



University
of Glasgow

Carter, J. A. and Mortini, D. (2020) Higher-order defeat in collective moral epistemology. In: Klenk, M. (ed.) *Higher Order Evidence and Moral Epistemology*. Series: Routledge studies in epistemology. Routledge: New York.

This is the Author Accepted Manuscript.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/192026/>

Deposited on: 8 August 2019

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

Higher-Order Defeat in Collective Moral Epistemology

J. Adam Carter & Dario Mortini
University of Glasgow

ABSTRACT. Recent work on the epistemology of moral deference suggests that moral knowledge must derive from a knower's own *ability* in a way that knowledge acquired easily through testimony need not. This paper transposes this idea to the collective level, and in doing so, shows how two leading accounts of collective knowledge, the *joint acceptance account* and the *distributed account*, would be best positioned to countenance group-level moral knowledge as knowledge creditable to group-level ability. The upshot is that we uncover some hitherto unnoticed puzzles to do with defeat in collective moral epistemology, puzzles which reveal collective moral knowledge to be surprisingly fragile *vis-à-vis* higher-order defeat compared to individual-level moral knowledge. A consequence of this disanalogy is that more work needs done if non-skeptical collective moral epistemology is to hold water.

I. INTRODUCTION

According to a popular view in individual moral epistemology called *pessimism*, there is something deeply problematic about believing a moral proposition purely on the basis of another's say so (e.g., Jones 1999; Nickel 2001; Driver 2006; Hopkins 2007)¹. For example, there is something amiss with believing that cruelty is wrong because your teacher told you so, but for no other reason.

Though pessimists disagree amongst themselves about why moral deference is problematic, a point of agreement is that believing a moral proposition purely on the basis of moral testimony violates an epistemic norm governing belief. Robert Hopkins (2007) articulates such a norm as follows:

Grasping norm (GN): You (epistemically) should believe a moral proposition only if you grasp the moral grounds for it.

Because a grasp of the moral grounds of a proposition can't be straightforwardly transmitted from speaker to hearer via testimony², some pessimists reason from GN that moral knowledge can't be acquired via testimony at all. For our purposes, we remain agnostic on this point.³ What we take to be more interesting, from an epistemological point of view, is what GN implies more generally about what moral knowledge demands.

On one way of thinking, knowledge, *per se*, doesn't require grasping grounds⁴, and so it follows from GN that there is no such thing as moral knowledge because nothing corresponding with (mere) knowledge is normatively constrained in this demanding way; yet, whenever GN is satisfied, one is in the market for moral *understanding*, for which such a grasp is essential.⁵ An alternative route open to pessimists is to maintain that knowledge, generally, doesn't require grasping grounds, however, it does require this in the special case of moral knowledge.⁶

Our main interest here is not to adjudicate this dispute. Rather, we will investigate how the foregoing predicament that gives rise to it also generates some interesting and hitherto unnoticed epistemological puzzles when transposed to the collective level, particularly with regard to the relationship between collective moral knowledge, disagreement, and defeat.

Here is the plan for the paper. §2 sharpens the foregoing problem in individual epistemology and argues that if there is moral knowledge, then it plausibly involves what robust virtue epistemologists (e.g., Greco 2010; 2012; Sosa 2009; 2015) call *cognitive achievement*, or cognitive success (e.g., true belief) creditable to the exercise of cognitive ability. An implication of this 'achievement' requirement on individual moral knowledge is that it is, at least in principle, more easily defeated than otherwise, given that more is required to retain moral knowledge than non-moral knowledge. However, this apparent fragility is not particularly worrying at the individual level. This is because, as we will show, individual-level abilities that give rise to achievements are generally stable.

Interestingly, as we will see, things are different at the collective level. In §3, we take, as a starting point that if there is collective moral knowledge, then it must plausibly be primarily creditable to (collective) ability.⁷ But what does this involve? In §4, we taxonomise two mainstream accounts of collective knowledge, the *joint acceptance model* and the *distributed model*, and then, in each case, we show how the idea that collective moral knowledge requires collective achievement would be plausibly modelled. §5 argues that, on both models, collective moral knowledge is

extremely fragile; it is only difficult to acquire⁸, but much easier than other kinds of group knowledge to defeat.

II. (INDIVIDUAL) MORAL KNOWLEDGE, CREDIT AND DEFEAT

The metaethical anti-realist might respond to pessimist's key idea by trying to undermine the initial puzzle motivating it. Perhaps there *is* no moral knowledge. This might be because surface-level moral claims are categorically false (i.e., error theory⁹) or non-truth evaluable (i.e. non-cognitivism). If this is right, then it's no wonder that something seems amiss about gaining moral knowledge through testimony. Epistemic anti-realist views that respond to the initial puzzle this way are committed to skepticism about moral knowledge (and, more generally, to skepticism about any kind of evaluative knowledge as it is traditionally (i.e., realistically) construed¹⁰). An assumption we will be making in what follows is *non-skeptical*: that moral knowledge, realistically construed, is possible, and that there is at least some moral knowledge.

With this assumption in play, let's focus in on what moral knowledge might involve over and above what non-moral knowledge involves. The most natural *general* answer is: a certain kind of ability not required by non-moral knowledge. Why? First, 'ability' is the sort of thing that can't be transmitted simply through testimony. (Consider: David Gilmour can tell you how to play a guitar solo; it doesn't follow that you are thereby able to play that guitar solo.¹¹) Second, grasping grounds implicates not only *something* that non-moral knowledge doesn't essentially involve, but an *ability* non-moral knowledge doesn't essentially involve.¹² Thirdly, if moral knowledge requires more by way of ability than non-moral knowledge, then this could help us to make sense of an important data point about the relationship between moral knowledge and moral action, to wit, that the moral goodness of an action plausibly depends not only on doing the right thing, but also on doing it for the right kind of reason. If moral knowledge, *as such*, requires abilities beyond what non-moral knowledge requires¹³, then we can easily make sense of the idea that moral knowledge is important to morally good action. If moral knowledge lacked any such abilities, then this connection would be harder to explain.

