Towards Collective Self-knowledge

Lukas Schwengerer

Received: 29 April 2019 / Accepted: 11 February 2020
© The Author(s) 2020

Abstract
We seem to ascribe mental states and agency to groups. We say ‘Google knows such-and-such,’ or ‘Amazon intends to do such-and-such.’ This observation of ordinary parlance also found its way into philosophical accounts of social groups and collective intentionality. However, these discussions are usually quiet about how groups self-ascribe their own beliefs and intentions. Apple might explain to its shareholders that it intends to bring a new iPhone to the market next year. But how does Apple know what it intends? How do groups get to know their own mental states? This is the question of collective self-knowledge. I argue that collective self-knowledge is a distinct phenomenon that deserves our attention. In particular I suggest: (1) that we should be interested in collective self-knowledge, because our behaviour indicates that we already engage with collective self-knowledge in practice; (2) that groups can collectively avow, which indicates that they have privileged and peculiar access to their own intentional states; and (3) that collective self-knowledge is not reducible to intentional states of individuals and therefore is an independent explanandum.

My plan is the following: I start with a rough and ready characterization of what I take a social group to be. In part 2 I introduce the phenomenon of self-knowledge. In part 3 I explain why we should be interested in collective self-knowledge and argue that groups indeed have access to their intentional states in a way that seems similar to individual self-knowledge. Part 4 focuses on an argument that shows that collective self-knowledge can be present with no individual member knowing the group’s intentional states. Part 5 gives a short overview of lessons to learn for a theory of collective self-knowledge.
1 Social Groups and Collective Mental States

Social groups (hereafter just groups) come in different shapes and sizes. We find large, structured corporations, middle sized philosophy departments, small clubs and even just three people pushing a car together. All of these are social groups in some sense. They are not just people arbitrarily taken together to treat them as one entity. Groups are taken as a single being for a more substantial reason. They bind their members together in virtue of some form of shared or joint intention or goal.¹ I take a social group to be first and foremost a set of people acting together. Acting together cannot be understood as merely having the same goal, or as a set of people showing some apparently coordinated behaviour. Two people might walk next to each other without walking together (Gilbert 1990). Moreover, some behaviour of masses might look like organized group behaviour, even though it is just a result of individual actions adding up without any shared or joint goal. This has already been highlighted by Weber (1978 (1913)) who recognises that homogeneous mass behaviour (Massenhandeln) is not a case of social action at all because every single actor only acts for and by themselves. These actions merely appear socially organized from the outside.

The precise nature of shared or joint intentions is contested. Bratman (1993, 2014) takes them to be a combination of intentions with a content referring to the group as ‘we’ plus common knowledge conditions. Searle (1995, 2010) and Tuomela (2004, 2005, 2013) propose different forms of a ‘we-mode’ for propositional attitudes. The furthest step away from individualism are suggestions by Gilbert (1989, 1990, 2009), Pettit (2001, 2007), List (2005), List and Pettit (2011) and Schmid (2009) to take the idea of a plural subject at face value without thereby admitting to any independently existing group mind.² For my purpose these details are not important. All I need is a rough and ready distinction between groups and mere collections of individuals. I am thereby not saying that these differences will never be important for questions about collective self-knowledge. However, my goal is less ambitious. I merely want to present reasons for taking questions of collective self-knowledge seriously. I do not provide a fully developed account for collective self-knowledge. Furthermore, I do not argue for the existence of mental states of groups. In order to focus on the topic of collective self-knowledge I take for granted that groups can have mental states. Arguing for the existence of these states would have to be either too quick to be convincing, or take too much space away from discussing collective self-knowledge. As a result, what I offer can be seen as a conditional claim: if social

¹ Huebner (2014) has a slightly different approach to understanding the mentality of groups that does not start from joint intentions. He focuses on parallels between the cognitive architecture of individuals and the cognitive architecture of distributed system. His general idea is that we can explain the mentality of groups by looking at cognitive processes without starting from the more controversial assumption of collective intentional states. These states might be a result of the group’s cognitive processes, but need not be taken for granted.

² Schmid (2017) holds that the distinction between we-mode and plural subject is not sustainable. Rather, the we-mode already is the subject, insofar as the mode itself involves some sort of self-awareness (cf. Schmitz 2018).
groups can have mental states, then we should be looking into collective self-knowledge and its features. Given that this antecedent claim is a live option in the current debate, I take an investigation of collective self-knowledge to be worthwhile.

2 Self-knowledge

Generally, I am in the best possible position to tell what mental state I am in. If I sincerely say ‘I feel pain in my leg,’ or ‘I want some ice cream,’ then rarely anyone ever doubts me. Moreover, there are some differences in our linguistic practice between statements such as ‘I feel pain in my leg and there is a tree behind the corner.’ It seems perfectly fine to ask ‘How do you know?’ in response to the tree statement, but it is comparatively odd to ask the same in response to my claim about feeling pain. It seems inappropriate to question the mental state self-ascription because the natural answer to the question is a reiteration of the initial claim. I know that I am feeling pain, because I do feel pain. Examples that emphasize the peculiar speech act of self-ascribing a mental state are widespread and commonly used to introduce the idea that self-knowledge appears to be special and privileged. Call these speech acts avowals.

It is still under debate whether these notions of self-knowledge being special and privileged are supposed to be explained on a doxastic or linguistic level. Wright (1998, 2015) and Finkelstein (2003) suggest that theories of self-knowledge should first and foremost explain the linguistic practice of avowing, whereas the majority of accounts in the literature aim to explain apparently peculiar and privileged beliefs and belief-forming processes (e.g. Moran 2001; Nichols and Stich 2003; Goldman 2006; Carruthers 2011; Cassam 2014; Byrne 2018). Following arguments from Borgoni (2019) and Schwengerer (Forthcoming) I assume for this paper that the phenomenon of self-knowledge ought to be understood primarily in terms of belief and belief-formation. I label this process of forming beliefs about one’s mental states introspection. When beliefs produced by introspection are said to be privileged and peculiar I take that to be a claim about the reliability and nature of introspection. When beliefs are said to be privileged then one’s beliefs about one’s own mental states are supposedly more likely to amount to knowledge than one’s beliefs about other’s mental states or the external world. Moreover, the claim of peculiarity of self-knowledge is a claim about a special method or way of knowing that produces self-knowledge (Byrne 2005). I want to highlight that one can hold that ordinary, adult human beings know their own mental states in an apparently privileged and peculiar way, without committing to the claim that this requires actually privileged and peculiar beliefs about one’s mental states. Hence, Gopnik (1993), Carruthers (2011) and Cassam (2014) all argue that self-knowledge is not as privileged and peculiar as usually assumed. Both Gopnik (1993) and Carruthers (2011) support

---

3 Which often is but need not be understood as a speech act.

4 ‘Apparently’ is meant to highlight that some of these views propose that self-beliefs are less privileged and special than we ordinarily assume. See Schwitzgebel (2008), Carruthers (2011) and Cassam (2014).
their position by drawing on psychological studies (e.g. false belief tasks and confabulation studies). Their proposals show that we have to distinguish the question of whether one has self-knowledge from the question of whether one has privileged and peculiar self-knowledge. You can affirm the former without affirming the latter.

