

ORIGINAL ARTICLE

Responsibility in epistemic collaborations: Is it me, is it the group or are we all to blame?

S. Orestis Palermos

Cardiff University, Cardiff, UK

Correspondence

S. Orestis Palermos, Cardiff University,
Cardiff, UK.

Email: palermosS@cardiff.ac.uk

Abstract

According to distributed virtue reliabilism (Palermos, 2020b), epistemic collaborations—such as Transactive Memory Systems and Scientific Research Teams—can be held epistemically responsible at the collective level. This raises the question of whether participants of epistemic collaborations are exempt from being held *individually* responsible. In response, this paper explores two possible ways in which attributions of individual responsibility may still be appropriate within epistemic collaborations: (I) Individuals can be held *epistemically* responsible for their individual shortcomings, but no amount of individual epistemic responsibility can replace collective epistemic responsibility. (II) Even if it is denied that participants of epistemic collaborations can be held *epistemically* responsible at the individual level, they may be held *structurally*, perhaps *morally*, and even *legally* responsible at the individual level for breaking joint commitments necessary for the effective coordination of the epistemic collaboration.

1 | INTRODUCTION

It is not uncommon to assume that the manifestation of responsibility is a necessary property for an entity to qualify as an agent. Thus, it not surprising that the notion of responsibility has assumed a central role in debates on collective mentality and collective agency—i.e., on whether collectives may qualify as *irreducible* cognitive agents in themselves. For example, several authors (who may otherwise support the idea of group cognition) have argued that collectives cannot qualify as cognitive agents in themselves by pointing out that collectives cannot be held responsible

in themselves (for example, as we shall soon see in more detail, Giere (2006), Huebner (2013), Michaelian & Arango-Muñoz (2018) all argue along these lines). Conversely, being in a position to claim that a collective *can* be held responsible in itself can go a long way towards motivating the claim that it can also be treated as an entity with an agency of its own (i.e., an emergent agency, over and above the agencies of its individual members).¹

Recently, I have argued (Palermos, 2020b) that *epistemic collaborations* such as scientific research teams and Transactive Memory Systems (TMSs) (i.e., groups of two or more individuals who collaboratively encode, store and retrieve information (Wegner et al., 1985) can manifest—at the collective level—the kind of *epistemic* responsibility necessary for generating propositional knowledge.² Thus, epistemic collaborations can qualify as collective *epistemic* agents in themselves, such that any resulting (collaboratively produced) knowledge can be attributed to them. If correct, this claim should be good news for proponents of collective propositional knowledge.³

Yet, a potential problem, concerning individual responsibility, might lurk beneath the surface: If epistemic collaborations can be held, in themselves, responsible for their failures and achievements, what does this mean with regards to their members' individual responsibility?⁴ That is, in cases where we attribute collective responsibility, can we not hold the participants of epistemic collaborations individually responsible as well? Or, if we can, in what sense can we hold them responsible and to what extent? Additionally, how does their individual responsibility relate to the epistemic responsibility associated with the collective? These are the questions I aim to address in this paper.

In more detail, section 2 sets the stage by explaining how the notion of responsibility has often been at the forefront of the debate on collective agency and mentality. Section 3 summarises my account of collective epistemic responsibility in the case of epistemic collaborations. Section 4 details the questions I here aim to address: In what sense and to what extent, if at all, can we hold the members of epistemic collaboration individually responsible? Then, section 5 provides an initial response: If it is possible to trace some epistemic failing back to one or more individual members of the collaboration, then we may hold the individual member(s) epistemically responsible at the individual level; however, at the same time, we should keep in mind that no amount or aggregate of individual epistemic responsibility can make up the entirety of the epistemic responsibility associated with epistemic collaborations, for there is always some amount of collective epistemic responsibility that needs to be attributed to the collaborations themselves. Initially, this may sound like a promising response; section 6, however, raises doubts about the possibility of attributing any amount of individual *epistemic* responsibility within epistemic collaborations. Thus, section 7 and 8 develop an error theory, which holds that our intuitions regarding the attribution of individual *epistemic* responsibility within epistemic collaborations may be explained away by the ability to attribute *structural*, perhaps *moral*, and even *legal* responsibility to the participants of epistemic collaborations. Perhaps, that is, it was not epistemic but structural, moral and legal individual responsibility that one had in mind when insisting that, within epistemic collaborations, responsibility should not be assigned at the collective level alone. Finally, section 9 concludes the discussion by summarising the problematic and offering an overview of the different positions one may hold in response.

2 | COLLECTIVE AGENCY AND RESPONSIBILITY

Among the most interesting and controversial topics within philosophy of mind and cognitive science concerns the debate surrounding the notions of distributed cognition and collective

mentality:⁵ Can groups of people be treated as irreducible distributed cognitive systems, and can they manifest mental or cognitive properties such as beliefs, desires, skills or intentions at the collective level?

In the previous section, it was noted that a common reason for which even proponents of the view doubt that such collectives can be treated as irreducible cognitive agents has to do with the notion of responsibility. For example, Ronald Giere (2006), who has argued that the way many scientific research teams *produce* knowledge can be best analysed in terms of distributed cognition, has nevertheless argued against the idea that, in such cases, there is a group-level agent who possesses the corresponding knowledge. Among Giere's reasons for this denial is the claim that cognitive agency is "tightly bound up" (2006, p. 215) with the concept of responsibility, which, Giere seems to think, cannot be manifested at the group level—though no sustained argument is provided for this claim. Giere also provides no specific reason why he thinks that responsibility is an important (perhaps even necessary) property for agency. Other authors, however, who make similar remarks regarding the link between agency, responsibility and collective agency, point out that ascriptions of agency require that any candidate entities be such that they can be held responsible for their actions, claims, cognitive successes and failures.

Bryce Huebner (2013), for example, who has extensively argued for the notion of collective mentality, expresses doubts about whether it is possible to treat collectivities as full-fledged mental entities, or as he refers to them, as entities with 'maximal mentality.' As Huebner (2013) explains: "This requires demonstrating that some collectivities can be held responsible for their claims and actions, and that this is possible in much the same way that we hold individuals responsible for their claims and actions" (p. 212). As he further notes, however, it is doubtful we can hold collectives responsible in this way.

Likewise, Michaelian and Arango-Muñoz's (2018) express similar scepticism against the idea that TMSs can *properly* qualify as cognitive agents in themselves. Their reason for doubting this, as they explain, is that:

There is no clear sense in which a TMS, for example, might be assigned responsibility for its cognitive success and failures, and it is in this sense we have suggested that extended and distributed memory systems do not qualify as cognitive agents: they may be cognitive agents simpliciter, but they are not responsible cognitive agents. (p. 242).