Granted, if *all* knowledge demands a lot of us by way of ability, then it would be hard to see how moral knowledge might be distinctive in what it demands of us. There is one view of knowledge

in particular—*robust virtue epistemology*—according to which *all* knowledge has to be primarily creditable to ability.¹⁴ As John Greco (2003) succinctly puts it:

To say that someone knows is to say that his believing the truth can be credited to him. It is to say that the person got things right due to his own abilities, efforts and actions, rather than due to dumb luck, or blind chance, or something else (2003, 116).

This view has a number of well-known advantages, not least that it offers an elegant way to deal with standard Gettier cases.¹⁵ However, one of the main disadvantages of this kind of ability-heavy view of knowledge is that it is too strong to reconcile with the prevalence of testimonial knowledge gained cheaply—e.g., by trusting a reliable source in the absence of defeaters. A point notably made by Jennifer Lackey (e.g., 2007) is that in paradigmatic cases of testimonial knowledge exchange (e.g., as when one asks for directions in a new town), it should be the *testifier* rather than the testimonial *recipient* who deserves the credit (if anyone does) for the recipient getting things right when they do.¹⁶

While testimony cases pose a serious obstacle for robust virtue epistemology as a general account of knowledge *simpliciter*, they offer an interesting vantage point to appreciate just how closely two things line up together: (i) moral knowledge, which we've shown plausibly requires ability in a way that non-moral knowledge does not; and (ii) knowledge *simpliciter* which, according to robust virtue epistemology, also requires the exercise of ability in a way that (as we've seen) is at tension with the thought that some knowledge is easily transmittable via testimony.

A plausible working hypothesis to draw here is that even if *all* knowledge does not require the kind of cognitive achievement (i.e., cognitive success due to ability) the robust virtue epistemologist identifies with knowledge and so takes to be necessary for acquiring it, *moral knowledge* in particular does require this. For ease of reference, call this idea the *Credit Condition on Moral Knowledge*.

Credit Condition on Moral Knowledge: *S* knows a moral proposition, p_m only if *S*'s believing p_m truly is primarily creditable to *S*'s exercise of (morally relevant) cognitive ability.

There are various ways that a credit condition on moral knowledge might be substantively glossed. For instance, those sympathetic to Hills' (2009) thinking might insist that the abilities involved in the credit condition include some of those that Hills takes to play a role in

understanding.¹⁷ Alternatively, one might view these abilities referenced by the credit condition more standardly along virtue reliabilist lines.¹⁸

Either way, a credit condition on moral knowledge, no matter how the details are filled in, is going to carry with it an important commitment to thinking about defeat and knowledge *asymmetrically* in cases of moral and non-moral knowledge. In particular, an implication is that moral knowledge is going to be—at least in principle—less resilient to being undermined by defeaters than non-moral knowledge, and this comparative fragility is on account of its being more demanding.

To make this idea more concrete, consider the following case:

COGNITIVE SABOTEUR: Through the exercise of his excellent moral reasoning, Theon comes to appreciate that selling high-risk subprime mortgages is morally wrong. Furthermore, through the exercise of his excellent math abilities, Theon comes to appreciate that Pythagoras's theorem is true—*viz.*, that the square of the hypotenuse of a triangle equals the sum of the squares of the other two sides. Unfortunately, Varys whispers to Theon 10 confusing moral claims and 10 confusing mathematical claims, with the sole purpose of sabotaging Theon's cognitive life. Varys's testimony has destabilized Theon's moral and mathematical abilities, which leads Theon to rightly begin to distrust them and even forget how to exercise them, even though—crucially—he retains his *beliefs* (which he now takes only on reliable testimony) that selling high-risk subprime mortgages is morally wrong and that Pythagoras's theorem is true.

A first point to note is that, even without the mathematical abilities he had before, he can *know* that Pythagoras's theorem (i.e., that *the square of the hypotenuse of a triangle equals the sum of the squares of the other two sides*) is true simply by continuing to trust mathematical experts.¹⁹ However, while Varys' testimony doesn't defeat Theon's mathematical knowledge, even if it has a deleterious effect on his mathematical abilities, it *does* seem to wreck not only Theon's moral abilities (*vis-à-vis* the subprime mortgage proposition) but *also*, via the credit condition on moral knowledge, to defeat his moral knowledge. After all, these moral reasoning abilities are undermined such that Theon is now believing what he does about the morality of subprime lending by simply trusting others.

Fortunately, this difference in fragility is not particularly concerning—despite initial appearances—at least at the *individual level* where we are considering it presently. This is because abilities demanded by any plausible unpacking of the credit condition on moral knowledge *must* be stable in such a way that they will in practice withstand all but the strongest kinds of SABOTEUR-style cases.

To see why this point holds, just consider how proponents of a credit condition on knowledge diagnose Plantinga's (1993) brain-lesion case, in which an undetected brain lesion happens to reliably cause the subject, 'Al', to believe that he has a brain lesion despite having no other evidence to support this. Is this an *ability* to which we can credit Al? If so, then (oddly) it looks as though Al's believing truly that he has a brain lesion is primarily creditable to ability rather than, say, luck. But this seems too permissive; Al's getting it right seems to have nothing to do with his abilities.²⁰

The way that robust virtue epistemologists such as Greco have dealt with such cases is to insist that the kind of abilities that can generate knowledge must be, as he puts it, *grounded in the subject's cognitive character*. According to Greco, this means they must:

be (a) stable [...] and (b) well integrated with other of the person's cognitive dispositions ... the cognitive process associated with the lesion is not well integrated with other aspects of the person's cognition. The process produces only a single belief, for example, and it is unrelated and insensitive to other dispositions governing the formation and evaluation of belief (2010, 152).