Even though I locate privilege and peculiarity on the level of belief, I accept that our ordinary linguistic practice provides a guide to features of self-beliefs. Hence, I am not completely abandoning Wright’s (1998, 2015) observation that avowals are authoritative, immediate, and salient. Wright distinguishes between phenomenal and attitudinal avowals, but I am here only interested in the latter. These are avowals of content-bearing states, such as ‘I believe that term ends on the 27th,’ or ‘I hope that noise stops soon’ (Wright 1998, p. 15). Wright argues that attitudinal avowals are weakly authoritative. They provide empirically assumptionless justification for the corresponding third-person claims. Even though on occasion one may doubt individual attitudinal avowals without doubting sincerity or understanding, such doubts are rare. Moreover, one cannot doubt that a person correctly avows attitudes in general. A person cannot be always unreliable about their own mental states. Attitudinal avowals are also taken to be immediate, which captures that it seems inappropriate to ask for reasons for the avowal. If I say ‘I want to eat a biscuit’ it seems infelicitous to ask ‘How do you know?’ Finally, attitudinal avowals are salient. If one is asked ‘Do you believe that p?’ it seems absurd to answer ‘I don’t know.’ It is clearly not absurd to answer that you suspend judgment, but it would be absurd if you did not know whether you believe that p, not believe that p, or are withholding judgement. These are general features of avowals, but they all allow for exceptions. Wright (2015) emphasizes that they depend at least on sincerity, and cognitive competences. Moreover, Borgoni (2019) shows that authority can be missing in contexts in which the avower is a victim of epistemic injustice.

I am not committing to any specific account of self-knowledge. All I take for granted is that self-beliefs appear to be privileged and peculiar in Byrne’s sense, and that this appearance tends to manifest in avowals as speech acts with unique characteristics in our ordinary linguistic practice. I do not assume that they are in fact privileged and peculiar. With these assumptions I can turn to the case of collective self-knowledge.

5 Wright (1998, 2001) uses groundlessness instead of immediacy, and transparency instead of salience.
6 A reviewer points to cases in which I have strong reasons in favour of believing that p, but also strong reasons against. These are cases in which I have not yet made up my mind with regard to p. It is at least not obvious whether these are cases of simply withholding judgment. Nevertheless, these are also cases in which it does not seem appropriate to claim that I do not know at all in which mental state I am. I am able to tell you about my doxastic situation regarding p—namely that I have not made up my mind yet given my reasons.
7 This emphasis is supposed to answer criticism by Snowdon (2012).
Towards Collective Self-knowledge

As stated in the very beginning, ordinary parlance allows for ascribing intentional states to groups and treating them as agents. Moreover, I explicitly started with the assumption of group states and group agents. Plausibly, with these assumptions in place groups as agents can also testify (Tollefsen 2007). From here it is only a short step to groups self-ascribing intentional states. If groups can testify that other groups have certain beliefs or intentions, then they should also be able to testify about their own beliefs or intentions. We find plenty of evidence for the antecedent. The investigative commission appointed to look into the Grenfell Tower fire might testify in their report that the construction company knew of the fire hazards in the building. The national football team of Croatia might state in a press conference that England’s team intends to play long balls into the box. We are familiar with groups ascribing mental states to other groups. We should therefore also accept groups ascribing intentional states to themselves.

One might interject here: Why should we care about collective self-ascriptions? I take it that we actually do care about these already. We judge people’s responsibility for their actions based on their epistemic access to relevant information about that action. We think that it makes a difference whether someone knowingly engages in behaviour with a high risk of harming others, or whether they were not aware of any such risk. Of course, the results of their actions matter, but so do the beliefs they had before and while acting. In law this makes the difference between mere negligence, and wilful misconduct. Generally, we think that wilful misconduct is in some sense worse than mere negligence. Take a variations of Peacocke’s (1998) administrator case.

ADMIN: A university administrator is in charge of admissions. She says that she believes that every applicant deserves the same chances regardless of their nationality and should be treated equally. However, in practice she prefers her compatriots.

Suppose we only have the avowal and the behaviour of the administrator available to make a judgment. The case stipulates that the admin acts in a biased way and is blameworthy with regard to her actions. However, is she also blameworthy with regard to her self-ascription of being egalitarian? Should we take her avowal that she believes that every applicant deserves the same chances at face value? Here we have a disagreement about the admin’s mental states. The admin claims to be in a mental state, but our evidence indicates that she is in a different state. In this case, she claims to have a particular belief B, but the evidence suggests that she has a different belief B* that is actually incompatible with B. Our task is to evaluate the situation given the authority of her avowal and the privilege of self-beliefs. If we take her to believe that compatriots are always preferable, then her actions become expressions of racist or nationalist beliefs. If we take her to believe what she avows, then her actions will be merely mistakes insofar as they do not contribute to satisfying the intention of treating all applicants equally. In the former case the admin is double blameworthy, compared to the latter: both
action and belief are blameworthy. Hence, determining what she likely believes becomes important for attributing blame. We have to ask how much evidence of biased behaviour we need to overrule the avowal. I take it to be more than in an ordinary disagreement that does not involve avowals at all. After all, at face value she seems to have privileged and peculiar access to her own states, so we need extraordinary (in strength or amount) evidence to overcome that privilege. Nevertheless, the privilege is not insurmountable. People are not infallible about their own mental states. We are prone to use quick and easy heuristics instead of answering difficult questions (Kahneman 2011), we sometimes come up with confabulated reasons (Nisbett and Wilson 1977), and we have a strong bias to see ourselves according to our self-conception (Schwitzgebel 2012). Her avowal is no guarantee of her belief. Hence, when confronted with an avowal we need to consider the nature of her access to her own beliefs and factors that can obstruct the access in the current situation. I am not providing an answer to this question here. Instead, I argue that the very same situation occurs in relation to social groups.

For instance, take the court case against BP regarding the Deepwater Horizon oil spill. One of the crucial points of the case was the question whether BP was merely negligent, or whether it was a case of wilful misconduct. That is, the question was precisely whether BP acted in a way that ignored relevant knowledge of the risks involved. One strategy to argue against the wilful misconduct case would have been to claim ignorance of the risks. In this case the situation is similar to the admin. BP might testify that they did not believe that there were any unusually high risks. How we should handle disagreement with BP’s testimony depends on whether we take BP to have any sort of privileged access to their intentional states.