On Michaelian and Arango-Muñoz's nuanced view, we may distinguish between two notions of cognitive agency—a rich one, that involves responsibility for one's cognitive successes and failures, and a weaker one that doesn't. Nevertheless, as their remarks make apparent, for proponents of group mentality and *a fortiori* for its opponents too, in order for a group to qualify as a cognitive agent proper—i.e., as an agent with rich agency—it must be capable of manifesting responsibility.

The above suggests, therefore, that there is an intuitive link between strong ascriptions of agency and the ability of the corresponding entity to be held responsible for its actions, claims, cognitive successes, and failures. It is obvious, then, that proponents of collective agency would be better off were they able to demonstrate that (some) collectives can manifest this type of responsibility at the collective level.

3 | JUSTIFICATION IN EPISTEMIC COLLABORATIONS

Building on the mainstream epistemological approach of agent reliabilism (Greco, 1999; 2003; 2010),⁶ I have recently proposed (Palermos, 2020b) *Distributed Virtue Reliabilism* to account for the way epistemic collaborations—such as Scientific Research Teams and Transactive Memory Systems (TMSs)—produce knowledge and justification.

A central claim of this view is that, contrary to the points outlined in the previous section, collectives that qualify as epistemic collaborations can manifest epistemic responsibility at the collective level. This claim can have several ramifications. For example, it has the potential to significantly motivate the existence of collective epistemic agents, but it also raises the central question of this paper: If it is appropriate to hold an epistemic collaboration responsible at the collective level, is it nevertheless still appropriate to also hold the individual members of the collaboration individually responsible; and, if so, in what sense? To answer, it will be helpful to first summarise in this section the argument for Distributed Virtue Reliabilism, so as to get a better grasp of the claim that epistemic collaborations can be held responsible at the collective level.

The initial step in introducing Distributed Virtue Reliabilism was to first highlight the distinctive feature of epistemic collaborations: i.e., their members' ongoing bidirectional interactions. The argument then proceeded by developing a hybrid account that combines Greco's agent reliabilism (1999, 2003, 2010) from mainstream epistemology with a mathematically informed version of the hypothesis of distributed cognition from philosophy of mind and cognitive science (Palermos, 2016).

In more detail, an important, starting assumption of the argument in (Palermos, 2020b) is that epistemic collaborations, such as scientific research teams and TMSs, are complex dynamical systems ultimately to be modelled on the basis of Dynamical Systems Theory (DST). With this assumption in place, the argument then proceeded to employ insights and formalisms offered by this widely used area of mathematics. First, it noted that, according to DST, when group members engage in continuous reciprocal interactions, *new properties emerge*, which belong to an integrated, distributed system, consisting of all interacting individuals. In the specific case of epistemic collaborations, group members' ongoing interactions can allow the group to manifest the emergent properties of *self-organisation* and *self-regulation*, which, respectively, allow the group to perform in an *epistemically reliable* and *responsible* manner. It is worth expanding on these points (if only briefly):

Self-organization is a diachronic process that extends over time. Initially, during its first stages, the members of the collaboration mutually interact, trying out several channels of action and communication, until they evolve into a relatively stable configuration. Normally, achieving this stable (though still malleable) structure signifies that group components have adapted to each other by limiting their co-ordinated interactions to those that make possible the regular attainment of their end (Heylighen et al., 2004). Thus, in the case of epistemic collaborations, self-organisation contributes over time—on the basis of several cycles of trial and error—to the reliability of the collective belief-forming process: If the dynamics of the group do not produce a preponderance of true over false beliefs, the collective does not regularly achieve its end; in result, it keeps rearranging into different configurations until it stabilizes into a successful (i.e., sufficiently reliable) one. Otherwise, if this is not feasible, eventually, the collective (most likely) disintegrates.⁷ So, overall, by mutually interacting, group members participate in a process of self-organisation, which is normally conducive to forming a group structure that can reliably produce true beliefs.

The epistemic benefits of members' ongoing bidirectional interactions do not stop at promoting self-organisation, however: They can also allow groups to *self-regulate*. Self-regulation is a synchronic process that takes place at the time of performance. Its main function is to allow the collaboration to produce beliefs in an epistemically responsible manner. It too works by relying on the continuous interactivity between the members of the group, which normally allows them to keep monitoring each other's performance in the following way: When there is something wrong with the overall process, the members' coordinated interactions get out of step; when such coordination breakdown occurs, the problem is likely to be spotted by at least one member of the coordinated performance, which then leads the group to issue an appropriate response. For example, if the members of a TMS are in synch, they will be able to notice whether there is anything wrong with their exchange. For instance, they may observe that their partner's non-verbal or paraverbal cues indicate that they misremember or that they are confused. Likewise, in a scientific experiment, if the experimentalist delivers data with weak peaks, the mathematician will spot this and will alert the physicist about it. As the above suggests, such synchronic monitoring processes are mainly powered by the individuals' coordinated activity. Essentially, if, on their basis, no member expresses doubts or concerns about the overall process as it unfolds over time, the group can qualify as epistemically responsible for its performance *by default*; and interestingly, this can be so, even if no member of the group, or even the group as a whole, fully grasps how its performance works.⁸

Returning now to the overall account, the next step was to point out that according to agent reliabilism, epistemic reliability and responsibility are the two main components of epistemic justification. Accordingly, given (1) the way DST understands group-level emergent properties as the product of members' ongoing bidirectional interactions and given that (2) to form beliefs, epistemic collaborations have their members engaging in continuous reciprocal interactions to *self-organise* and *self-regulate*—processes which contribute to the reliability and responsibility of the corresponding beliefs, respectively—we get the following result: When epistemic collaborations are epistemically responsible and reliable, the group's epistemic reliability and responsibility—i.e., the group's justification—are emergent properties belonging to the group as a whole.⁹ Consequently, according to Distributed Virtue Reliabilism (Palermos, 2020b), any resulting knowledge will be collective knowledge, attributable to the whole group, consisting of all the densely interacting (i.e., collaborating) individuals.

Of course, this is but a rough summary of Distributed Virtue Reliabilism and for the argument in its support; but the intention here has only been to demonstrate some of the main points involved. Specifically, for present purposes, it suffices to have outlined the sense in which, on Distributed Virtue Reliabilism, members' collaborative interactions normally allow the group to manifest epistemic responsibility *at the group level*. Put another way, on this view, when epistemic collaborations are responsible, their responsibility is an emergent property that belongs to the group *as a whole*.