This point about stability has an important ramification for how we should think about cases like COGNITIVE SABOTEUR. What's important is that knowledge-generating abilities will be, in virtue of being well-integrated into the subject's cognitive character, *highly resilient* to being undermined in the way described in COGNITIVE SABOTEUR. An implication is that whilst, say, testimony might suffice to undermine a *poorly* integrated disposition that falls short of *bona fide* ability by the lights of a plausible credit condition, the wrecking of well-integrated cognitive abilities (at least, at the individual level) through the acquisition of new beliefs will *not* be easy to do at all.²¹ And this is welcome news: it means that—at the individual level, at least—the comparative asymmetry in resilience to defeat between moral and non-moral knowledge is unlikely to generate any serious skeptical threat to the (individual) moral knowledge we have. Things, however, are very different at the collective level. And it's to this point that we'll now turn.

III. COLLECTIVE MORAL KNOWLEDGE: PARITY PRINCIPLES

Just as individuals know things, so do groups. For example, the FBI knows where the President is at all times. Chevrolet knows that airbags must be put in cars before they can be sold. CERN knows that the 125 GeV/c² particle discovered in 2012 is a Higgs Boson.

On one way of thinking about what group knowledge involves, the above knowledge ascriptions come out true provided at least one (or perhaps several or most) individuals of the target group possesses the relevant item of knowledge. This view is called *summativism*: what's key to summativism is that group knowledge reduces to individual knowledge.

Non-summativism, by contrast, is a more philosophically interesting way of thinking about group knowledge. According to non-summativism, which is gaining traction in social epistemology²², groups can have epistemic properties even if no individual in the group possesses them, including the property of knowledge.²³ In the next section, we'll review some of the standard ways in the non-summativist literature to make good on this idea (and how these connect with different ways of thinking about non-summativist moral knowledge, in particular).

But first we want to make explicit four assumptions that we will be making, and which allow us to engage in some new ways with puzzles that arise in connection with (non-summative) moral knowledge, disagreement and defeat. Firstly, we are going to assume that there is non-summativist knowledge. That is to say, we assume that there are group subjects that know things, where this group-level knowledge is not reducible to a summation of the individual knowledge of its members.²⁴ Second, we are going to assume that there is not *merely* non-summativist *non-moral* knowledge, but also non-summativist *moral* knowledge—viz., that in some cases, a group can know a moral proposition.

The third and fourth assumptions will be the most important in what follows. The third assumption is what we'll call 'Parity Principle 1':

Parity Principle 1: If an individual S knows a moral proposition, p_m only if S 's believing p_m truly is primarily creditable to S 's exercise of (morally relevant) cognitive ability, then the same goes for (non-summative) groups agents.

Parity Principle 1 is a special case of a more general parity principle that is more or less universally accepted in collective epistemology. This more principle says (roughly) that *epistemic* conditions (e.g., justification) on individual knowledge carry over, *mutatis mutandis*, to the collective level.²⁵ What we're calling Parity Principle 2 below is just another instance of the general principle: if individual *moral* knowledge requires that a credit condition be satisfied, then, *mutatis mutandis*, so does group moral knowledge. Parity Principle 1, in conjunction with the *Credit Condition on Moral Knowledge*, jointly *imply* 'Parity Principle 2':

Parity Principle 2: A group g (non-summatively) knows a moral proposition, p_m only if g 's believing p_m truly is primarily creditable to g 's exercise of (morally relevant) cognitive ability.

Parity Principle 2 is, effectively, a group-level version of the individual credit condition on moral knowledge. The individual-level version of this principle—while it made moral knowledge easier to defeat than non-moral knowledge—did not do so substantially. This was due to the stability of individual level abilities of the sort that are capable of generating individual level knowledge. Whether the same holds for group-level knowledge remains to be seen.

IV. COLLECTIVE MORAL KNOWLEDGE: TWO VARIETIES

In this section, we first outline two strategies for fleshing out non-summativist knowledge, generally speaking: (i) the *joint acceptance model* and the (ii) *distributed model*. Then, for each account, we show in outline what it would take to countenance Parity Principle 2. Given the differences between the two accounts, the shape the credit condition will take in each case will be different.

4.1 Joint-Acceptance Model

According to the *joint-acceptance model* of group knowledge, knowledge is built out of group belief (e.g., Gilbert 1987; 2002; 2013). A group belief is, itself, a function of conditional commitments on the part of its individual members. The key features of the joint commitment account of group belief are as follows:

Joint Acceptance-Belief (JAB): (i) A group, g , believes p iff the members of g jointly accept p ; (ii) the members of g jointly accept that p when the members conditionally commit to

accept that p ; (iii) members of g conditionally commit to accept that p when each is committed to acting as if p provided the others do.

For example, according to JAB, a jury believes that the defendant is guilty provided the members of the jury commit to act, in their capacities as jury members, as if the defendant is guilty. This will include voting ‘guilty’, e.g., by raising their hand at the appropriate time, responding in a way that is consonant with a guilty vote when queried by the foreman or judge, etc. Note that on Gilbert’s JAB model, this conditional commitment does not extend to the individual members’ believing the defendant is guilty in a private capacity, just to acting as if the client is guilty in their capacity as jurors.²⁶

Nothing about JAB prevents a group from believing a proposition that is false; group *knowledge*, then, at least requires that the group belief be true, as well as justified. Notice, though, that the ‘justification’ of a JAB-style group belief won’t simply be a matter of the whether the individuals’ beliefs are justified.²⁷ After all, as a non-summativist model, JAB (as per the jury example illustrates) does not *require* individual belief for group belief.