I want to flag that this court case is about a past event and the mental state ascription is about a state BP had in the past. This makes a difference, because the privileged nature of self-beliefs might be diminished when memory is involved. I take it that we often still treat a self-ascription of a recent (but not current) mental state as privileged. I am likely the best person to ask about what I believed in the recent past. However, memory introduces new ways in which one might be wrong about one’s own past mental states, so ideally I find a case without the noise of memory. I think the same kind of case can be made about current mental states held by a group. Consider the following case:

ROXXON: Roxxon is an oil company that drills at an area on the ocean floor close to Iceland. The area is especially unsafe to drill, but this is not common knowledge. Greenpeace files a complaint hoping that the Icelandic government stops the drilling due to the high risk. They also file a charge for wilful misconduct against Roxxon, claiming that Roxxon knows about these risks but keeps drilling anyway. Roxxon answers in a press release by sincerely stating that they believe that there is no unusually high risk involved and it is safe to drill.

The Icelandic court in this case is confronted with two claims. First, that the area is too risky to drill; and second, that Roxxon knows about the risks. The first part can be answered by looking at evidence about the area and the risks. However, the second claim involves a disagreement about Roxxon’s intentional states. Even if the
area is shown to be risky, what exactly is required to show that Roxxon knows about these risks, if Roxxon itself testifies that it believes the area is safe? The difference here is between merely stopping the drilling for oil or holding Roxxon accountable for willingly endangering the ecosystem and anyone living close to the drilling platform. And to figure out how to handle such a case we need to know whether Roxxon has privileged access to their own intentional states. We need to know to what extent Roxxon is the best subject to ask about Roxxon’s mental states in order to know what evidence is required to overrule Roxxon’s testimony. In other words: we need an account of collective self-knowledge.8

However, that might be a little too quick. Here are two worries with these court and complaint cases as intuition pumps: The first potential problem is that these cases introduce an incentive for a guilty speaker to be insincere. It seems obvious for BP to lie, even if they knew that they were aware of the risks. To counteract this incentive you can imagine cases in which amnesty will be given to anyone who broke the law in any way. For instance, the Truth and Reconciliation Commission in South Africa was able to give amnesty by default for a subset of harms done.9 These cases will be less common and therefore our intuitions might not be as clear as we would like them to be, but help to remove some of the noise of the incentive to lie.

Similar incentives do not only occur in court cases. Corporations have general incentives to present themselves in a good light, both to the outside and to their own members. However, that does not seem to be all that different from individuals wanting to see themselves according to their self-conceptions. Hence, just as we do in evaluating avowals by individuals we need to consider the nature of the group’s access to the group’s beliefs, factors that can obstruct the access in the current situation, and incentives to consciously lie.

The second worry is that what matters most in these cases is not whether BP or Roxxon knew, but whether they were in a position to know, or should have known.10 That might be right. It certainly matters to us whether they could have known and should have known. However, that does not rule out that it also matters to us whether they did know. It is possible to blame Roxxon for (1) drilling in a risky area, (2) drilling in an area Roxxon should know that is risky to drill, and (3) actually knowing that the area is risky to drill. In order to blame for (3) we need to overrule Roxxon’s claim that Roxxon does not believe that the area is risky to drill. We need to overrule Roxxon’s self-ascription of a mental state. In case we cannot do that Roxxon can only be blamed for (1) and (2). These three ways of blaming Roxxon

8 Relatedly, we also need an account that tells us something about the knowledge that a member of the group has about the group’s states. For instance, when complaints about the sexist and dangerous work culture in the Dallas Mavericks were made public in 2018 by Sports Illustrated the Maverick’s owner Mark Cuban claimed that he did not know about these practices. Supposing that these practices indicate some mental states of the group we should ask whether that is a credible avowal of Cuban. There is an intuition that a hands-on owner such as Cuban has to know something about what is going at the workplace and what general group beliefs regulate everyday work in the Mavericks. I plan to look at such cases in detail in other work.

9 Thank you to Raphael van Riel for this example.

10 Thank you to an anonymous reviewer for raising this issue.
in the case might not be exhaustive. However, other aspects of the case that provide reasons to blame Roxxon will not remove the importance of Roxxon knowing or not knowing. We just add another reason to blame Roxxon for. Hence, we still need to know what to do with Roxxon’s self-ascribing the belief that there is no unusually high risk involved. We still need an account of collective self-knowledge.

After these somewhat muddy cases about blame and practical considerations I want to consider a more theory driven motivation for looking into collective self-knowledge. Even if you are not convinced by the previous cases, there are theoretical reasons to be interested in collective self-knowledge. Self-knowledge seems to be closely connected to agency under some understandings of agency. For instance, List and Pettit (2011) notice that a newspaper such as The Economist can self-ascribe their intentions and goals. They proceed to tell us that such self-ascriptions are clear markers for group agency (List and Pettit 2011, p. 41), but unfortunately do not develop their view on the role of self-knowledge further.

The relation of self-knowledge to agency becomes more evident in a take on intentional action championed by Anscombe (1957). Anscombe proposed that we should understand intentional φ-ing as involving non-observational knowledge of φ-ing and knowing non-observationally that such-and-such is the cause for one’s φ-ing. If I intentionally throw a ball, I need not observe myself to know what I am doing and why I am doing it. I just know. One natural way to understand non-observational knowledge here is to take it to be a kind of self-knowledge. After all, it seems to be knowledge that is special and privileged. No one else knows that I am intentionally φ-ing as well and by the same method as I do. Usually, I am the best person to be asked about what I am doing, or why I am doing it. If we take this observation at face value it seems that this view leads towards a dependence of intentional action on a kind of self-knowledge. Hence, if we want to say that groups can act intentionally, we better be able to say something about the kind of self-knowledge involved in intentional actions for groups. This is a challenge that so far only Schmid (2014, 2016) and Rödl (2018) engaged with explicitly.