4 | ARE INDIVIDUALS EXEMPT?

Now, if the foregoing is correct, then significant dialectical space is made for the possibility of group agency in the case of epistemic collaborations—at least as far as Giere's, Huebner's, Michaelian and Arrango-Muñoz's worries are concerned. But, on the flipside, there is a potential worry that needs to be directly addressed. According to the above, when things go well or badly in epistemic collaborations, responsibility can be assigned to the group *as a whole* (i.e., there is an

irreducible type of collective responsibility that belongs to the group in itself). But does this mean that none of the individual members can be held responsible? Put another way, are individuals exempt?

To some, this might sound like the right thing to say. To most, however, it is likely to sound problematic (see, for example, Winsberg et al, 2014). While many proponents of collaborative knowledge and group agency might be willing to ascribe responsibility to groups as wholes, they will insist that it is highly counterintuitive to entirely deny responsibility at the individual level. I find the force of this worry to be strong and I think it is incumbent on any theory of group agency to explain how the individual members of the group can be held responsible too.

The question I need to address then is how one can argue for this in a way that agrees with the notion of collective epistemic responsibility outlined above. To do so, in the remainder of the paper, I will attempt to offer two possible explanations of, and clarify how, in the case of epistemic collaborations, responsibility can be attributed at both the individual and the collective level simultaneously. In doing so, we will also get a better understanding of the notion of collective epistemic responsibility.

5 | INDIVIDUAL AND COLLECTIVE EPISTEMIC RESPONSIBILITY

Perhaps the most straightforward way to think about the situation is along the following lines: Grant, as we have, that the responsibility associated with the overall distributed belief-forming process belongs to the collaboration as a whole, but note that if it is possible to trace an error back to a specific participant, then epistemic responsibility can be attributed to them for failing to successfully carry out *their part of the process*.

I say *if* it is possible to trace back an error to a specific participant, because it might not be: Collaboratively produced knowledge and the associated epistemic responsibility are the products of dense interactions between group members. Thus, a potential worry springs from the fact that such cooperative processes make it impossible to use “piecemeal decomposition and additive reassembly” (Clark, 2008, p. 116) to trace parts of the final product back to specific, isolated contributions by group members. That is to say: collaborative knowledge is resistant to an aggregative analysis, because the final product is an epistemic amalgam that makes it impossible to isolate how the knowledge or actions of specific individual members of the group were involved. Indeed, if the picture outlined in section 3 is correct, the resulting collaborative knowledge cannot be the sum of the knowledge possessed by the members of the group, because its two main epistemic properties (i.e., reliability and responsibility) are at least partly the result of the members’ coordinated interactions.

Nevertheless, even though, on Distributed Virtue Reliabilism, it is impossible to break down collaborative knowledge in terms of individual epistemic contributions, it might still be possible—at least sometimes—to highlight significant individual failings, such as the introduction of a false belief that propagated in the group collaboration with negative epistemic results. In such cases, it might make good sense to attribute epistemic responsibility to the corresponding individual. The crucial question, however, is: In what sense and to what extent?

In response, we should first note that, given the foregoing points on the nature of epistemic collaborations, no individual member or set of individual participants can be individually attributed the entirety of the responsibility for the collaboration’s failure—only a certain part of it. That part may still be a considerable amount of responsibility, but it cannot be the entirety of it. The reason for this is that if we are indeed talking about an epistemic collaboration that is otherwise capable

of self-regulating, then the sum of individual epistemic responsibility cannot make up the entirety of epistemic responsibility: No matter how big the role of some individual(s) might be in the failing of an epistemic collaboration, there is always some responsibility that should be directed at the collaboration itself. To claim otherwise would amount to actively ignoring that the collaboration's self-regulatory channels of interaction have in some way failed, and that, in the absence of this failing, the individual shortcoming(s) could have been detected and dealt with.

In other words, we should make sure that the attribution of epistemic responsibility will not be exhausted at the individual level. Instead, when it comes to the failings of epistemic collaborations, we must not forget about collective responsibility, so as to highlight that ameliorative measures should be directed at the group's organization too. Otherwise, we run the risk of addressing all individual shortcomings—by replacing, for example, all individually responsible members with new ones—while leaving the epistemic collaboration vulnerable to similar organizational shortcomings in the future. As noted above, epistemic collaborations are supposed to be self-organizing and self-regulating entities that can thereby 'learn' from their mistakes (in fact, the ability to learn in this way is precisely where their collective responsibility arises from).¹⁰ Thus, if we failed to hold epistemic collaborations responsible at the collective level (or, worse, if epistemic collaborations failed to do so themselves) would undermine what seems to be their most distinctive feature—i.e., their dynamically self-organizing and self-regulating structure. Eventually, such a failure would likely lead to their disintegration.

We can put the above points together in the form of the following principle:

Distinction Between Emergent and Aggregative Responsibility: The aggregate of all individual epistemic responsibility never amounts to the collaboration's emergent epistemic responsibility.

In summary, collective responsibility might be the only kind of responsibility that we can intelligibly attribute in many epistemic collaboration failings. Nevertheless, sometimes, it might still be possible to also attribute individual responsibility to certain members of the group. If this is possible, we should also not forget: (1) No amount of individual responsibility can replace collective responsibility; (2) Therefore, we need to also attribute responsibility at the collective level—otherwise we effectively fail to treat the epistemic *collaboration* as such.

6 | SCEPTICISM ABOUT INDIVIDUAL EPISTEMIC RESPONSIBILITY WITHIN EPISTEMIC COLLABORATIONS

The above response allows for the attribution of individual epistemic responsibility in epistemic collaborations, while making no concessions with regards to the attribution of collective epistemic responsibility. Nevertheless, proponents of collective epistemic responsibility may insist that the attribution of individual epistemic responsibility is unwarranted in the case of epistemic collaborations. It could be pointed out, for example, that participation in an epistemic collaboration amounts to an invitation to turning one's attention away from regulating one's individual epistemic states and focusing instead on one's participation in the group. Should everything go to plan, any individual errors or shortcomings will be spotted and eradicated by means of the coordinated activity of group members, and doing so is the sole responsibility of the group *as a whole*. Individual members, therefore, should not be held epistemically responsible for their individual shortcomings or their propagated effects on the distributed belief-forming process.