So what, then, is the source of group-level justification, when a group-level belief is justified? A natural answer here is a simple reliabilist answer, one that does not require *additional* group beliefs to function as group reasons.²⁸ On the simple reliabilist model, a JAB-style group belief is justified just in case the process of joint acceptance is one that reliably gets to true beliefs.²⁹ On such a model, then, we might think of group knowledge provisionally as a JAB-style belief that is both true and which arises from truth-reliable joint acceptance.

There is, unfortunately, a lurking problem. Joint acceptance issues in group-level propositional outputs with both mind-to-world and world-to-mind directions of fit. In this way, it is importantly disanalogous from traditionally reliable belief-forming processes at the individual level (e.g., perception) which, when functioning normally, issue only mind-to-world outputs. As such, it’s not clear how ‘joint acceptance’ is plausibly a reliable process in a sense that would mimic the kind of reliable processes that we expect to issue in individual knowledge.

Jeroen de Ridder (2014) proposes a way to get around this problem. Taking scientific group knowledge as a paradigm for group knowledge, de Ridder maintains that a group is justified in believing something, p , only if the group belief is properly based on a reliable process of *inquiry*,

where a process of inquiry implicates a joint commitment to getting to the truth about whether or not p . This caveat avoids the worry facing a flat-footed reliability account because, while unqualified ‘joint acceptance’ as a process type isn’t a viable candidate for reliability, specifically *inquiry-directed* joint acceptance by comparison is. Even more, de Ridder’s basing condition closes an important potential gap between the doxastic output issued by joint acceptance of the group and the reliable inquiry-directed process.

Let’s suppose that something like de Ridder’s account is on the right track and that, e.g., with suitable supplementation with an anti-Gettier proviso, it offers a workable non-summative account of group knowledge built out of JAB-style group belief. Even if these assumptions are granted, it follows from Parity Principle 2 that *additional* conditions will be needed to be satisfied if the group is to have specifically *moral* knowledge.

After all, the credit condition that features in Parity Principle 2 isn’t going to be secured simply through a basing condition such as de Ridder’s, *viz.*, where the relevant group output must be based on a reliable inquiry-directed process. For one thing, not all processes are abilities (as the discussion of cognitive integration in §2 reveals at the individual level). And so, what follows from Parity Principle 2 is that a workable JAB account needs to be supplemented with a further account of group ability if it is to countenance group knowledge. For another, such an account of group ability needs to be put to work in the account in such a way that the account can explain how (when a group knows a moral proposition) the group ability is what primarily explains why the group’s doxastic output is true.

4.2 A Distributed Model

On a *distributed* model on non-summative group knowledge, a group can know something even though none of the individuals knows it, and even if the group doesn’t jointly commit to the proposition in the way required by JAB. What’s important, on the distributed model, is principally that the group generates propositional outputs in a reliable way.

For example, on Alexander Bird (2010)’s model, what’s key to group knowledge is that the group plays functional roles that are analogous to the knowledge-generating cognitive powers of individuals.³⁰ In particular, according to Bird, group knowers will have the following properties:

- (i) they have characteristic outputs that are propositional in nature (*propositionality*);

- (ii) they have characteristic mechanisms whose function is to ensure or promote the chances that the outputs in (i) are true (*truth-filtering*);
- (iii) the outputs in (i) are the inputs for (a) social actions or for (b) social cognitive structures (including the very same structure) (*function of outputs*) (2010, 43-4).

A group, for Bird, is functionally integrated when there is dependence between them for their own proper functioning, and such dependence is what determines group membership. This is of starkly different from the Gilbert-style requirement of joint commitment. One similarity, though, concerns a broadly reliabilist requirement that features in both models. On the JAB model (glossed with de Ridder's justification condition), the reliability requirement on group knowledge is a matter of a group belief's being properly based on a reliable process of inquiry. On a distributed model like Bird's, the reliability requirement is captured in clause (ii)—viz., Bird's truth-filtering requirement.

Let's now assume for the sake of argument that Bird's account—supplemented with a suitable anti-Gettier proviso—offers a viable account of non-summativist distributed group knowledge. Even if this is assumed, it follows from Parity Principle 2, that an additional credit condition will be needed to be satisfied if the group is to have specifically *moral* knowledge, and not merely non-moral knowledge.

What would it take to satisfy this further credit condition on a distributed model? Given that it's possible for all of Bird's conditions to be met while the credit condition is not met, further elaboration is needed. One suggestion here can be extracted from recent work on group knowledge by S. Orestis Palermos (2015). Palermos, in discussing the conditions under which group knowledge might be an achievement on virtue reliabilist lines, proposes a kind of *modal* condition according to which “getting to the truth of the matter as to whether p (or not- p) could only be collectively achieved and is *thereby* creditable only to the group as a whole.³¹” Transposed to the language of Bird's model: getting to the truth of the matter as to whether p (for some moral proposition p) is primarily creditable to g 's exercise of a (morally) relevant ability if and only if the group's truth-filtering mechanisms are not only sufficient but also *necessary* for getting to the truth of the matter as to whether p .

With this kind of supplement, then, we can in principle make sense of differential demands within a distributed model for (i) non-moral knowledge; and (ii) moral knowledge, respectively, where the latter, but not the former, is meant to countenance Parity Principle 2.

V. COLLECTIVE MORAL KNOWLEDGE AND DEFEAT: TWO NEGATIVE RESULTS

In this section, we want to show how *both* the accounts of group knowledge sketched in §4—*viz.*, the joint acceptance and distributive accounts—face a dilemma concerning higher-order defeat: in short, each turns out to be able to accommodate group-level moral knowledge only by making such knowledge problematically fragile.³²³³

5.1 Joint commitment model

Recall that, on a JAB-style account of group knowledge, it needs to be clear how—when a group knows a *moral* proposition—a *group ability* is what primarily explains why the group’s doxastic output is true. But how, exactly, is a group going to acquire any abilities it might have on a joint-acceptance model?