At this point I have shown that groups can self-ascribe mental states and that we should be interested in these self-ascriptions. Or, put on a doxastic level, that if groups have knowledge of any kind, they also have self-knowledge. However, I have not shown that collective self-knowledge has the distinctive epistemic features that individual self-knowledge is supposed to have. In particular, I have not shown that groups have apparently privileged and peculiar access to their intentional states, or that speech acts self-ascribing mental states for groups show authority, salience, and immediacy. However, I think we have clear cases of speech acts for groups that fit with Wright’s characterisation of avowals. A paradigmatic case I have in mind is a group’s mission statement. Consider The Vienna Forum for Analytic Philosophy (hereafter Forum), a small philosophical society at the University of Vienna. The members of Forum might come together and publish together a mission statement like the following:

MISSION STATEMENT: The Vienna Forum for Analytic Philosophy aims to support professional work in analytic philosophy. It is based at the University of Vienna and was founded in 2009. Our principle objective is to intensify
Towards Collective Self-knowledge

research, exchange and discussion among students and scholars. For this purpose, we organize regular meetings, lectures, workshops and conferences and foster national and international contact with individuals and groups interested in analytic philosophy. The Vienna Forum for Analytic Philosophy provides a platform for work committed to an analytical approach in areas including philosophy of language and logic, philosophy of mathematics, epistemology and philosophy of mind.11

Along some historical references to the founding, the mission statement asserts what the society intends to do, what they aim for and what their core beliefs are.12 I suggest that this mission statement is similar to an individual expressing their mental states in a sincere speech act. For the individual case it seemed inappropriate to challenge their avowal. What about Forum’s self-ascription? Here the epistemic position of individual group members has to be kept sharply distinct from the epistemic position of the group as a whole. The parallel I want to draw is one between individuals avowing their intentional states and groups avowing their intentional states. An observation about individual group members avowing the intentional states of the groups they are a part of is not going to impact the parallel at all. Hence, we need to be careful to only look at differences concerning the group avowal13 as a speech act done by a group. Schmid (2014, 2016, 2018) notices that group ascriptions done by individual members usually do not seem to be authoritative, immediate and salient avowals.14 However, he is more sympathetic to a sort of privileged position with regard to collective states for the group itself. I think this fits our intuitions on the self-ascriptions in the Forum example. As long as we take Forum as a group to be sincere and linguistically competent it appears to be inappropriate to challenge the mission statement with regard to the intentions expressed in it. In ordinary circumstances a challenge of the Forum’s claim about their aims will only come in form of challenging the sincerity of the claim made. On this axis the mission statement appears to be an authoritative group avowal.

However, is it also inappropriate to ask Forum how the group know that this is what it intends and aims for? Is the group avowal immediate? For individual avowals an answer to ‘How do you know what you intend?’ would merely be a reiteration of the intentions. I just know what I intend. How else am I supposed to answer than by stating my intention? For collective avowals an appropriate answer might be different. If Forum is asked how the group knows what it intends there seems to be

11 Taken from wfap.philo.at (Accessed 2/11/2018).
12 A different reading might take mission statements and similar collective avowals to be expressions of commitments and not assertions. In this case the same strategy I propose here can be put forward, but with different background assumptions on self-knowledge. For instance, one might argue that collective avowals are similar to individual avowals insofar both of them express commitments. This requires an acceptance of a broadly agentalist account of individual self-knowledge. However, views in the spirit of Moran (2001), Bilgrami (2006) or Coliva (2016) could still make a structurally similar argument for collective self-knowledge.
13 I use group avowal and collective avowal interchangeably.
14 Schmid (2014, 2016) does not use Wright’s terminology, but discusses features that are very similar. These differences do not matter for my purpose.
a plausible answer: ‘we discussed and deliberated on our goals’ or perhaps ‘we discussed our intentions and decided by vote.’ Here again, I emphasise the difference between asking a member of the group and asking the group. Of course, it is appropriate to ask an individual member how the member knows what the group intends (cf. Schmid 2014, 2016). But the situation we imagine is not one in which a member is asked. This situation is irrelevant for the features of collective avowals. What I am interested in is whether it would be an appropriate response to the collective speech act. Would it be appropriate to ask the group ‘How do you know what you intend?’ Here the intuitions are less clear. There is a question in the vicinity, one asking how the group came to an intention. That question is definitely felicitous. One can ask the group how they settled on a particular intention, and the group can answer by pointing to discussions, voting, or other forms of generating intentions as a group. Is this enough for the group to also be able to state how the group knows what it intends? I am inclined to say ‘yes’. It is an appropriate question with a satisfying answer, even when directed to the group. If this is right, then this points to an interesting connection between the formation of an intentional state in a group, and the self-ascription of the intentional state. The evidence a group gives for self-ascribing a state is, at least in this case, not based on merely observing the group’s behaviour. The evidence is the group’s own intention-forming process. I will come back to this at the very end. At this point, I have illustrated a case of a group avowal that is authoritative, but does not seem immediate.

Let me look at a second example. Consider a situation before an election. Political parties work on their programs or manifestos that express their political goals, intentions, and values. The labour party might publish their manifesto which states that the party wants to increase the funding for the National Health Service. Would it seem appropriate to ask the labour party “Do you really want that?” I do not think so. This is not because there is no individual to ask. We could—and journalists and opposing parties often do—direct questions towards the labour party as a group. Rather, I suggest, it is because the manifesto functions as a collective avowal and avowals are authoritative. But here the analogy with individual avowal seems to break down. When I avow “I want ice cream” it seems unlikely that someone challenges me. However, when the labour party manifesto states that the party wants to increase funds, a challenge seems quite ordinary. The manifesto is not taken at face value! To explain what is going on here it is worth looking back to Wright’s formulation of authority. Wright is not claiming that avowals cannot be challenged. In fact, he explicitly allows that individual avowals can be challenged on some dimensions. However, they cannot—or only in rare cases—be challenged with regard to brute errors. When introducing avowals as authoritative he claims that “A doubt about such a claim has to be a doubt about the sincerity or the understanding of the one making it” (Wright 1998, p. 14). And even though the authority of attitudinal

---

15 Another example of this happening even in a formal environment are questions to the commission of the European Union. When members of the European Parliament submit such questions they are always addressed at the commission as a whole, not individual commission member.

16 Thank you to an anonymous reviewer who has pressed this point.
avowals allows for exceptions, he still holds that “You may not suppose me sincere and comprehending, yet chronically unreliable, about what I hope, believe, fear, and intend” (Wright 1998, p. 17). In these formulations sincerity is always noted as a condition for an avowal to be authoritative. Or in other words: it is rarely appropriate to challenge a sincere avowal, but it is perfectly fine to challenge the sincerity of an avowal. This is what often happens when one is confronted with a political manifesto. It is not unusual, nor inappropriate to challenge the sincerity of a political manifesto. But once you grant sincerity challenging the manifesto as being simply mistaken about the party’s intentional states seems to be infelicitous. Assuming sincerity, the labour party is the best subject to tell us about the labour party’s intentions. If their manifesto states that they intend to φ, we believe that they intend to φ, as long as we do not take the party to be insincere or have unusually strong evidence against the party intending to φ. Again, the avowal seems to qualify as authoritative.

So far collective avowals seem to be authoritative, but not immediate. Are they salient? This feature of individual avowals does not seem to have a straightforward analogue in the group case. Wright claims

[...] that, except where the matter is one of interpretation, we think a subject ought to know without further ado what she believes, or desires, etc., so that any profession of ignorance or uncertainty, unless coupled with a readiness to allow that the matter is not basic but calls for (self-) interpretation, will seem perplexing (Wright 1998, pp. 16–17).