As an illustration of this point, take the example of Wikipedia. With regards to its scientific entries, Wikipedia is considered to be nearly as reliable as Encyclopaedia Britannica.¹¹

Nevertheless, in the case of Wikipedia, the reliability of these entries cannot be accounted for in terms of the reliability of the people writing them. Unlike Encyclopaedia Britannica, Wikipedia has been operating on the basis of an open editability policy, meaning that anyone (often anonymously) can write, edit and modify its entries, without having to demonstrate any credentials of relevant expertise. Given this operational mode, however, the crucial question is: Where does the reliability of Wikipedia's entries originate from? Apparently, the answer is from 'the power of many eyes' (Noveck, 2007): The huge number of participants in combination with the fact that Wikipedia's entries can be searched by anyone on the Web allows for its entries to continuously receive a degree of quality control that no single individual could ever provide on their own, irrespective of their level of expertise. Thus, anyone can contribute to Wikipedia, because, when people post inaccuracies or falsehoods, most likely, there is someone else who will take notice and deal with them. In other words, Wikipedia generates knowledge not by maximizing the probabilities of contributors posting true statements, but by minimizing the probabilities of the survival of falsehoods.¹²

Now, given Wikipedia's *modus operandi* (and especially its open editability policy), if one posted a falsehood on one of its pages, and this falsehood survived detection long enough to be consumed by other individuals with detrimental epistemic effects, should we hold the individual from whom the falsehood originated epistemically responsible for their action? The worry is that many would point out that epistemic responsibility, in such cases, lies entirely with the epistemic collaboration, rather than with any of its individual members—not even with the one who started it all.

Of course, it is unclear whether denying, in this way, individual epistemic responsibility in epistemic collaborations should be generally accepted (or indeed whether it should even be accepted in the case of Wikipedia). Assuming, however, that there is merit to the above kind of scepticism with regards to individual epistemic responsibility within epistemic collaborations, in the rest of the paper, I will develop an error theory that can explain the sense in which we might be inclined to attribute individual responsibility in epistemic collaborations. Specifically, I will argue that even if we cannot attribute *epistemic* responsibility to the individual members of epistemic collaborations, there is a variety of alternative ways in which we may hold them individually responsible when the collaboration fails epistemically. However, to set the stage for such an error theory, a short detour is first needed to the two main models for accounting for the ability of certain collectives to produce propositional knowledge—*viz.*, the commitment model and the distributed model.¹³

7 | THE DISTRIBUTED MODEL AND THE COMMITMENT MODEL

The distributed model can be seen as a family of views whose common point is their focus on some form of 'distribution' of the epistemic effort, deriving from the target community's, or group's, *division of labour*. Alexander Bird (2010, 2014) and Jeroen de Ridder (2014) have both provided their own versions of the distributed model and my account of epistemic collaborations (2020b) is yet another version of it.

Likewise, the commitment model is a family of views whose common point is their focus on *joint commitments* by the members of the group. Margaret Gilbert (1987, 2000, 2004, 2014), Raimo Tuomela (2004), and Catherine Rolin (2008) have all provided different versions of the commitment model. While the focus of proponents of the commitment model might be directed to different joint commitments, ranging over commitment to specific beliefs, commitment to reasons in support of the beliefs of the group or commitment to background assumptions, they all

rely on something along the following characterisation of a joint commitment towards specific propositions:

Joint Belief

- (i) A group *G* believes that *p* if and only if the members of *G* jointly accept that *p*.
- (ii) Members of a group *G* *jointly accept* that *p* if and only if it is common knowledge in *G* that the individual members of *G* have openly expressed a conditional commitment jointly

to accept that *p* together with the other members of *G*. (Gilbert, 1987, p. 195).

According to Gilbert, one of the advantages of the commitment model is its ability to capture what Emile Durkheim referred to as the ‘coercive power’ of groups. According to Gilbert:

once a group believes that *p*, then, *ceteris paribus*, group members are personally obliged not to deny that *p* or to say things that presuppose the denial of *p* in their ensuing interactions with each other. If someone does say something which implies that *p* is not the case, this person should give some sort of explanation, or qualify the statement, saying something like: “In my own view, *p* is not the case”. A violation of these obligations is understood to be grounds for rebuke (Gilbert, 1987, 193–194).

Of course, as Gilbert further notes, the underlying joint commitment need not be the result of an explicit agreement between group members: “Evidently one can communicate one’s conditional commitment to participate in joint acceptance in quite subtle, nonverbal ways” (Gilbert, 1987, p. 199). Even in such cases, however, the resulting *right to rebuke* is a significant mechanism of group *cohesion*.

8 | THE TWO MODELS, DISTRIBUTED VIRTUE RELIABILISM AND INDIVIDUAL STRUCTURAL, MORAL AND LEGAL RESPONSIBILITY

According to Distributed Virtue Reliabilism, to argue for the collective nature of collaborative knowledge, one only needs to focus on the emergent nature of the collective’s epistemic reliability and responsibility. That is to say: the knowledge produced by epistemic collaborations is collective not because the belief component of the corresponding knowledge is collective. Rather, collaboratively produced knowledge is collective, because *the justification component* (i.e., the reliability and responsibility of the distributed belief-forming process) is collective. Thus, on Distributed Virtue Reliabilism, there is no need to appeal to the commitment model in order to explain the collective nature of collaborative knowledge.

Nevertheless, closer inspection reveals that to fully understand *the workings* of epistemic collaborations we need to employ elements of the commitment model too (Palermos, forthcoming). Briefly, the reason for making this concession has to do with the fact that humans are complicated entities inhabiting an informationally busy and often confusing world. Thus, it is not uncommon to suddenly find ourselves presented with doubts (possibly misleading ones) that call into question what we believe or take ourselves to know. Importantly, such doxastic changes have significant effects on our actions and intentions, such as the kind of inquiries and projects we are prepared to engage in. For example, if a scientist lost faith in one of their paradigm’s metaphysical assumptions, they would be less inclined to participate in collaborative projects that rely on it. Thus, unsurprisingly, engaging in collaborative projects requires agreement on certain standards,

values, assumptions, attitudes toward certain kinds of evidence and other propositions, rejection of which would hinder the collaboration. This is where joint commitments toward certain propositions come into play: They solidify the collaboration by precluding participants from doxastically diverging—i.e., from abandoning beliefs, whose acceptance is essential for sustaining their coordinated epistemic performance.^{14,15} What might such commitments be? Think of propositions such as “the laws of physics are the same across all inertial frames of reference”; “physicalism is true”; “gravity can be explained in terms of quantum mechanics”; “gravitons exist”; “there are parallel universes”; “the naturalist approach to consciousness studies is correct”; “the hard problem of consciousness is not a real problem”; “the symbol-system hypothesis within AI is true”; “artificial general intelligence is possible”.