Miranda Fricker (2010) offers a suggestion. In her work on collective character traits, Fricker shows how we can make room for collective character virtues within a Gilbert-style joint-acceptance model, according to which joint acceptance is what fixes the non-summative properties of a group. On Fricker’s proposal, a group *g* has a collective character virtue when the members of *g* jointly commit to a good motive as a body. For instance, in the case of a committee, a committee that jointly commits to being open-minded or impartial when undertaking some task type, *T*, can be considered open-minded *vis-à-vis* that task type. It is contestable whether character virtues must themselves be reliable.³⁴ Abilities, at least those that are knowledge-generating, on the other hand, must be. Though a reliability requirement can naturally be incorporated into the kind of character account Fricker is proposing: the idea in short is that a group has an ability just in case it (i) jointly commits to achieving some good end, where ‘good’ is relative to a given domain (e.g., moral, epistemic, aesthetic, etc.); and (ii) is reliably enough successful at bringing this good end about.

On this kind of a template view, then, a JAB model could accommodate Parity Principle 2 by insisting that when a group knows a *moral* proposition, what primarily explains why the group’s doxastic output is true is a group ability that is, itself, fixed in part by joint commitments to achieving some epistemically good end.

Unfortunately, no matter how we fill in further details with this proposal, there is a looming problem. To bring this problem to light, consider the following case:

DISAGREEMENT:³⁵ A bioethics policy committee, C, is tasked with determining whether it is morally acceptable to use BrainEx technology to perform a perfusion on a dead human brain.³⁶ The committee (with reference to JAB) jointly commits to two things: firstly, they jointly commit to investigating the matter of whether performing such perfusions is morally acceptable in an intellectually rigorous and open-minded way that accords with scientific standards; then—following their detailed investigation—the group makes a second joint commitment: to the truth of the proposition that using BrainEx Technology to perform a perfusion on a dead human brain is *not* morally acceptable. During the course of the committee’s deliberation, one of the committee’s members (member A) registered reservations about how the committee was weighing evidence about what counts as ‘brain death’, and this led to a disagreement with another member, member B. But A agreed with B and the rest that the total evidence overwhelmingly supported what the group jointly committed to (e.g., that such perfusions are not morally acceptable); thus, A joined the rest in the eventual joint commitment that was made.

A first thing to note about the DISAGREEMENT case is that it (assuming for the sake of argument that the target proposition is true) looks like about as good a candidate for non-summative moral knowledge within a JAB framework as you could expect to encounter in practice. It is, after all, not reasonable to expect that there will be *no* disagreements whatsoever between a group about such things as the how certain kinds of evidence (in this case, concerning usage of the term ‘brain death’) should be interpreted scientifically.

But, even so—and here is the crux of the problem—the kind of dispute we find in DISAGREEMENT about the standards being used to evaluate the evidence turns out to be enough to defeat the group’s moral knowledge. The reasoning here is as follows: the defeasibility conditions for group ability on a JAB model are highly fragile. If, in DISAGREEMENT, some members submit during the course of the group investigation that the group is not following agreed-to scientific standards, then it trivially follows that (at least) some members are not *acting in their capacity as group members as if the group is following such standards*. But since joint commitments are on the JAB model *conditional* commitments this means that even minority dissent has the power to release others from their conditional commitments.

Of course, a tempting gloss of the situation just described is to emphasize that the minority dissent that features in in DISAGREEMENT is just a kind of ‘higher-order’ disagreement about the *methods* used by to group to reach its conclusion. There was no first-order disagreement within the group about the truth of target proposition itself. But, crucially, an attempt to minimize the epistemic significance of this (even lone) higher-order dispute is simply not on the table if Parity Principle 2 is to be upheld. Parity Principle 2, to reiterate, articulates a sense in which moral knowledge of a group must be creditable primarily to group *ability* (which we’ve fleshed out on the joint acceptance model along Fricker’s lines). Whenever that ability is undermined, then so, thereby, is group moral knowledge. And this is the case even if undermining a group ability is not sufficient for undermining *non-moral* group knowledge. So long as a lone higher-order disagreement within a group can suffice to undermine group ability, then it can suffice to undermine group moral knowledge.

What we find, then, is an important disanalogy between the comparative fragility of (i) individual moral knowledge; and (ii) group moral knowledge, at least when group moral knowledge is theorized about within a joint commitment framework. A proponent of a joint-commitment model is forced to make a choice that (as we’ve suggested in §2) a proponent of individual moral knowledge is not forced to make. The choice is: (i) Either give up the idea that group *moral* knowledge requires ability in a way that non-moral knowledge does not (i.e., give up Parity Principle 2) and sever the connection at the group level between moral knowledge and ability that we find well motivated at the individual level, *or* accept Parity Principle 2 and accept that group moral knowledge is highly fragile³⁷, so fragile that it will be likely to be undermined in cases that feature even outlying higher-order disagreement of the sort we find in DISAGREEMENT. Either route is problematic.

5.2 *The distributive model*

A variation on the dilemma just sketched faces a proponent of *distributed* group moral knowledge. As we showed in §4.2, it’s possible for all of Bird’s conditions to be met while the credit condition implied by Parity Principle 2 is not met. The elaboration suggested, drawing from work by Palermos (2015), was as follows: a group g knows a moral proposition only if the group’s truth-filtering mechanisms are not only sufficient but also *necessary* for getting to the truth of the matter as to whether p (for some moral proposition, p).

The combination of Bird's distributed account of group knowledge with a Palermos-style construal of a credit condition captures an important intuition about credit: the group deserves the credit if no subset of individuals could get the desired result alone. And, indeed, in some cases of moral knowledge, that will be the case, especially perhaps difficult moral knowledge (e.g., as in the case of DISAGREEMENT, or perhaps even in the scientific case of CERN discovering the Higgs Boson through widely distributed collaboration).