This claim about salience does not seem to hold on the group level. It is perfectly plausible that a group like Forum is asked about their intentions and the group cannot “without further ado” know what the group intends. In some cases Forum cannot immediately testify if they believe p, non-p, or suspend judgment. Paradigmatic candidates for these instances are cases in which the proposition has not been the topic of any discussions in the group. The reason for this lies in the group’s structure. The states of the group depend on interactions within the group. The mission statement was formed after deliberation and discussion within Forum. But asking the group about a proposition p that has not been the topic of discussions within the group is different. The group has not made up its mind yet with regard to p. In this case, Forum does not know whether it believes that p, because the question whether p is not settled for the group. Moreover, the group might be asked about its states when it is not in a position to settle the question whatsoever. Suppose an e-mail is sent to Forum asking whether the group believes that p, but almost all members of the group are on vacation and merely two members are present. These two on their own are unable to settle the question in any way. They cannot even legitimately state

---

17 Though “without further ado” might already be mistaken for some cases of individual self-knowledge. See footnote 6.

18 This relates to discussions about self-knowledge and the deliberative stance present in Moran (2001) and settling questions as discussed in Vierkant (2015). Collective self-knowledge seems to be most plausible in case the group has settled a question already.

19 When exactly that is the case varies. See also Gilbert (1989, p. 311) on the role of conventions in similar cases.
that the group suspends judgement, because they cannot speak for the group at this point. The question reached members of the group, but not the group itself, which is why we cannot expect an answer. Such a divergence is not possible in case of the individual. When the individual receives the question they will (assuming sufficient cognitive and linguistic resources) be in a position to settle the question, even if only by suspending judgment. However, for groups there might still be a weaker form of salience available. If the group receives the question they will be in a position to answer whether they believe p, non-p, suspend judgement, or are in the process of judging. It would seem odd for a fully operative group to answer that they do not know which of these fits the group’s current situation. Notice that this does not entail that a group is aware of all its mental states. The claim is only about the felicity of a certain sort of response to a question that reached the group. In this condition it appears to be absurd for the group to answer “We do not know.” One might use the term ‘suspending judgment’ also for cases in which the group is still in the process of judging. I do not think these terminological questions matter too much. The benefit of keeping them apart for group cases is that forming mental states as a group usually seems to be slower than for individuals, and hence it might be useful to pick out the state of being in the process of settling for a state more prominently. 

So far I argued that there is a similarity between individual avowals and collective avowals. They share the feature of authority, and both seem to have some form of salience. Even though collective avowals do not seem to be immediate, they are different enough from ordinary assertions that they deserve the name avowal and should be treated accordingly. If one takes the individual avowal to be an indication of self-beliefs, then one should also take the collective avowal as an indication of collective self-beliefs. In the good case it will constitute collective self-knowledge. However, even if social groups can make speech acts that are relevantly similar to avowals, one might be reluctant to take the next step and attribute privileged and peculiar second-order beliefs. I want to mention two lines of response here. First, features on the level of intentional states can explain why these avowals have the features they have. Hence, the group having a privileged belief about its intentional states seems to be a good candidate for an inference to the best explanation. Second, we can find some arguments for privileged belief beyond authoritative speech acts for individual self-knowledge (Borgoni 2019; Schwengerer (Forthcoming)). We can then extend the parallels between individual and collective self-knowledge by attending to these arguments. For instance, Borgoni (2019) argues for the existence of self-knowledge without the possibility of authoritative avowals for the individual case. The idea is that authority can be cancelled out in case the speaker is a victim of epistemic injustice. In a sexist society a woman’s avowal might not be taken seriously. Nevertheless, in these situations we would intuitively attribute self-knowledge. We can conceive of similar situations in which a social group is not in a position to testify their mental states with authority, but the group plausibly still has privileged self-knowledge. Imagine a family of refugees who fled together from a war-torn area, trying to reach a safe place together. The family testifies that the group believes to be in danger. Suppose further, that they

---

20 This can be construed as a case of silencing. For a discussion of silencing see Dotson (2011).
fled into a xenophobic society in which people do not take their testimony seriously. In this society people expect refugees to exaggerate their fears and therefore the group is a victim of testimonial injustice. The family might write together a statement to a media outlet, asserting their belief to be in danger. The statement is published, but no one believes it. Here the group’s testimony is not authoritative. Nevertheless, it seems plausible that the group of refugees knows that the group believes to be in danger. After all, the best explanation for their action of writing the statement together is that they believe as a group that they are in danger. The family as a group knows that they believe to be in danger, but the group avowal does not show authority.

A final note on the similarities and differences of individual and collective avowals. Schmid (2014, 2016) observes that collective avowals do not always come with a use of the terms ‘we’, ‘our’, ‘us’, etc. Moreover, even in case they do, ‘we’ seems importantly different than the use of ‘I’ in individual avowals. ‘I’ is taken to be immune from error of misidentification. It’s impossible to identify someone else as the one I am talking about when I avow ‘I believe that p.’ On the other hand, it is not fully clear whether ‘we’ is immune to error of misidentification when used in a group avowal. Clearly, often ‘we’ is not immune to such an error when used by an individual. If I say ‘We need to leave’ the term ‘we’ is picking out a set of people and I might make an error in determining whether a particular person is part of the set or not. However, given that the relevant speech act is the group avowal the observation from the individual case will not get us anywhere. And for group avowals ‘we’ still provides problems, because ‘we’ can be used to refer to social groups, but also to refer to a selection of people that do not satisfy the conditions for constituting a social group at all. In the latter case, it is just a shortcut for referring to a set of people, and can surely misidentify. I am not sure whether ‘we’ used by a social group in a collective avowal is in the business of picking out the group. I suspect that it is at least not merely picking out the members of the group. Recently, Schmid himself proposed that there is a sort of self-identification that might be close enough to the singular case (Schmid 2018, pp. 238–239). Moreover, Smith (2018) is convinced that some uses of ‘we’ are immune to error of misidentification just like most uses of ‘I’ are. Ultimately, I do not think it matters. Collective avowals can be authoritative even when the subject of the avowal is identifying the speaker and is not immune to error of misidentification. After all, immunity to error of misidentification alone seems neither necessary nor sufficient to explain authority or privileged access. The literature on self-knowledge is full of proposals that require something in addition to this immunity [e.g. a second kind of immunity in Bar-On (2004)] or that function without any special role of this immunity to explain privileged access (e.g. Nichols and Stich 2003; Goldman 2006; Carruthers 2011; Cassam 2014; Byrne 2018). If this is correct, then it should not worry us much that we find collective avowals using ‘we’, ‘us’, ‘our’ intermixed with references by name.21

21 However, we should still be sceptical of taking collective self-ascriptions without any indexical referring to the speaking group as collective avowals. After all, we accept avowals as (partial) explanations of action, and explanations of actions seem to require some indexicality in the intentional states (cf. Perry 1979).
4 Collective Self-knowledge and Reducibility

I have argued that collective self-knowledge exists and can in paradigmatic cases be expressed in an authoritative fashion. This indicates that there is at least some sort of privileged access present for social groups. The next natural question to ask is how collective self-knowledge relates to knowledge of the group members. Is the group’s privileged access located on the group level, or is it always instantiated in an individual? And can groups know their own intentional states if no single member knows the group state? These questions show similarities to the discussions around summativism. Summativists claim that a group has a mental state \( M \) only if some of its member has the state \( M \). Non-summativists deny this, and propose that group states are independent of matching individualistic mental states (Gilbert 1989).22

We can think of a proposal similar in spirit that proposes that a group only has a second-order belief (knowledge) about the group’s intentional state in case that an individual member has a belief (knowledge) about the group’s intentional state. This is a second-order summativist view that could still be compatible with non-summativism for first-order states. Call this second-order summativism.