Thus, I recently pointed out (Palermos, forthcoming), it is important to acknowledge that to pursue and sustain their projects, members of epistemic collaborations often undertake several such joint commitments, violation of which would warrant other members’ rebuke. Note, however, that many of these commitments (such as the ones mentioned above) are propositions for which no (or at least no *direct*) epistemic justification and/or evidence exists. In fact, the absence of any direct epistemic credentials in support of such propositions, which would allow members of epistemic collaborations to embrace them on the basis of purely epistemic grounds, is the very reason participants need to agree—i.e., jointly commit—that they will let them stand as the view of their group. That is to say, in the absence of their joint commitments to such propositions, collaborators could very easily change their minds about them, given that there are no direct epistemic reasons for accepting them. This is an important note, because it means that speaking against the content of such commitments—i.e., openly doubting the relevant propositions—would hardly warrant any *epistemic* criticism on the part of the other members. From a purely epistemic point of view, individuals could very well choose to either accept or reject them. Thus, when collaborators speak or act in a manner that goes against such propositions—which is likely to have a detrimental effect for the epistemic collaboration—the ensuing attribution of irresponsibility and rebuke is unlikely to be epistemic in nature. If that’s correct, however, we need to figure out what the nature of the associated irresponsibility and rebuke really is. Put another way, when one breaks such joint commitments, what kind of obligation has one failed to observe?

Quite plausibly, one may claim that the nature of the relevant obligations and the associated entitlement to rebuke others (when these obligations are not met) is *moral*—especially if one assumes that breaking promises, agreements and, more generally, joint commitments, is always *morally* bad. Interestingly, however, in several places (e.g., 2014, chs. 1, 4, 9, 13, 17), Gilbert denies this claim, for the specific reason that one might be jointly committed to act in evil or impermissible ways.^{16,17} In such cases, Gilbert holds, one may still be *normatively* obliged to act according to their commitment, and they may be rightly rebuked for failing to do so by the rest of the participants. Such failures, however, do not amount to having done anything *morally* wrong, for given the immoral nature of the relevant commitments, *morality*, in such cases, dictates that one fail to act according to them.

But if it is neither epistemic nor moral grounds that bind one to act according to their joint commitments, what is the nature of the relevant obligation? Gilbert notes (2014, p. 224) that, instead of morality, the normative force of joint commitments derives from the participant’s capacity to act rationally—i.e., as Gilbert specifies, in a manner that is reasons-responsive. Specifically, on Gilbert’s view, when agents jointly commit to acting in certain ways, they ‘owe’ each other to act in the agreed manner. Thus, when ‘all else is equal’ (e.g., there are no overriding moral reasons for acting against one’s joint commitment), this ‘owing’ to others should be sufficient reason for participants to act in line with their joint commitment; and failure to do so would rightly invite the

other members' rebuke for acting against what the structure of their group rationally demands of its members. In the absence of a better term, we may here call the obligation to act on one's joint commitment *structural obligation* and the responsibility for failing to abide by this obligation *structural responsibility*.

Admittedly, it is not entirely clear whether breaking joint commitments does not always also amount to a moral failure (on top of the aforementioned rational failure imposed by the group's structure). Whatever one's intuitions on this may be, however, the upshot is that when epistemic collaborations fail or malfunction due to having some of their members' acting 'out of line' by breaking joint commitments, these members can be held *structurally* responsible—and perhaps even *morally* responsible, if one disagrees with Gilbert on this point—at the *individual level*.

Now, before concluding this section, it is also worth raising a further point regarding epistemic collaborations' joint commitments: To sustain their teamwork, epistemic collaborations do not solely rely on joint commitments to accept specific *propositions*. Rather, they also seem to make use of joint commitments toward *intentions to act* in certain ways, much in the way Michel Bratman suggests. According to Bratman (1993, 106), two (and possibly more) people intend to *J* if and only if:

1. (a) I intend that we *J* and (b) you intend that we *J*
2. I intend that we *J* in accordance with and because of 1a, 1b, and meshing subplans of 1a and 1b; you intend that we *J* in accordance with and because of 1a, 1b, and meshing subplans of 1a and 1b.
3. 1 and 2 are common knowledge between us

The reason why such joint intentionality is necessary in the case of collaborative knowledge is because collaboration requires coordination and coordination requires that each member of the team act in such a way that their actions interlock (i.e., mesh) with the actions of other members. Of course, as Bratman further notes, in such cases

I need neither know nor seek to know of all your subplans for us to have a shared intention; nor need we already have arrived at complete, meshing subplans. What is required is that I intend that we *J* by way of meshing subplans. I can so intend even though there are as yet no specific, meshing subplans such that I intend that we *J* by way of them. You and I may not yet have filled in each of our subplans, or we may have filled them in in ways which do not yet mesh. We may have conflicting preferences concerning subplans and be involved in negotiations about how to fill in our plans even while we have already started to *J*. (Bratman, 1993, p. 106)

In other words, the kind of shared intentionality at play here concerns the mutually binding and reinforcing will to keep collaborating and to preserve the collaboration by acting in ways that will facilitate its proper functioning. Thus, if a member of the team starts acting in ways that do not so facilitate the team effort, then other members have the right to reproach them for breaking their commitment to act as a member of the team.¹⁸ Now, as with the aforementioned joint commitments to accept certain propositions, joint commitments to act in certain ways do not seem to involve any kind of individual *epistemic* responsibility but a kind of individual *structural* (and perhaps even *moral*) responsibility. Additionally, however, the associated responsibility may occasionally be much more than that: If joint commitments to act in this manner are ensured, as they may sometimes be, by means of legal contracts, then breaking them might further warrant legal action. In result, when epistemic collaborations fail or malfunction due to having some of

their members acting 'out of line' by breaking such joint commitments, these members may be held structurally, perhaps morally, and even legally responsible at the individual level.

9 | SUMMARY AND CONCLUSION

Relying on previous work, I have here assumed that epistemic collaborations can manifest epistemic responsibility at the collective level. Specifically, I have noted that by means of members' dense bidirectional interactions, epistemic collaborations act as complex dynamic systems capable of self-organising and self-regulating. This, in turn, allows their distributed belief-forming process to be both reliable and epistemically responsible. Given that this kind of epistemic reliability and responsibility is (at least partly) the product of members' ongoing bidirectional interactions, mathematical modelling on the basis of DST would treat them as irreducible, emergent properties.