The view, however, struggles when it comes to making sense of *easy* moral knowledge. To make this idea concrete, consider the following case:

CREDIT SWAMP: A bioethics policy committee, C, is tasked with determining whether it is morally acceptable to allow, as a method of deterrence, the whipping of young children in hospitals who do not follow the doctors' advice. The committee realize that this will be a short meeting. In a manner that satisfies all of Bird's three conditions—including a truth-filtering condition—the committee, working collaboratively and through the normal distribution of tasks across committee members, produces a (very) prompt and professional report detailing why the answer is 'no'.

CREDIT SWAMP looks *ex ante* like a clear-cut case of distributed moral knowledge. The problem is that there's no straightforward way for a proponent of distributed knowledge to make sense of this while, *at the same time*, upholding Parity Principle 2. In short, the situation is this: it's simply *false* that, in CREDIT SWAMP getting the right result could only be collectively achieved. Granted, the tasks are *in fact* distributed across the members of the group as in a typical case of distributed knowledge, but given how easy the moral question is under consideration, *any one individual* we may assume would have been able to get this correct result (i.e., that children shouldn't be whipped, even if it were a successful deterrent!). But once this point is appreciated, the prospects that the credit condition in this case is satisfied look dim: after all, the truth-filtering mechanisms of the group were on display here, but is the group's getting it right *primarily* creditable to this distributed mechanism? It's not true that the result could only have been collectively achieved through such collective mechanisms. Given the widespread agreement both in individual beliefs and individual abilities, each has suitable epistemic coverage to do the relevant cognitive work that other members of the group happen to be doing.

The proponent of distributed moral knowledge might press back against our dilemma and simply *deny* that any moral knowledge is *easy* moral knowledge. Perhaps, as this line of thought would go, all moral knowledge is difficult. We reject this claim. However, our response to the worry doesn't require that we do. In fact, all that's needed to generate problems for a proponent of distributed knowledge who wants to uphold Parity Principle 2 is to point out that some moral knowledge is easy enough to come by that a collective effort is not needed to achieve it, even if a collective effort is *used* to achieve it. Our CREDIT SWAMP case is meant to be such an example—e.g., where the distributed filtering abilities are reliable and in fact used to get to the truth but—given the wide coverage of the individual abilities and beliefs that bear on the target proposition—superfluous.

6. CONCLUDING REMARKS

Our overarching aim here has been to motivate some new puzzles to do with defeat in collective moral epistemology, puzzles which have ultimately revealed collective moral knowledge to be surprisingly fragile compared to individual-level moral knowledge. We take as a starting point, in individual epistemology, a distinction between (i) moral knowledge, and (ii) the kind of (non-moral) knowledge that is easily transferable by testimony. On the assumption that there *is* moral knowledge, we've argued that the best way to countenance it is with reference to the kind of 'credit condition' that robust virtue epistemologists, albeit mistakenly, think that *all* knowledge, moral as well as non-moral, must answer to. We then argued, by parity, from the individual to the collective level as follows: if moral knowledge demands the satisfaction of a credit condition at the individual level, then the same should hold at the collective level. In particular, the key 'parity principle' we defend maintains the following: that a group g (non-summatively) knows a moral proposition, p_m only if g 's believing p_m truly is primarily creditable to g 's exercise of (morally relevant) cognitive ability.

With this parity principle in play, we then outlined the two most prominent template accounts of non-summativist moral knowledge: the joint acceptance account, notably defended by Gilbert, and the distributed model, defended by Bird. We showed what it would take to satisfy the credit parity principle on moral knowledge on each of these accounts, with reference to their substantive differences. The upshot was, in each case, a dilemma an analogue to which we don't find at the individual level. Joint acceptance accounts turned out to be capable of vindicating group-level moral

knowledge in a way that satisfies the parity principle only at a substantial cost—*viz.*, by making group level moral knowledge highly fragile, so fragile that (it was argued) a single intragroup dispute about methodology would (in the case DISAGREEMENT) be sufficient for undermining it. Distributed accounts faced a similar dilemma: for proponents of distributed group moral knowledge, the dilemma was to either reject the parity principle (and thus sever the connection at the group level between moral knowledge and credit that we find well motivated at the individual level) *or* retain that connection and accept that group moral knowledge is undermined (via ‘credit swamping’) whenever the moral knowledge at issue is too easy to require a collective effort.

These puzzles place the burden of argument on non-skeptical collective moral epistemologists to show us how (non-summative) knowledge is possible, and in a way that avoids serious theoretical costs. We hope to have shown what some of these costs are, and what kinds of considerations the non-skeptical collective moral epistemologist will need to grapple with in order to meet the challenge in a plausible way.³⁸