Second-order summativism A group \( G \) knows (believes) that \( G \) has mental state \( M \) only if some of \( G \)’s members know (believe) that \( G \) has mental state \( M \).

My aim now is to show that second-order summativism is false or at least not suitable to explain the features of collective self-knowledge. I do so by considering an argument by Bird (2010, 2014) for social knowledge that can be adapted as a counterexample to second-order summativism. Bird’s (2010, 2014) original case is the following:

**SCIENCE:** Dr X, a physicist, and Dr Y, a mathematician, collaborate on a project to demonstrate the truth of the conjecture that \( q \). Their project can be broken down into three parts. Part one is a problem in physics, the problem of showing that \( p \), which will be the work of Dr X alone. Part two is a problem in pure mathematics, that of proving that if \( p \) then \( q \), for which Dr Y takes sole responsibility. Part three is an application of modus ponens to the results of parts one and two. They arrange for an assistant to publish the paper if and only if the assistant receives from Dr X the demonstration that \( p \) is true and from Dr Y the proof of \( p \rightarrow q \) (the brief final part with the application of modus ponens has been prewritten). We can imagine that Drs X and Y have no other communication with each other or with the assistant and so do not know at the time of publication that \( q \) has been proven.

In this case the group comes to know that \( q \), even though no individual member of the group knows that \( q \). X and Y are not aware that the other researcher succeeded in their part, and the assistant does not understand the individual parts sufficiently to know. Nevertheless, the group acts on their belief that \( q \) insofar as the group publishes the paper. Moreover, they know that \( q \), because the case is set up such that

---

22 Lackey (2016) falls somewhere in-between.
q is true and the group belief justified. Hence Bird suggests that “[…] we cannot reduce states of social knowing to states of individuals” (2014, p. 58). Kallestrup (2016) argues in support of Bird’s intuition that the belief in SCIENCE is even produced competently and therefore satisfies requirements of a virtue reliabilist account of knowledge.

SCIENCE is a counterexample to summativism and can be used to target second-order summativism explicitly. The adaption required is merely a fourth part with an addition of a self-ascription of intentions for future research that is conditional on finding out that q.

REFLECTIVE SCIENCE: Dr X, a physicist, and Dr Y, a mathematician, collaborate on a project to demonstrate the truth of the conjecture that q. Their project can be broken down into four parts. Part one is a problem in physics, the problem of showing that p, which will be the work of Dr X alone. Part two is a problem in pure mathematics, that of proving that if p then q, for which Dr Y takes sole responsibility. Part three is an application of modus ponens to the results of parts one and two. Part four is a plan for future research on r that the research group is intending to pursue if q is proven true. This plan has been developed by X and Y together and is prewritten before they have any knowledge of whether q is true, nor of the success of the first two parts. They arrange for an assistant to publish the paper if and only if the assistant receives from Dr X the demonstration that p is true and from Dr Y the proof of p→q (the brief part with the application of modus ponens and the part about future research intentions have been prewritten). The paper includes the self-ascription of intentions for future research on r. We can imagine that Drs X and Y have no other communication with each other or with the assistant and so do not know at the time of publication that q has been proven, nor that they intend to do future research on whether r.

In REFLECTIVE SCIENCE almost everything stays the same. Drs X and Y still do not believe that q individually. Neither does the assistant who lacks the concepts required to understand q properly. Moreover, no member of the group has a privileged belief that the group intends to research whether r. X and Y lack this belief, because they have no idea whether q is true. The assistant cannot believe that the group intends to research whether r, because the assistant lacks understanding and concepts for q and r. At best the assistant can come to believe that the group intends to research something described as r, which is a related but different proposition. Given that no member of the group believes that the group intends to research whether r, is this even a case of collective self-knowledge? I think so. Here the parallel with SCIENCE is important. If you take SCIENCE to show that the group knows something that no member knows, than you should also accept that in REFLECTIVE SCIENCE the group has the self-belief that it intends to research r. After all, the publication states that the research group intends to research r. This claim seems to be authoritative. It would be inappropriate to challenge the group on this point, besides questioning the sincerity. Hence, REFLECTIVE SCIENCE appears to be a case of collective self-knowledge with no individual members having knowledge of the group state.
One might try to put some pressure against this line of argument in two ways. Here is the first one: I conceded that the assistant might have the belief that the group has the intentions that are described as researching r. I take it that there is a gap between this description and a belief with the content that the group intends to research r. However, for the sake of argument suppose that this is actually sufficient to have the belief that the group has the intention to research r. This would satisfy the condition of the second-order summativist. Nevertheless, there is still something fishy about this reading of the case. The self-ascription on the group level in the publication would be authoritative. However, the assistant’s belief lacks privilege and cannot be the basis of the authoritative group avowal. Their belief is merely based on the evidence that there is a published paper stating that the group intends to research r. It is a belief formed by observing the group’s behaviour: an act of mind-reading, not of introspection. It seems perfectly appropriate to challenge the assistant on the truth of the claim. The reason for this is that the assistant does not play a role in decision making and the formation of intentional states of the research group. The assistant is not an operative member of decision making for the research group (Tuomela 2013, pp. 86–87). That is, the assistant is merely responsible for putting group decisions into action. Drs X and Y on the other hand are in this case operative members of decision making for the group. Their actions determine what the group does, believes and intends. Hence, even if the assistant has the second-order belief that the group intends to research r, we cannot explain the features of the second-order group state in virtue of the assistant’s second-order belief. The authority of the group’s testimony cannot be explained in virtue of the assistant’s second-order belief, because the assistant has no authority. Moreover, if we continue the story and include the assistant’s death shortly after the publication, the remaining group would still have a collective avowal with all its authority indicating self-knowledge. Hence, collective self-knowledge seems to be present without any individual members having privileged access to the group’s states. Second-order summativism is unable to explain collective self-knowledge.

