade epistemic status. I've indicated that the right way to think of these cases is in terms of *meta*-epistemic defeat. In this section, I want to consolidate some of the key features of meta-epistemic defeaters, sharpen their aspects of contrast with first-order defeaters, and in the process raise some new puzzles that I think future work on meta-epistemic defeat needs to address.

It will be helpful to begin by noting one very crucial difference between traditional first-order defeaters and the kind of meta-epistemic defeaters considered in Sects. 2 and 3, as this difference is relevant to how we should articulate a key point about the mechanisms of meta-epistemic defeat. To a first approximation, while both first-order and meta-epistemic defeaters defeat epistemic status (construed generally), meta-epistemic defeaters do so in a way that occurs entirely at the *reflective* level, but not the object level. For a given belief, p, facts about what other beliefs I have can cause first-order epistemic status simply not to obtain, *vis-à-vis*, p. Second-order status is different. The matter of whether this status *actually* obtains it is not sensitive to what I believe. And so therefore, lest we be radical constructivists about meta-epistemic status, for any belief q that I might form, if q in some way defeats the meta-epistemic status M of a target belief, p, it will not do so by actually causing p to lack M[33].

But, as I've indicated in Sects. 2 and 3, there remains an important sense in which the overall epistemic status of certain particular beliefs *is* sensitive to the acquisition of new meta-epistemic beliefs I acquire, such as ERROR THEORY and RELATIVISM. And this is so *even though* the acquisition of such beliefs does not defeat first-order status, *nor* does it cause second-order status to fail to obtain given that (again, lest we be radical constructivists) the obtaining of second-order status is insensitive to our acquisition of particular beliefs about second-order status.

The right way to make sense of the mechanisms by which meta-epistemic defeaters (like ERROR THEORY and RELATIVISM) defeat the meta-epistemic status of a given belief is as follows: for some individual S, and first-order target belief, F, S's acquisition of a meta-epistemic defeater M defeats the meta-epistemic status, α, of F not by causing α to fail to obtain, with respect to F, but by reflectively undermining for S, α, with respect to F. α is reflectively undermined for S, with respect to F, when α is caused to be rationally unavailable as a first-order mode of presentation, with respect to F, for S.

A mode of presentation (in the Fregean sense) is a way of thinking[34]. One notable way of thinking about something is *as* oneself, as in the case of *de se* attitudes. When I think of something under this mode, certain dispositions follow, e.g., I am inclined to

---

[33] Here is a simple example to make this point. Suppose, *ex hypothesi*, that knowledge claims really are objective. If that's the case, then when, in cases like *Oven*, someone comes to believe otherwise, the status of one's knowledge-claims doesn't thereby lose that status in any way. This is to say that, if knowledge is objective, then for any item of knowledge K I have, K has an objective status, and it has this status no matter what I believe about knowledge as a kind of attitude, and no matter whether I believe that all my knowledge is relative (in the sense claimed by MacFarlane). First-order status, (e.g., whether I know or am justified in believing that the oven is off) is sensitive to what I believe in a way that second-order or meta-epistemic status (e.g., whether my would-be knowledge that the oven is off is objective in character) is not.

[34] For discussion, see for example Evans (1982). Cf., Stanley (2011, Ch. 4).

use first-person pronouns[35]. Another much discussed mode of presentation is *practical*. Stanley and Williamson (2001), in defending the thesis that knowledge-how is a kind of propositional knowledge under a practical mode of presentation, put it this way:

> Thinking of a person as oneself entails being disposed to behave in certain ways, or form certain beliefs, given relevant input from that person. Similarly, thinking of a place as here entails being disposed to behave in certain ways, or form certain beliefs, given relevant input from that place. Analogously, thinking of a way under a practical mode of presentation undoubtedly entails the possession of certain complex dispositions. It is for this reason that there are intricate connections between knowing-how and dispositional states. (2001, p. 429)

Likewise, thinking of one's first-order epistemic status *as* objective plausibly entails certain complex dispositions. In the case of *Oven*, one such disposition (on the part of *A*) will be to regard *B* as making a mistake when *B* says that *A* does not know that the oven is off. Thinking about one's knowledge claim as objective will also entail a disposition to regard things like local social contingencies and personal knowledge-standards to be irrelevant to discovering the truth of one's knowledge claim[36].