If correct, this approach would be a significant step towards arguing for epistemic group agency and collective knowledge (in the strongest metaphysical sense possible). Nevertheless, arguing in this way raises an important worry: If, in epistemic collaborations, epistemic responsibility is an emergent collective property, does this mean that individual members are exempt from attributions of responsibility? In other words, is there any room for attribution of individual responsibility in epistemic collaborations?

Assuming that the answer must be positive, I here set out to explore possible ways for making sense of individual responsibility in epistemic collaborations in a way that is in line with the above view on epistemic collaborations. One way to go, I argued, is to note that, *if* it is possible to trace back an epistemic error to a specific participant, then epistemic responsibility can be attributed to them for failing to successfully carry out their part of the process. Nevertheless, I further noted, given that we are willing to treat the corresponding group as an epistemic collaboration, then we must also accept (1) that no amount of individual responsibility can replace collective responsibility; and that (2) we thereby need to also attribute responsibility at the collective level. Otherwise, we run the risk of addressing all individual shortcomings—perhaps by replacing every member of the group—while neglecting possible problems with the collaboration's channels for self-regulation and self-organization: A kind of negligence that would leave the collaboration vulnerable to similar (structural) failures in the future.

Nevertheless, I further noted, one might argue that, in epistemic collaborations, it is impossible to trace back individual epistemic shortcomings; or they may point out that the attribution of individual epistemic responsibility is unwarranted in such cases, because membership to an epistemic collaboration amounts to an invitation to turn one's attention away from regulating one's individual epistemic states and to focusing instead on the epistemic coordination of the group. Even so, I noted, there is an alternative way to make sense of attributions of individual responsibility in epistemic collaborations. To effectively coordinate, members of epistemic collaborations undertake joint commitment towards certain propositions, for which no epistemic justification and/or evidence exists, as well as towards intentions to act in specific ways that are conducive to their coordinated performance. When members fail to honour these commitments, we may hold them structurally, perhaps morally, and even legally responsible. Essentially, this amounts to an error theory that provides a plausible, alternative way to explain away our intuitions on individual epistemic responsibility in epistemic collaborations: Perhaps, after all, it was not epistemic but structural, moral and legal individual responsibility that one had in mind when insisting that, within epistemic collaborations, responsibility should not be assigned at the collective level alone.

Ultimately, whether one should accept both responses or only the second one (thereby being an error theorist about individual *epistemic* responsibility in epistemic collaborations) is a question to which no set answer might always be provided. Rather, the appropriate way to respond might be context-sensitive, depending on the specifics of the epistemic collaboration and the problem at hand. For example, each time, we may need to ask: Is it possible to isolate a specific person's epistemic shortcomings in the case we are interested in; moreover, was there an explicit or, at least implicit, understanding that participating in the collaboration excludes one from accounting for their individual epistemic failures?

Whatever the answer might be in each case, however, what is overall important is that, according to the above, there is a variety of ways in which we can accommodate our intuitions on individual responsibility within epistemic collaborations, even on the assumption that, in such cases, there is also an emergent kind of epistemic responsibility that must be attributed at the collective level.

NOTES

- ¹In what follows I focus on collective *epistemic responsibility* and collective *epistemic agency*. However, similar considerations may apply with regards to collective *moral agency*. See, for example, Copp's (2007) discussion of collective moral responsibility and his attempt to make the case for the 'collective moral autonomy thesis.'
- ²For precursors to this idea, see Palermos and Pritchard (2013, 2016); Palermos (2015, 2017).
- ³An indicative list of proponents of collective propositional knowledge would include the following authors (although, notably, their views on how (at least some) propositional knowledge may amount to collective knowledge differ significantly): Bird (2010, 2014); de Ridder (2014); Gilbert (2004); Palermos (2017, 2020b); Rolin (2008); Tuomela (2004); Wray (2007).
- ⁴Despite what the title of the paper suggests, due to skepticism on basic desert, I consider responsibility only in the forward-looking sense. Any talk of blame or responsibility is here intended in the attributive sense alone and with the sole aim of focusing on the relevant agent for future purposes (excluding all kinds of retribution).
- ⁵See, for example, Barnier et al. (2008); Cooke et al. (2013); Duarte et al., (2013a, 2013b); Heylighen et al. (2004); Hutchins (1996); Palermos (2016, 2020a); Palermos and Tollefsen (2018); Sutton (2008); Sutton et al. (2010); Theiner (2013a, 2013b); Theiner and O' Connor (2010); Theiner et al. (2010); Tollefsen (2006, 2015); Tollefsen and Dale (2012); and Wilson (2005).
- ⁶For my take on agent reliabilism (and virtue reliabilism, more generally), see (Palermos, 2021).
- ⁷I here claim that if the group cannot reconfigure into a sufficiently reliable stable structure, then it will *most likely*—but not necessarily—disintegrate in order to pre-empt a possible worry that usually comes up at this point: It is certainly possible to imagine that an epistemic collaboration could preserve itself despite continuous failures. For example, it is possible, though rather implausible, that there could be a group of people that has secured enough funding and willingness to keep operating, despite repeatedly falling short of producing a preponderance of true over false beliefs. So, the mere existence of a self-organising epistemic collaboration does not conceptually guarantee that, as time goes by, it will either disintegrate or it will become reliable enough... even though, these are the most likely outcomes.
- ⁸Additionally, it is also worth noting that the distinction between self-organization and self-regulation is essentially a theoretical artifact. In reality, in well-functioning epistemic collaborations, every instance of epistemic self-regulation contributes to the process of epistemic self-organization and vice versa: The diachronic structure of the group is continuously shaped by and at the same time shapes its ongoing performance. Interestingly, this also suggests an organic link between the epistemic concepts of collective reliability and collective responsibility within epistemic collaborations, though this topic is well beyond the scope of the present paper.
- ⁹It may be thought that the requisite continuous reciprocal interactions between team members are only present in the case of small groups, such that larger scientific collaborations are excluded. Nevertheless, Knorr-Cetina's (1999) ethnographic study of high energy physics suggests otherwise. For more discussion on this point, see also Palermos (2020b).