The second line of resistance against REFLECTIVE SCIENCE is to accept the existence of the group avowal, but reject that any collective intentional state should go with that avowal. That is, even though the group is able to avow that they intend to research r, the group does not believe that they intend to research r. This move allows one to hold on to second order summativism while still accepting both Bird’s case and my adjusted version. The drawback of this solution is that we thereby lose our explanation of why the avowal is authoritative. A convenient way to explain authority is to propose that the avowal is based on an especially reliably formed belief. When I say I believe that p, you accept my assertion by default because I say this in normal circumstances only if I have a reliably formed belief that I believe that p. Reliability is by no means the only way to explain authority, but it is one of the most popular explanations found for instance in Armstrong (1968), Nichols and Stich (2003), Goldman (2006), Carruthers (2011), Cassam (2014), and Byrne (2018). Giving up an attractive explanation of authority is a significant price to pay.
Towards Collective Self-knowledge

5 Lessons for a Theory of Collective Self-knowledge

I end by providing some suggestions for how to build a theory of collective self-knowledge. Any such theory ought to explain why group avowals are authoritative and weakly salient, where salience is understood in the limited sense that when a group is successfully asked whether it has an intentional state, it is infelicitous for the group to answer “We do not know”. Moreover, at least in some cases the group avowal is not immediate. It can be appropriate to ask the group “How do you know that you intend to φ?” The group can answer by citing its process of forming the intention in question. As I pointed out in the Forum case, the role of the process forming the intention strikes me as especially interesting and potentially insightful. The group does not know what it intends by merely observing any connected group behaviour. It would be odd if the answer to “How do you know that you intend to φ?” were “We observed our actions in the past year and inferred that their best explanation is an intention to φ.” Felicitous answers look more akin to “We discussed the issue and put it up to a vote” or “We talked internally about our core values and agreed that those demand us to φ.” It is specific behaviour inside the group that constitutes the grounds for the group’s self-ascription of mental states. Behaviour that generates the very mental state that is self-ascribed.

Pointing to a procedure that generates a mental state in order to self-ascribe that state has a tradition in theories of self-knowledge. The thought is prominently linked to an idea of Gareth Evans:

[I]n making a self-ascription of a belief, one’s eyes are, so to speak, or occasionally literally, directed outward—upon the world. If someone asks me ‘Do you think there is going to be a third world war?’, I must attend, in answering him, to precisely the same the same outward phenomena as I would attend to if I were answering the question ‘Will there be a third world war?’ I get myself in a position to answer the question whether I believe that p by putting into operation whatever procedure I have for answering the question whether p. (Evans 1982, p. 225).

Evan’s idea here is to link first-order belief-forming processes with second-order belief-forming processes. The same procedure that generates my belief that p also generates my belief that I believe p. Theories building on this idea are transparency accounts of self-knowledge. I suggest that this framework is fruitful for collective self-knowledge. If I am correct and a group can answer the “How do you know that you intend to φ?” question by pointing to the procedure that generated the intention, then there seems to be an important link between knowing the intention and forming the intention.

There is more evidence to support this approach. When I discussed salience for group avowals I pointed out that salience seems to be lacking when the group has not settled on a question and is not in a position to settle the question. Being in a position to settle a question is equal to being in a position to generate an adequate intentional state. Without the procedure to form the first-order state available, the group is unable to provide any report of its doxastic position towards that state. For instance, without the procedure for answering the question whether p, the group
cannot tell whether it believes that p. A transparency theory explains the weaker form of salience for group avowals by showing that it matters whether a group is in a position to form first-order states.

Furthermore, transparency is compatible with REFLECTIVE SCIENCE. The group’s intention-forming process can function without any group member having the intention individually. If collective self-knowledge is as closely connected to the forming first-order states as transparency accounts suggest, then the group’s self-knowledge can also be generated by the same procedure. Hence, the group’s self-knowledge also does not require specific mental states of the group members. Transparency accounts for collective self-knowledge can deny second-order summativism.

Finally, identifying the first-order process of forming mental states with the second-order process of self-ascribing these states is an undemanding explanation. To build a transparency account of collective self-knowledge we need not assume any additional belief-forming processes that are not already independently accepted. Self-ascriptions of first-order mental states are done by the same procedures as forming these first-order mental states. Nothing else required. And if groups can have mental states, as I have explicitly presupposed in the beginning, then we already have procedures for forming first-order mental states in place. Hence, a transparency account of collective self-knowledge does not come with any additional ontological commitments.

At this point I have merely gestured towards a theory that explains collective self-knowledge as being generated by the same procedure as first-order group states. I do not have the space to develop the approach in detail. However, I suggest we try to build such an account and see how far this approach can take us.

6 Conclusion

I argued that we should take collective self-knowledge seriously. Our linguistic practice includes collective speech acts that are sufficiently similar to avowals made by individuals. Social groups can self-ascribe their intentional states in avowals that appear to be at least equally authoritative, and weakly salient. I proposed that this indicates collective self-knowledge, just as avowals of individuals are taken to be evidence for self-knowledge for individuals. It is not conclusive evidence. There is room for explanations that point towards differences between the individual and the collective case. However, the burden of proof is on proponents of these alternative explanations. Accepting collective avowals and collective self-knowledge captures phenomena such as mission statements or manifestos. Any alternative has to capture these at least equally well.

I further provided an argument in favour of taking collective self-knowledge at face value and to some degree as irreducible. In some circumstances social groups seem to be able to know their own intentional states even if no members know these intentional states. The upshot here is a desideratum for any future account of collective self-knowledge. Theories of collective self-knowledge ought to explain how groups can acquire self-knowledge even if no individual member knows the group’s states.
Towards Collective Self-knowledge

I did not offer such a theory of my own in this work, but merely pointed towards the benefits of a transparency account. There is still a lot of work to be done.