REFERENCES

- Baehr, J. (2011). *The Inquiring Mind*. Oxford: Oxford University Press.
- Bird, A. (2010). 'Social Knowing: The Social Sense of 'Scientific Knowledge'', *Philosophical Perspectives*, 24(1), 23–56.
- Brady, M. S., & Fricker, M. (Eds.). (2016). *The Epistemic Life of Groups: Essays in the Epistemology of Collectives*. Oxford: Oxford University Press.
- Carter, J. A. (2015). 'Group Knowledge and Epistemic Defeat', *Ergo* 2(28), 711-735.
- (2008). 'Robust Virtue Epistemology as Anti-Luck Epistemology: A New Solution', *Pacific Philosophical Quarterly*, 97(1), 140-155.
- Carter, J. A., & Pritchard, D. (2015). 'Knowledge-How and Epistemic Value', *Australasian Journal of Philosophy*, 93(4), 799-816.
- Carter, J. A., & Navarro, J. (2017). 'The Defeasibility of Knowledge-How', *Philosophy and Phenomenological Research*, 95(3), 662-685.
- Chrisman, M. (2012). 'Epistemic Expressivism', *Philosophy Compass*, 7(2), 118-126.
- De Ridder, J. (2014). 'Epistemic Dependence and Collective Scientific Knowledge', *Synthese*, 191(1), 37-53.
- (2018). 'Representations and Robustly Collective Attitudes', J. A. Carter et. al., (eds.), *Socially Extended Knowledge*. Oxford: Oxford University Press, 36–58
- Driver, J. (2006). 'Autonomy and the Asymmetry Problem for Moral Expertise', *Philosophical Studies*, 128(3), 619-644.
- Enoch, D. (2014). 'A Defense of Moral Deference', *The Journal of Philosophy*, 111(5), 229-258.
- Fricker, M. (2010). 'Can There Be Institutional Virtues?', *Oxford Studies in Epistemology* 3, 223-235.
- Gilbert, M. (1987). 'Modelling Collective Belief', *Synthese*, 73(1), 185–204.
- (2000). *Sociality and Responsibility: New Essays in Plural Subject Theory*. Rowman & Littlefield.
- (2002). 'Belief and Acceptance as Features of Groups', *Protosociology*, 16, 35–69.
- (2013). *Joint Commitment: How We Make the Social World*. Oxford: Oxford University Press.
- (2014). 'Collective Epistemology', *Oxford Bibliographies Online*
doi:10.1093/acprof:oso/9780199970148.003.0008
- Graham, P. J. (2006). 'Can Testimony Generate Knowledge?', *Philosophica*, 78, 105-127.
- Greco, J. (2003). 'Knowledge as Credit for True Belief', In M. DePaul & L. Zagzebski (eds.), *Intellectual Virtue: Perspectives From Ethics and Epistemology*. Clarendon Press. pp. 111-134.
- (2010). *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity*. Cambridge: Cambridge University Press.
- Grimm, S. R. (2014). 'Understanding as Knowledge of Causes', In *Virtue epistemology naturalized* (pp. 329-345). Dordrecht: Springer.
- Hills, A. (2009). 'Moral Testimony and Moral Epistemology', *Ethics*, 120(1), 94-127.
- (2015). 'Understanding-Why', *Noûs* 49(2), 661-688.
- Hopkins, R. (2007). 'What is Wrong with Moral Testimony?', *Philosophy and Phenomenological Research*, 74(3), 611-634.
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Jones, K. (1999). 'Second-Hand Moral Knowledge', *The Journal of Philosophy*, 96(2), 55-78.

- Kallestrup, J. (2016). 'Group Virtue Epistemology', *Synthese*, 1-19. doi: 10.1007/s11229-016-1225-7.
- Kallestrup, J., & Pritchard, D. (2014). 'Virtue Epistemology and Epistemic Twin Earth', *European Journal of Philosophy*, 22(3), 335-357.
- Kvanvig, J. L. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.
- Lackey, J. (2007). 'Why We Don't Deserve Credit for Everything We Know', *Synthese*, 158(3), 345-361.
- (2014). (Ed.). (2014). *Essays in Collective Epistemology*. New York: Oxford University Press, USA.
- (2016). 'What is Justified Group Belief?', *Philosophical Review*, 125(3), 341-396.
- Lillehammer, H. (2014). 'I—Moral Testimony, Moral Virtue, and the Value of Autonomy', In *Aristotelian Society Supplementary Volume*, 88(1), 111-127.
- List, C. & Pettit, P. (2006). 'Group Agency and Supervenience', *Southern Journal of Philosophy*, 44, 85-105.
- (2011). *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.
- Mackie, J. (1990). *Ethics: Inventing Right and Wrong*. Penguin UK.
- Mathiesen, K. (2006). 'The Epistemic Features of Group Belief', *Episteme*, 2(3), 161–175.
- Montmarquet, J. A. (1992). *Epistemic Virtue and Doxastic Responsibility*. Rowman & Littlefield.
- Nickel, P. (2001). 'Moral Testimony and its Authority', *Ethical Theory and Moral Practice*, 4(3), 253-266.
- Olson, J. (2014). *Moral Error Theory: History, Critique, Defence*. Oxford: Oxford University Press.
- Palermos, S. O. (2015). 'Active Externalism, Virtue Reliabilism and Scientific Knowledge', *Synthese*, 192(9), 2955-2986.
- Plantinga, A. (1993). *Warrant and Proper Function*. Oxford: Oxford University Press.
- Pritchard, D. (2012). 'Anti-Luck Virtue Epistemology', *The Journal of Philosophy*, 109(3), 247-279.
- Pritchard, D., Carter, J. A., & Turri, J. (2018). 'The Value of Knowledge', *Stanford Encyclopedia of Philosophy*.
- Poston, T. (2016). 'Know How to Transmit Knowledge?', *Noûs*, 50(4), 865-878.
- Schmitt, F. F. (1994). 'The Justification of Group Beliefs', In *Socializing Epistemology: The Social Dimensions of Knowledge*, ed. F.F. Schmitt, 257–87. Lanham, MD: Rowman and Littlefield.
- Silva, P. (2018). 'Justified Group Belief is Evidentially Responsible Group Belief', *Episteme*, 1-20.
- Sliwa, P. (2012). 'In Defense of Moral Testimony', *Philosophical Studies*, 158(2), 175-195.
- Sosa, E. (2000). 'Reliabilism and Intellectual Virtue', In G. Axtell (ed.), *Knowledge, Belief, and Character: Readings in Virtue Epistemology*. Rowman & Littlefield Publishers, pp. 33-40.
- (2009). *A Virtue Epistemology: Apt Belief and Reflective Knowledge (Vol. 1)*. Oxford: Oxford University Press.
- (2015). *Judgment and Agency*. New York: University Press, USA.
- Turri, J. (2011). 'Manifest Failure: The Gettier Problem Solved', *Philosopher's Imprint*, 11(8), 1-11.
- Zagzebski, L. T. (1996). *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.