- ¹⁰I here intend 'learn' in the same sense that connectionist, machine-learning networks 'learn.' For more on the possible parallels between connectionist networks and self-organising distributed cognitive systems, see also Heylighen et al. (2004).
- ¹¹See Giles (2005). See also encyclopedia Britannica's response (http://corporate.britannica.com/britannica_nature_response.pdf) and *Nature's* counter response (http://www.nature.com/press_releases/Britannica_response.pdf).
- ¹²For further discussion on Wikipedia, see (Palermos, 2017).
- ¹³The distinction between these two models is introduced in Bird (2014).
- ¹⁴See also section 3 of (Weatherall and Gilbert, 2016) where they also note the consolidating power of joint commitments in the case of collaborative research within science. I should note, however, that unlike my view on collaborative knowledge, which focuses on well-organised groups, Weatherall and Gilbert discuss research within larger scientific communities such as the set of scientists who ascribe to string theory. Though I doubt that such wide and loosely organised groups of people can produce collaborative knowledge in my sense of the term, I largely agree with Weatherall and Gilbert's claim that joint commitments play an important coordinating role in instances of collaborative research.
- ¹⁵It is worth noting here that joint commitments towards intentions to *act* in certain ways (as opposed to joint intentions toward specific propositions) appear to also play a similar role in (collaborative) Group Know-How: In the case of collaborative Group Know-How, joint commitments towards certain actions prevent members from *practically* (as opposed to doxastically) diverging. For more details, see section V of (Palermos & Tollefsen, 2018). Upon closer reflection, however, this kind of joint intentions towards specific ways of *acting* seem to also play a similar practical role in the case of epistemic collaborations producing collaborative propositional knowledge (as opposed to collaborative Group Know-How). I say more on this below.
- ¹⁶I am very thankful to Deborah Tollefsen for calling my attention to this aspect of Gilbert's view.
- ¹⁷In Gilbert's (2014) words: "I should emphasize that the rights and obligations in question appear to exist by virtue simply of the existence of a particular commitment that is joint. It may for this reason be misleading to describe them as moral obligations insofar as they need only the will of the parties to bring them into being. This may be capricious or arbitrary, and its content may bear little or no positive relationship to the true, the beautiful, or the good" (p. 35).
- ¹⁸Though Bratman (1993, pp. 110-12) claims that shared intentions to act do not always (i.e., necessarily) bestow on members mutual *obligations* to act accordingly, Bratman is willing to admit that, in most normal cases, joint intentions to act do have normative consequences similar to those generated by the joint commitments that Gilbert is interested in (i.e., joint commitments towards certain propositions).

ACKNOWLEDGEMENTS

I am thankful to the audience of the Cogito Epistemology Workshop at Glasgow University for their feedback on a previous version of this paper. I am also thankful to Deborah Tollefsen and John Greco for their helpful and insightful comments.

REFERENCES

- Barnier, A. J., Sutton, J., Harris, C. B., & Wilson, R. A. (2008). A conceptual and empirical framework for the social distribution of cognition: The case of memory. *Cognitive Systems Research*, 9(1-2), 33-51. <https://doi.org/10.1016/j.cogsys.2007.07.002>
- Bird, A. (2010). Social knowing: The social sense of 'scientific knowledge'. *Philosophical Perspectives*, 24(1), 23-56. <https://doi.org/10.1111/j.1520-8583.2010.00184.x>
- Bird, A. (2014). When is there a group that knows? Distributed cognition, scientific knowledge, and the social epistemic subject. In Jennifer Lackey (Ed.), *Essays in collective epistemology* (pp. 42-63). Oxford University Press.
- Bratman, M. (1993). Shared intention. *Ethics*, 104(1), 97-113. <https://doi.org/10.1086/293577>
- Clark, A., (2008), *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford University Press.
- Cooke, N. J., Gorman, J. C., Myers, C. W., & Duran, J. L. (2013). Interactive team cognition. *Cognitive science*, 37(2), 255-285. <https://doi.org/10.1111/cogs.12009>