I've said that meta-epistemic defeaters defeat (in their own idiosyncratic way) by reflectively undermining meta-epistemic status, and further that meta-epistemic status is reflectively undermined when is rationally unavailable as a first-order mode of presentation. This is not to say that it is not available *at all* as a first-order mode of presentation. For example, in the case of *Oven*, even after *A* comes to believe that MacFarlane's arguments for a relativist treatment of "knows" are persuasive, *A* is nonetheless *capable* of continuing to think of *A*'s knowledge claim—i.e., that *A* knows the oven is turned off—under an objective mode of presentation, and so can potentially continue to manifest the complex dispositions that are entailed by thinking of *A*'s knowledge claim under an objective mode. But doing so is no longer rational by *A*'s own lights, given *A*'s having the belief RELATIVISM[37]. This is why being reflectively undermined is articulated in terms of a mode of presentation being *rationally* unavailable (as opposed to *merely* unavailable). The epistemic loss one incurs *via* meta-epistemic defeat is best understood in terms of this kind of rational unavailability of the objective mode, and hence not (implausibly) in terms of the non-obtaining of the objectivity of the first-order knowledge-claim.

I now turn to a question that was raised in Sect. 2, in the discussion of error-theoretic defeaters, though never satisfactorily resolved. In Sect. 2, it was shown that error-

---

[35] See Lewis (1979) for the canonical presentation. See also Ninan (2010) for an overview.

[36] Furthermore, it's plausible to suppose that thinking of one's first-order epistemic status under an objective mode disposes one to take certain additional claims as pragmatic presuppositions (in the sense of Grice 1991 and Stalnaker 1974; 1978) which influence the kind of conversational moves one is inclined to make when discussing epistemic status. See Carter (2016, Ch 1) for discussion.

[37] Consider an analogy to *de se* attitudes. Suppose due to a brain injury, I come to believe that I am D.B. Cooper. I begin thinking of D.B. Cooper as myself, and come to think of D.B. Cooper by using first-person pronouns. Under such a circumstance, I might acquire good evidence—e.g., from doctors and historians—that I am not D.B. Cooper. However, if we suppose my conviction that I am D.B. Cooper is recalcitrant, I can continue thinking of myself as D.B. Cooper, as a matter of psychological possibility. I can just no longer continue to do so rationally, provided I believe the testimony from the doctors and historians.

theoretic defeaters are not defeated by the production of the very kinds of reasons that defeat paradigmatic no-reason defeaters. And so the conclusion drawn in that section was that error-theoretic defeaters are thereby *not* no-reason defeaters.

That point was meant to dispel the idea that cases like *Lemonade* can be accounted for in terms of traditional (first-order) mechanisms of defeat. But this invites a more general question. Can meta-epistemic defeaters be defeated, and if so, how? As it turns out, this question generates some interesting puzzles, the solutions to which might—potentially, at least—be special cases of more general solutions to certain problems in the epistemology of disagreement (a point, I'll return to).

Let's begin by running an idealization on the *Lemonade* case, focusing on structure. Suppose A believes that p, on the basis of reason R. A then comes to acquire a belief, q, where q is the belief that epistemic error theory is true. On the diagnosis I'm offering, articulated in the language introduced in this section, A's acquisition of q is a meta-epistemic defeater, *vis-à-vis* the meta-epistemic status of p, for A, by (in short) reflectively undermining its meta-epistemic status for A, by rendering rationally unavailable for A a realist mode of presentation of this epistemic status (i.e., that it is based on a reason) of A's belief that p[38]. Now, how might this meta-epistemic defeater *itself* be defeated?

The most natural candidate answer is to suppose that A could accomplish this through the acquisition of a further belief $q*$, where $q*$ is just the denial of ERROR-THEORY. After all, in *Lemonade*, if I acquire beliefs that MOUNTAIN DEW and HALLUCINATION are false, both of these new beliefs are paradigmatic defeater defeaters for MOUNTAIN DEW and HALLUCINATION. So wouldn't the acquisition of the belief that error theory is false suffice to defeat error theory in its capacity as a meta-epistemic defeater?