NOTES

¹ For some recent critiques of moral pessimism, see for example Sliwa (2012) and Enoch (2014).

² See, e.g., Hills (2009), Hopkins (2007) and Lillehammer (2014).

³ For a more thorough discussion of moral testimony, moral deference and its relationship to higher-order evidence, see Lee, Robson and Sinclair (this volume).

⁴ See Lackey (2007) and Pritchard (2012). Cf., Greco (2012) for a reply.

⁵ See Nickel (2001) and Hills (2009) for developments of such an ‘understanding reply’ to the problem of moral deference.

⁶ There is also logical space for the line that one epistemically should not have (mere) moral knowledge. Because this conflicts with platitude that knowledge is at least epistemically permissible, we’ll bracket this possibility.

⁷ Kallestrup (forthcoming) maintains that all collective knowledge requires collective achievement.

⁸ It is, however, difficult to acquire on the two models for very different reasons.

⁹ See, e.g., Mackie (1977) and Olson (2014).

¹⁰ We say ‘traditionally construed’ because it is available to the non-cognitivist to embrace, along with ethical non-cognitivism, also *epistemic* non-cognitivism, according to which knowledge attributions are expressions of epistemic approval. For discussion of this kind of a view, see Chrisman (2012).

¹¹ Similar examples have been raised in recent work in the epistemology of know-how. See, for example, Carter and Pritchard (2015) and Poston (2016).

¹² For discussion of grasping as a kind of ability, see Kvanvig (2013) and Grimm (2014).

¹³ This is of course not to imply that moral knowledge is always or even generally more difficult than non-moral knowledge to acquire. Some non-moral knowledge is *obviously* more difficult to acquire than some of the easiest moral knowledge (e.g., Goldbach’s conjecture versus the wrongness of gratuitous evil). Rather, the idea is that moral knowledge categorically requires ability in a way that non-moral knowledge does not.

¹⁴ See, for example, Greco (2010; 2012) and Sosa (2009; 2015) for representative defences; cf., Zagzebski (1996) for a stronger version of the position which requires that these abilities be not only reliable dispositions, but also that they feature distinctive motivations.

¹⁵ See, for example, Greco (2010, Ch. 6), Sosa (2010), Turri (2011), and Carter (2016) for representative discussions. Robust virtue epistemology also has the advantage of explaining why knowledge, qua achievement, has the kind of value often ascribed to it. For discussion, see Pritchard et al. (2014).

¹⁶ For related points, see Pritchard (2012) and Kallestrup and Pritchard (2014).

¹⁷ See, for example, Hills (2015) for a development on the view of abilities and moral understanding from her (2009).

¹⁸ See, for example, Sosa (2010) for an account of virtue reliabilist abilities as *competences*, which are dispositions with three components: seat/shape/situation (see also Sosa (2015)). A canonical presentation of virtue reliabilist abilities is given in Greco (2010).

¹⁹ For discussion, see Graham (2006). While the idea that one can know mathematical propositions via testimony is widely accepted in the epistemology of testimony, it is contentious in the philosophy of mathematics, according to which proof is required for knowledge. Thanks to Justin Clarke-Doane for raising this point.

²⁰ What goes for AI plausibly goes for other kinds of ‘meta-incoherence’ cases in the classic reliabilist literature. For discussion, see, e.g., Sosa (2000).

²¹ For a recent discussion of how abilities can be defeated, see Carter and Navarro (2017), who engage with this issue in the context of anti-intellectualism about know-how.

²² See Gilbert (2014) for an overview.

²³ See especially the collections of papers in Lackey (ed.) (2014) and Brady and Fricker (eds.) (2016).

²⁴ See, along with Lackey (ed.) 2014, and Brady and Fricker (eds.) (2016), Gilbert (2014).

²⁵ Note that ‘belief’ is not an epistemic condition, per se.

²⁶ For critical discussion on this point, see Mathiesen (2006) and Carter (2015).

²⁷ See, however, Lackey (2016) for an interesting kind of amalgamation of summativism and non-summativism in an account of group justification. For a criticism of Lackey’s account, see Silva (2018).

²⁸ Though such an account has been defended. See, for example, Schmitt (1994). Cf., however, Lackey (2016, 346-7).

²⁹ For a more sophisticated version of this view, see de Ridder (2014).

³⁰ See Bird (2010, §4.3) for discussion of how this kind of functionalist approach draws from Durkheim's functionalism and the organismic analogy.

³¹ Our italics.

³² DiPaolo (this volume) discusses how a group of fanatics, given their dogmatic beliefs, might be more resilient to higher-order defeat than other more epistemically virtuous groups.

³³ For a more precise characterization of higher-order evidence which is congenial to the project we pursue here, see Bennett (this volume, section 2).

³⁴ This is a topic of longstanding debate in individual epistemology, in the literature on virtue responsibilism. See, for example, Montmarquet (1992), Zagzebski (1996) and Baehr (2011).

³⁵ For the sceptical significance of moral disagreement as an instance of higher-order evidence, see Terman and Risberg (this volume). See also Turnbull and Sampson (this volume) for a steadfast account of moral disagreement.

³⁶ For a recent discussion in *Nature* of some of the ethical quandaries surrounding perfusion on mammalian brains more generally, see <https://www.nature.com/articles/d41586-019-01168-9>.

³⁷ We are neutral (here, and elsewhere) with respect to how these claims about knowledge would best interface with a linguistic theory of knowledge attributions.

³⁸ Thanks to Justin Clarke-Doane and Michael Klenk for helpful comments on a previous version of this paper.