- Copp, D. (2007). The collective moral autonomy thesis. *Journal of Social Philosophy*, 38(3), 369–388. <https://doi.org/10.1111/j.1467-9833.2007.00386.x>
- De Ridder, J. (2014). Epistemic dependence and collective scientific knowledge. *Synthese*, 191(1), 37–53. <https://doi.org/10.1007/s11229-013-0283-3>
- Duarte, R., Araújo, D., Folgado, H., Esteves, P., Marques, P., & Davids, K. (2013a). Capturing complex, non-linear team behaviours during competitive football performance. *Journal of Systems Science and Complexity*, 26(1), 62–72. <https://doi.org/10.1007/s11424-013-2290-3>
- Duarte, R., Araújo, D., Correia, V., Davids, K., Marques, P., & Richardson, M. J. (2013b). Competing together: Assessing the dynamics of team–team and player–team synchrony in professional association football. *Human movement science*, 32(4), 555–566. <https://doi.org/10.1016/j.humov.2013.01.011>
- Giere, R. (2006). ‘The Role of Agency in Distributed Cognitive Systems’. *Philosophy of Science*, 73(5), 710–719. <https://doi.org/10.1086/518772>
- Gilbert, M. (1987). Modelling collective belief. *Synthese*, 73(1), 185–204. <https://doi.org/10.1007/BF00485446>
- Gilbert, M. (2000). Collective Belief and Scientific Change, in *Sociality and Responsibility: New Essays in Plural Subject Theory*. (pp. 37–49). Rowman & Littlefield.
- Gilbert, M. (2004). Collective epistemology. *Episteme*, 1(2), 95–107. <https://doi.org/10.3366/epi.2004.1.2.95>
- Gilbert, M. (2014). *Joint commitment: How we make the social world*. Oxford University Press.
- Giles, J. (2005). Internet encyclopaedias go head to head, *Nature*, 438(7070), 900–901. <https://doi.org/10.1038/438900a>
- Greco, J. (1999). Agent Reliabilism, in J. Tomberlin (Ed.), *Philosophical Perspectives* (pp. 273–296). Ridgeview Press.
- Greco, J. (2003). Further thoughts on agent reliabilism: Replies to Cohen, Geivett, Kvanvig, and Schmitt and Lahroodi. *Philosophy and Phenomenological Research*, 66(2), 466–480. <https://doi.org/10.1111/j.1933-1592.2003.tb00275.x>
- Greco, J. (2010). *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity*. Cambridge University Press.
- Heylighen, F., Heath, M., & Van, F. (2004). The Emergence of Distributed Cognition: a conceptual framework. In *Proceedings of Collective Intentionality IV*.
- Huebner, B. (2013). *Macrocognition: Distributed minds and collective intentionality*. Oxford University Press.
- Hutchins, E. (1996). *Cognition in the Wild*. MIT Press.
- Knorr, K. C. (1999). *Epistemic cultures: How the sciences make knowledge*. Harvard University Press.
- Michaelian, K., & Arango-Muñoz, S. (2018). Collaborative memory knowledge: A distributed reliabilist perspective. In M. Meade, A. Barnier, P. Van Bergen, C. Harris, & J. Sutton. (Eds.), *Collaborative Remembering: How Remembering with Others Influences Memory* (pp. 231–247). Oxford University Press.
- Noveck, B. S. (2007). Wikipedia and the Future of Legal Education. *J. Legal Educ.*, 57(1), 3–9.
- Palermos, S. O. (2015). Active externalism, virtue reliabilism and scientific knowledge. *Synthese*, 192(9), 2955–2986. <https://doi.org/10.1007/s11229-015-0695-3>
- Palermos, S. O. (2016). The dynamics of group cognition. *Minds and Machines*, 26(4), 409–440. <https://doi.org/10.1007/s11023-016-9402-5>
- Palermos, S. O. (2017). Social machines: a philosophical engineering. *Phenomenology and the Cognitive Sciences*, 16(5), 953–978. <https://doi.org/10.1007/s11097-016-9489-4>
- Palermos, S. O. (2020a). Bees Do It: Distributed Cognition and Psychophysical Laws. In M. Neese, E. Salas, M. R. Endsley (Eds.), *Foundations and Theoretical Perspectives of Distributed Team Cognition* (pp. 115–132). CRC Press.
- Palermos, S. O. (2020b). Epistemic Collaborations: Distributed Cognition and Virtue Reliabilism. *Erkenntnis*, 87, 1481–1500. <https://doi.org/10.1007/s10670-020-00258-9>
- Palermos, S. O. (2021). System Reliabilism and basic beliefs: defeasible, undefeated and likely to be true. *Synthese*, 199(3), 6733–6759. <https://doi.org/10.1007/s11229-021-03090-y>
- Palermos, S. O. (forthcoming). Collaborative Knowledge: Where the Distributed and Commitment Models Merge. *Synthese*. <https://doi.org/10.1007/s11229-022-03459-7>
- Palermos, S. O., & Pritchard, D. (2013). Extended knowledge and social epistemology. *Social Epistemology Review and Reply Collective*, 2(8), 105–120.
- Palermos, S. O., & Pritchard, D. (2016). The distribution of epistemic agency. In P. Reider, (Ed.), *Social epistemology and epistemic agency: De-centralizing epistemic agency* (pp. 109–126) Rowman and Littlefield.

- Palermos, S. O., & Tollefsen, D. P. (2018). Group know-how. In A. J. Carter, A. Clark, J. Kallestrup, S. O. Palermos, D. Pritachrd (Eds.), *Socially extended epistemology*, (pp. 112–131) Oxford University Press.
- Rolin, K. (2008). Science as collective knowledge. *Cognitive Systems Research*, 9(1-2), 115–124. <https://doi.org/10.1016/j.cogsys.2007.07.007>
- Sutton, J. (2008). Between Individual and Collective Memory: Interaction, Coordination, Distribution. *Social Research*, 75(1), 23–48.
- Sutton, J., Harris, C. B., Keil, P. G., & Barnier, A. J. (2010). The psychology of memory, extended cognition, and socially distributed remembering. *Phenomenology and the Cognitive Sciences*, 9(4), 521–560. <https://doi.org/10.1007/s11097-010-9182-y>
- Theiner, G. (2013a). Onwards and Upwards with the Extended Mind: From Individual to Collective Epistemic Action. In L. Caporael, J. Griesemer, & W. Wimsatt (Eds.), *Developing Scaffolds* (pp. 191–208). MIT Press.
- Theiner, G. (2013b). Transactive Memory Systems: A Mechanistic Analysis of Emergent Group Memory. *Review of Philosophy and Psychology*, 4(1), 65–89. <https://doi.org/10.1007/s13164-012-0128-x>
- Theiner, G., Allen, C., & Goldstone, R. L. (2010). Recognizing group cognition. *Cognitive Systems Research*, 11(4), 378–395. <https://doi.org/10.1016/j.cogsys.2010.07.002>
- Theiner, G., & O'Connor, T. (2010). The Emergence of Group Cognition. In Corradini, A., & O'Connor, T. (Eds.), *Emergence in science and philosophy*. Routledge.
- Tollefsen, D. P. (2006). From extended mind to collective mind. *Cognitive Systems Research*, 7(2–3), 140–150. <https://doi.org/10.1016/j.cogsys.2006.01.001>
- Tollefsen, D. P. (2015). *Groups as agents*. Polity Press.
- Tollefsen, D., & Dale, R. (2012). Naturalizing joint action: A process-based approach. *Philosophical Psychology*, 25(3), 385–407. <https://doi.org/10.1080/09515089.2011.579418>
- Tuomela, R. (2004). Group knowledge analyzed. *Episteme*, 1(2), 109–127. <https://doi.org/10.3366/epi.2004.1.2.109>
- Weatherall, J. O., & Gilbert, M. (2016). Collective Belief, Kuhn, and the String Theory Community. In M. Fricker & M. Brady (Eds.), *The Epistemic Life of Groups* (pp. 191–217). Oxford University Press.
- Wegner, D. M., Giuliano, T., & Hertel, P. T. (1985). Cognitive Interdependence in Close Relationships. In D. W. Ickes (Ed.), *Compatible and Incompatible Relationships* (pp. 253–276). Springer New York.
- Wilson, R. A. (2005). Collective memory, group minds, and the extended mind thesis. *Cognitive Processing*, 6(4), 227–236. <https://doi.org/10.1007/s10339-005-0012-z>
- Winsberg, E., Huebner, B., & Kukla, R. (2014). Accountability and values in radically collaborative research. *Studies in History and Philosophy of Science Part A*, 46, 16–23. <https://doi.org/10.1016/j.shpsa.2013.11.007>
- Wray, K. B. (2007). Who has scientific knowledge? *Social Epistemology*, 21(3), 337–347. <https://doi.org/10.1080/02691720701674288>

How to cite this article: Palermos, S. O. (2022). Responsibility in epistemic collaborations: Is it me, is it the group or are we all to blame?. *Philosophical Issues*, 1–16. <https://doi.org/10.1111/phils.12230>