Perhaps not. In order to appreciate this point, first consider that various epistemically-irrelevant psychological factors could *cause* you to lose your ERROR-THEORY belief, where these causal facts are not rational descriptions of how the defeater is defeated[39]. But once the aim in play is clarified as that of giving a rational description of how a meta-epistemic defeater might itself be defeated, it becomes much less straightforward to see how the acquisition of a belief that is the negation of the meta-epistemic defeater could itself defeat the meta-epistemic defeater in question[40]—at least in the error-theoretic and relativist cases envisioned.

---

[38] Following Miller (2014) I am taking realism, for a given domain, to have commitments regarding existence and independence. Error-theory, on Miller's taxonomy, does not deny that *if* there were epistemic reasons, they would be mind-independent in the relevant sense. The error-theorist denies the existence leg of realism, by insisting that the property of 'being an epistemic reason' is never instantiated, and that there are no 'epistemic reason' facts. For discussion, see Carter (2016 Ch. 1).

[39] You might, for example, suffer brain trauma that causes you to simply forget you ever heard of Jonas Olson's theory, much less that you were persuaded by it. In such a circumstance, ERROR- THEORY, as a meta-epistemic defeater, is simply 'removed', in the way that ordinary first-order defeaters can be removed by similar non-rational mechanisms. But this isn't a rational description of the mechanisms of defeater defeaters.

[40] The same point holds, *mutatis mutandis* for the acquisition of a belief that indicates the process leading to the meta-epistemic defeater was reliably formed.

Consider, after all, once you believe ERROR- THEORY, it's rationally unavailable for you to think of your newly acquired reasons for believing the denial of error theory *as* reasons for believing the denial of error theory, just as it is rationally unavailable for you to regard your reasons for believing there is lemonade in the professor's glass *as* reasons for believing there is lemonade in the professor's glass. But then, it's not clear how, *even if* you come (on the basis of believing a new proposition, which is the denial of ERROR THEORY) to no longer believe the ERROR THEORY belief, that your doing so qualifies as rational by your own lights. Accordingly, it's not clear how the acquisition of a belief that is the negation of ERROR THEORY would suffice to defeat ERROR- THEORY in its capacity as a meta-epistemic defeater. I'm using the case of an illustrative example, though a parallel kind of problem arises in the case of *Oven*, with respect to the matter of how to rationally defeat the belief that knowledge-ascriptions are relative, by coming to form a new belief which by your own lights could aspire only to relative knowledge[41].

I have, to be clear, not suggested that meta-epistemic defeaters are indefeasible. Surely meta-epistemic defeaters *can* be defeated. More weakly, I've suggested that the mechanisms by which meta-epistemic defeaters stand subject to defeat themselves is not straightforward, and in an important respect, puzzling. Here, I think, is the source of conundrum: the acquisition of a new meta-epistemic belief (such as ERROR THEORY and RELATIVISM) can reflectively undermine (in the sense I've articulated) not *merely* the meta-epistemic status of beliefs we *already* have. But, moreover, the acquisition of such a belief can influence which modes of presentation are rationally available and which are rationally unavailable *vis-à-vis* new beliefs we might accept. And this is why defeating a meta-epistemic defeater isn't as simple as just coming to believe its negation, at least, in so far as we're interested in how meta-epistemic defeaters could be rationally (by the lights of the agent) defeated or neutralised.

To the extent that the general account of meta-epistemic defeat I've proposed is on the right track, the matter of articulating the mechanisms by which meta-epistemic defeaters are themselves defeated raises important questions that work on meta-epistemic defeat needs to address. I want to conclude by gesturing toward a potential

---

[41] With reference to *Oven,* when *A* comes to embrace MacFarlane's arguments for an assessment-sensitive semantics for knowledge attributions, *A*'s new belief—i.e., RELATIVISM—reflectively undermines for *A* the meta-epistemic status of *A*'s first-order knowledge claim, that *A* knows that the oven is off, by rendering rationally unavailable 'objectivity' as a first-order mode of presentation, with respect to *A*'s knowledge claim, for *A*. This is the application of my view at any rate. Now, how might *this* (relativist) meta-epistemic defeater be defeated? Again, let's consider the straightforward idea: that *A* comes to acquire a new belief, NOT- RELATIVISM which is just the denial of RELATIVISM. Does *A* thereby acquire a defeater defeater for RELATIVISM in RELATIVISM's capacity as a meta-epistemic defeater? Things here are (at best) not straightforward. Again, to reiterate, for a defeater be rationally defeated, we need a story for how the subject can rationally, in light of what she already believes, dismiss the original meta-epistemic defeater on the basis of her acquisition of some new belief—i.e., the defeater defeater. But, once *A* believes RELATIVISM, *A* can rationally interpret, for any belief, *x*, which *A* considers, *x* as a belief that aspires to mere relative (i.e., assessment-sensitive) knowledge, for *A*. This includes beliefs whose content are the very denial of RELATIVISM, which means that by *A*'s lights, the content RELATIVISM *is false* is something that aspires for *A* to the status of mere relative knowledge, such that, if *A* knew the content, it would not get an absolute truth value by *A*'s own lights. It is accordingly at least not straightforward how *A*'s acquisition of this new belief, then, is something which would entitle *A* to rationally dismiss the meta-epistemic defeater *A* had previously acquired.