Acknowledgements Thank you to Jesper Kallestrup and Aidan McGlynn for early discussions on the topic. A first version of this paper has been presented at the University of Duisburg-Essen at an Oberseminar led by Raphael van Riel and the Social Ontology 2019 conference at the University of Tampere. Thank you to all attendees for their feedback. Finally, thank you to the anonymous reviewers who helped me improve the paper. Open Access funding provided by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Anscombe, G. E. (1957). *Intention*. Oxford: Basil Blackwell.
Armstrong, D. M. (1968). *A materialist theory of mind*. London: Routledge & Kegan Paul.
Bar-On, D. (2004). *Speaking my mind*. Oxford: Oxford University Press.
Bilgrami, A. (2006). *Self-knowledge and resentment*. Cambridge, MA: Harvard University Press.
Bird, A. (2010). Social knowing: The social sense of ‘scientific knowledge’. *Philosophical Perspectives*, 24(1), 23–56.
Bird, A. (2014). When is there a group that knows? Scientific knowledge as social knowledge. In J. Lackey (Ed.), *Essays in collective epistemology* (pp. 42–63). Oxford: Oxford University Press.
Borgoni, C. (2019). Authority and attribution: the case of epistemic injustice in self-knowledge. *Philosophia*, 47(2), 293–301.
Bratman, M. E. (1993). Shared intention. *Ethics*, 104(1), 97–113.
Bretman, M. E. (2014). *Shared agency: A planning theory of acting together*. Oxford: Oxford University Press.
Byrne, A. (2005). Introspection. *Philosophical topics*, 33(1), 79–104.
Byrne, A. (2018). *Transparency and self-knowledge*. Oxford: Oxford University Press.
Carruthers, P. (2011). *The opacity of mind: An integrative theory of self-knowledge*. Oxford: Oxford University Press.
Cassam, Q. (2014). *Self-knowledge for humans*. Oxford: Oxford University Press.
Coliva, A. (2016). *The varieties of self-knowledge*. London: Palgrave Macmillan.
Dotson, K. (2011). Tracking epistemic violence, tracking practices of silencing. *Hypatia*, 26(2), 236–257.
Evans, G. (1982). *The varieties of reference*. (J. McDowell, Ed.) Oxford: Oxford University Press.
Finkelstein, D. (2003). *Expression and the inner*. Cambridge, MA: Harvard University Press.
Gilbert, M. (1989). *On social facts*. London: Routledge.
Gilbert, M. (1990). Walking together: A paradigmatic social phenomenon. *Midwest Studies in Philosophy*, 15(1), 1–14.
Gilbert, M. (2009). Shared intention and personal intention. *Philosophical Studies*, 144(1), 167–187.
Goldman, A. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. Oxford: Oxford University Press.
Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences*, 16, 1–14.

23 First tentative steps towards an alternative have been taken by Schmid (2014, 2016).
Huebner, B. (2014). *Macrocognition: A theory of distributed minds and collective intentionality*. Oxford: Oxford University Press.

Kahneman, D. (2011). *Thinking, fast and slow*. London: Allen Lane.

Kallestrup, J. (2016). Group virtue epistemology. *Synthese*. https://doi.org/10.1007/s11229-016-1225-7.

Lackey, J. (2016). What is justified group belief? *Philosophical Review*, 125(3), 341–396.

List, C. (2005). Group knowledge and group rationality: A judgment aggregation perspective. *Episteme*, 2(1), 25–38.

List, C., & Pettit, P. (2011). *Group agency: The possibility, design, and status of corporate agents*. Oxford: Oxford University Press.

Moran, R. (2001). *Authority and estrangement*. Princeton: Princeton University Press.

Nichols, S., & Stich, S. (2003). *Mindreading: An integrated account of pretence, self-awareness, and understanding other minds*. Oxford: Oxford University Press.

Nisbett, R., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231–259.

Peacocke, C. (1998). Conscious attitudes, attention, and self-knowledge. In C. Wright, B. Smith, & C. Macdonald (Eds.), *Knowing our own minds*. Oxford: Oxford University Press.

Perry, J. (1979). The problem of the essential indexical. *Nôus*, 13(1), 3–21.

Pettit, P. (2001). Collective intentions. In N. Naffine, R. Owens, & J. Williams (Eds.), *Intention in law and philosophy* (pp. 241–254). Farnham: Ashgate.

Pettit, P. (2007). Rationality, reasoning and group agency. *Dialectica*, 61(4), 495–519.

Rödl, S. (2018). Joint action and plural self-consciousness. *Journal of Social Philosophy*, 49(1), 124–136.

Schmid, H. B. (2009). *Plural action: Essays in philosophy and social science*. Dordrecht: Springer.

Schmid, H. B. (2014). Expressing group attitudes: On first person plural authority. *Erkenntnis*, 79, 1685–1701.

Schmid, H. B. (2016). On knowing what we’re doing together. *The epistemic life of groups: Essays in the epistemology of collectives* (pp. 51–72). Oxford: Oxford University Press.

Schmid, H. B. (2017). What kind of mode is the we-mode? In G. Preyer & G. Peter (Eds.), *Social ontology and collective intentionality: critical essays on the philosophy of raimo tuomela with his responses* (pp. 79–94). Cham: Springer International Publishing.

Schmid, H. B. (2018). The subject of “we intend”. *Phenomenology and the Cognitive Sciences*, 17(2), 231–243.

Schmitz, M. (2018). Co-subjective consciousness constitutes collectives. *Journal of Social Philosophy*, 49(1), 137–160.

Schwengerer, L. (Forthcoming). Beliefs over avowals: Setting up the discourse on self-knowledge. *Episteme*. https://doi.org/10.1017/epi.2018.56

Schwitzgebel, E. (2008). The unreliability of naive introspection. *Philosophical Review*, 117(2), 245–273.

Schwitzgebel, E. (2012). Self-Ignorance. In J. Liu & J. Perry (Eds.), *Consciousness and the self: New essays* (pp. 184–197). Cambridge: Cambridge University Press.

Searle, J. (1995). *The construction of social reality*. New York: The Free Press.

Searle, J. (2010). *Making the social world: The structure of human civilization*. New York: Oxford University Press.

Smith, J. (2018). The first-person plural and immunity to Error. *Sciencecco*, 10(49), 141–167.

Snowdon, P. (2012). How to think about phenomenal self-knowledge. In A. Coliva (Ed.), *The self and self-knowledge* (pp. 243–262). Oxford: Oxford University Press.

Tollefsen, D. (2007). Group testimony. *Social Epistemology*, 21(3), 299–311.

Tuomela, R. (2004). Group knowledge analyzed. *Episteme*, 1(2), 109–127.

Tuomela, R. (2005). We-intentions revisited. *Philosophical Studies*, 125(3), 327–369.

Tuomela, R. (2013). *Social ontology: Collective intentionality and group agents*. Oxford: Oxford University Press.

Vierkant, T. (2015). How do you know that you settled a question? *Philosophical Explorations*, 18(2), 199–211.

Weber, M. (1978 (1913)). Some categories of interpretative sociology. In M. Weber, G. Roth, & C. Wittich (Eds.), *Economy and society* (pp. 1375–1380). Berkeley: University of California Press.

Wright, C. (1998). Self-knowledge: The Wittgensteinian legacy. In C. Wright, B. Smith, & C. Macdonald (Eds.), *Knowing our minds* (pp. 13–45). Oxford: Oxford University Press.

Wright, C. (2001). *Rails to infinity: Essays on themes from Wittgenstein’s philosophical investigations*. Cambridge, MA: Harvard University Press.
Towards Collective Self-knowledge

Wright, C. (2015). Self-knowledge: the reality of privileged access. In S. C. Goldberg (Ed.), *Externalism, self-knowledge, and skepticism* (pp. 49–74). Cambridge: Cambridge University Press.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.