way forward, by drawing some parallels with some of the literature on disagreement, and in particular, cases featuring deep disagreement.

Let's call a disagreement 'deep' if it has the following features: (i) there is a disagreement about the truth of some target proposition, p; and (ii) there is a disagreement about what kind of epistemic standards are relevant to adjudicating whether p[42]. One very famous such disagreement is the early 17th century dispute between Galileo and Cardinal Bellarmine. The "p" in this dispute can be represented as the claim that heliocentrism is true (a claim Galileo accepted and Bellarmine denied). What's epistemologically noteworthy about this case is that, given Galileo's and Bellarmine's very different background beliefs about what kinds of evidence are relevant to adjudicating matters of the cosmos (viz., scientific or scriptural), it looked to one another, and to some commentators, that neither was in a position to rationally persuade the other[43]. In short, neither was in a position to think of what the other regarded as evidence (*vis-à-vis p*) *as* evidence[44].

The matter of whether this kind of deep dispute could potentially be rationally resolvable has been a notable point of contention[45]. For our purposes, we needn't take a stand. What's relevant is that the philosophical problem of determining how meta-epistemic defeaters might themselves be defeated has features which closely mimic some of the well-established difficulties that feature in cases of deep disagreements.

In particular, consider that—in Bellarmine vs. Galileo—neither was in a position to think of what the other regarded as evidence (*vis-à-vis p*) *as* evidence. This can be appreciated as a social-epistemic variant of our own problem, of accounting how, for a given individual, a meta-epistemic defeater might itself be defeated through the acquisition of new beliefs. In both cases, the acquisition of such beliefs (which seem like they would be best candidates to, respectively, induce rational persuasion, and to rationally defeat a meta-epistemic defeater) are reflectively unavailable under the kind

---

[42] Hales (2014) regards disagreements which have this feature irreconcilable, and takes the existence of such disagreements to motivate relativism. For critical discussion of this argument, see Carter (2016, Ch. 4).

[43] For example, given Galileo's background beliefs, the evidence of planetary and lunar movements issued by his telescope supported the heliocentric model and counted against the geocentric model. But, given Bellarmine's background beliefs, including his beliefs about what counts as evidence for propositions about the cosmos, the information issued by the telescope did not point to heliocentrism, given that scripture was taken to be the basic source of celestial evidence and the information provided by the telescope seemed at odds with it. Galileo and Bellarmine were deadlocked.

[44] For instance, one thing which stands in the way of (for instance) Bellarmine being disposed to be rationally persuaded by the telescopic evidence presented by Galileo is that Bellarmine is already rationally committed (in light of what he already believes *about what counts as evidence*) to interpret new information he uncovers in a particular way. The appreciation of the telescopic evidence *as* evidence for heliocentricm was, to use the language introduced previously, rationally unavailable to him, as a mode of presentation. He could not rationally, by his own lights, think of the evidence *as* evidence. It is no surprise, then, that no particular claim which Galileo could put forward would be a candidate for swaying him.

[45] Various epistemological interpretations of this dispute have been noted by Rorty (1980), Boghossian (2006), Siegel (2011), MacFarlane (2009), Hales (2014) and Carter (2016), among others. In particular, what has been a point of dispute is two related issues. Firstly, are disagreements with this kind of structure rationally resolvable? Secondly, if no, does this kind of 'epistemic incommensurability' motivate epistemic relativism? For a recent negative answer, see Carter 2016, Ch. 4.

of mode of presentation which would make these objectives possible, in light of the agents' background beliefs.

In sum, a satisfying account of how rational persuasion might be possible in cases of deep disagreements would very plausibly shine light on how we could model the defeasibility of meta-epistemic defeaters. I indicated that a solution to the latter might in fact be a special case of the solution to the former, which is why I think future work on meta-epistemic defeat would benefit from interfacing with social-epistemic debates about disagreement.

5 Conclusion

In this paper, I've attempted to locate a kind of epistemic defeat—i.e., *meta-epistemic defeat*—which is epistemically interesting in its own right, and which can downgrade epistemic status through mechanisms that are importantly different from more traditional kinds of epistemic defeaters. Meta-epistemic defeaters *reflectively undermine* meta-epistemic status, by rendering such status rationally unavailable as a first-order mode of presentation. Further, it's been shown that meta-epistemic defeaters raise difficult puzzles when it comes to accounting for how meta-epistemic defeaters themselves can be defeated. I concluded by gesturing toward a kind of epistemic parallel between this problem and a related social-epistemic problem in the literature on deep disagreement; future work on meta-epistemic defeat could accordingly benefit by exploring further points of connection between meta-epistemic defeat and deep disagreements in social epistemology.

References

Alfano, M. (2012). Expanding the situationist challenge to responsibilist virtue epistemology. *Philosophical Quarterly*, *62*(247), 223–249.

Alfano, M. (2014). Expanding the situationist challenge to reliabilism about inference. In A. Fairweather (Ed.), *Virtue epistemology naturalized* (pp. 103–122). New York: Springer.

Baghramian, M., & Adam C, J. (2015). Relativism. In EN Zalta (ed), Stanford encyclopaedia of philosophy (pp. 1–46). Fall.

Baker-Hytch, M., & Benton, M. A. (2015). Defeatism defeated. *Philosophical Perspectives*, *29*(1), 40–66. doi:10.1111/phpe.12056.

Bergmann, M. (1997). Internalism, externalism and the no-defeater condition. *Synthese*, *110*(3), 399–417.

Boghossian, P. (2001). How are objective epistemic reasons possible? *Philosophical Studies*, *106*(1), 1–40.

Boghossian, P. (2006). *Fear of knowledge: Against relativism and constructivism*. Oxford: Oxford University Press.

Burnyeat, M. (1976). Protagoras and self-refutation in plato's theaetetus. *The Philosophical Review, 85*, 172–95.

Carter, J. A. (2015). *Group knowledge and epistemic defeat. Ergo*, *2*(28), 711–735.

Carter, J. A. (2016). *Metaepistemology and relativism*. London: Palgrave Macmillan.

Carter, J. A. (2017). Epistemological implications of relativism. In J. J. Ichikawa (Ed.), *Routledge handbook of epistemic contextualism*. London: Routledge.

Cohen, S. (1986). Knowledge and context. *Journal of Philosophy*, *83*(10), 574–583.

Cuneo, T. (2007). *The normative web: An argument for moral realism*. Oxford: Clarendon Press.

DeRose, K. (1992). Contextualism and knowledge attributions. *Philosophy and Phenomenological Research*, *52*(4), 913–929.

Evans, G. (1982). *Varieties of reference*. Oxford: Oxford University Press.

Feldman, R. (2000). The ethics of belief. *Philosophy and Phenomenological Research*, *60*(3), 667–695.

Forgas, J. P., Goldenberg, L., & Unkelbach, C. (2009). Can bad weather improve your memory? An unobtrusive field study of natural mood effects on real-life memory. *Journal of Experimental Social Psychology*, *45*(1), 254–257.

Grice, P. (1991). *Studies in the ways of words*. Cambridge, MA: Harvard University Press.

Hales, S. D. (2014). Motivations for relativism as a solution to disagreements. *Philosophy*, *89*(1), 63–82.

Hawthorne, J. (2004). *Knowledge and lotteries*. Oxford: Oxford University Press.

Hazlett, A. (2010). The myth of factive verbs. *Philosophy and Phenomenological Research*, *80*(3), 497–522.

Kölbel, M. (2002). *Truth without objectivity*. London: Routledge.

Kölbel, M. (2003). Faultless disagreement. *Proceedings from the Aristotelian Society*, *104*, 53–73.

Kompa, N. (2002). The context sensitivity of knowledge ascriptions. *Grazer Philosophische Studien*, *64*(1), 1–18.

Kusch, M. (2010). Epistemic replacement relativism defended. In M. Suarez (Ed), *EPSA epistemology and methodology of science: Launch of the European Philosophy of Science Association* pp. (165–75).

Lackey, J. (1999). Testimonial knowledge and transmission. *The Philosophical Quarterly*, *49*(197), 471–490.

Lackey, J. (2003). A minimal expression of non-reductionism in the epistemology of testimony. *Nous*, *37*(4), 706–723.

Lackey, J. (2005). Memory as a generative epistemic source. *Philosophy and Phenomenological Research*, *70*(3), 636–658.

Lewis, D. (1979). Attitudes de dicto and de se. *The Philosophical Review*, *88*(4), 513–543.

Lewis, D. (1996). Elusive knowledge. *Australasian Journal of Philosophy*, *74*(4), 549–567.

MacFarlane, J. (2005). The assessment sensitivity of knowledge attributions. *Oxford Studies in Epistemology*, *1*, 197–233.

MacFarlane, J. (2009). Nonindexical contextualism. *Synthese*, *166*(2), 231–250.

MacFarlane, J. (2010). Relativism and knowledge attributions. In S. Bernecker & D. Pritchard (Eds.), *Routledge companion to epistemology*. London: Routledge.

MacFarlane, J. (2014). *Assessment sensitivity: Relative truth and its applications*. Oxford: Oxford University Press.

Mackie, J. (1977). *Ethics: Inventing right and wrong*. London: Harmondsworth.

Miller, A. (2014). Realism. In E. N. Zalta (Ed), *The stanford encyclopedia of philosophy*. Winter.

Ninan, D. (2010). De se attitudes: Ascription and communication. *Philosophy Compass*, *5*(7), 551–567.

Olson, J. (2011). Error theory and reasons for belief. In A. Reisner, & A. Steglich-Petersen (eds.), *Reasons for belief* (pp. 75–93). Cambridge: Cambridge University Press.

Olson, J. (2014). *Moral error theory: History, critique, defence*. OUP Oxford: Oxford.

Plantinga, A. (1993). *Warrant and proper function*. New York: Oxford University Press.

Pollock, J. L. (1986). *Contemporary theories of knowledge*. New York: Rowman and Littlefield.

Railton, P. (1997). On the hypothetical and non-hypothetical in reasoning about belief and action. In G. Cullity, B. Gaut (eds), *Ethics and practical reason* (pp. 53–79). Cambridge: Cambridge University Press.

Richard, M. (2004). Contextualism and relativism. *Philosophical Studies*, *119*(1), 215–242.

Rorty, R. (1980). *Philosophy and the mirror of nature*. Princeton, NJ: Princeton University Press.

Seidel, M. (2014). *Epistemic relativism: A constructive critique*. London: Palgrave Macmillan.

Shah, N. (2011). *Can reasons for belief be debunked?*. In *Reasons for belief*. Cambridge: Cambridge University Press.

Siegel, H. (2013). Relativism refuted: A critique of contemporary epistemological relativism. Dordrecht: Springer Science & Business Media.

Siegel, H. (2011). Epistemological relativism: Arguments Pro and Con. In S. D. Hales (Ed.), *A companion to relativism*. Oxford: Wiley-Blackwell.

Stalnaker, R. (1974). Pragmatic presuppositions. In R. Stalnaker (Ed.), *Context and content*. Oxford: Oxford University Press.

Stalnaker, R. (1978). Assertion. *Syntax and Semantics*, *9*, 315–322.

Stanley, J. (2005). *Knowledge and practical interests*. Oxford: Oxford University Press.

Stanley, J. (2011). *Know How*. Oxford: Oxford University Press.

Streumer, B. (2012). Can We believe the error theory? *Journal of Philosophy, 110*, 1–25.

Sudduth, Ml. (2008). Defeaters in epistemology. *Internet Encyclopedia of Philosophy*. http://www.iep.utm.edu/ep-defea/.

Williamson, T. (2005). Contextualism, subject-sensitive invariantism and knowledge of knowledge. *The Philosophical Quarterly*, *55*(219), 213–235.

Wright, C. (2007). *The perils of dogmatism* (pp. 25–48). Themes from GE Moore: New essays in epistemology and ethics.

Wright, C. (1994). *Truth and objectivity*. Cambridge, MA: Harvard University Press.