



# Collective vice and collective self-knowledge

Lukas Schwengerer<sup>1</sup> 

Received: 18 July 2022 / Accepted: 5 November 2022 / Published online: 6 January 2023  
© The Author(s) 2023

## Abstract

Groups can be epistemically vicious just like individuals. And just like individuals, groups sometimes want to do something about their vices. They want to change. However, intentionally combating one's own vices seems impossible without detecting those vices first. Self-knowledge seems to provide a first step towards changing one's own epistemic vices. I argue that groups can acquire self-knowledge about their epistemic vices and I propose an account of such collective self-knowledge. I suggest that collective self-knowledge of vices is partially based on evidence that a group can generate by performing internal promptings. Whereas these promptings are done mentally in individual self-knowledge, these promptings are done by interactions of group members in the collective case. The group can then acquire inferential self-knowledge of their vices based on the evidence generated by the interactions within the group. Groups thereby bring themselves into a position from which they can combat and change those vices intentionally.

**Keywords** Collective vice · Group vice · Epistemic vice · Self-knowledge · Collective self-knowledge

## 1 Introduction

Self-knowledge seems to be a straightforward step towards fighting one's own epistemic vices. If I am unaware of my closed-mindedness, I will likely not be able to do anything about that vice. A problem I am unaware of can hardly be tackled by me – except by pure luck. In order to intentionally work on my vices, I need to know them. This much seems like a truism. Of course, knowing one's vices can be difficult. I'd like to see myself in a good light and recognizing my faults takes effort. Moreover, some epistemic vices undermine one's ability to detect them in the first place. They are 'stealthy', as Quassim Cassam calls them (2019, p. 145). Nevertheless, that

---

✉ Lukas Schwengerer  
Lukas.Schwengerer@uni-due.de

<sup>1</sup> Fakultät für Geisteswissenschaften, Institut für Philosophie, University of Duisburg-Essen, Universitätsstr. 12, 45141 Essen, Germany

self-knowledge can at least sometimes enable one to combat one's own vices seems undeniable. And if the self-knowledge route to a more virtuous person works for the individual case, it seems that it is worth trying the same path for groups. Groups can be epistemically vicious just like individuals. They can be dogmatic, closed-minded, or gullible. They can also form beliefs in virtue of vicious thinking styles, such as wishful thinking. And just like in the individual case these vices can at least sometimes be detected by the group itself. This is what I aim to show in this paper. Epistemic vices in groups can sometimes be detected by collective self-knowledge. Groups thereby bring themselves into a position from which they can combat and change those vices intentionally.

The paper is structured as follows: I start with a quick overview of the notion of a social group I am working with. I then describe the notion of epistemic vice that I use, which is based on Cassam's (2019) obstructivism. Finally, I discuss collective self-knowledge and a group's ability to detect their own epistemic vices. I do this by suggesting a way to apply Cassam's (2014) account of self-knowledge to social groups. Cassam's accounts of epistemic vices in individuals and self-knowledge for individuals are meant to fit together. I show that the same fitting combination can be established if we build collective versions of Cassam's accounts.

## 2 Social groups

Social groups (hereafter just *groups*) are easy to find in our everyday life. Some are large, structured corporations like Apple or Microsoft, others are middle sized philosophy departments or small reading groups. Even two people going for a walk together already qualify as a group in some sense. What is distinctive about groups is that they are bound together in virtue of some form of shared or joint intention or goal. They are people acting *together*. It is important to emphasize acting together to contrast groups in this sense with mere collections of people that might have the same goal, or showing some apparently coordinated behaviour. Gilbert (1990) has convincingly shown that there is a significant difference between two people going for a walk together and two people with the same target destination walking next to another by chance. Having the same goal is not enough to form a group. Similarly, apparently coordinated behaviour is insufficient to form a social group. Individual actions can add up in a way that creates the appearance of coordination without the individuals acting together. A mass of people entering a convention centre when it opens might appear as if the mass of people acts in a coordinated way, even though any individual only acts by themselves. Weber (1978 (1913)) already points to similar behaviour as homogeneous 'mass behaviour' (*Massenhandeln*). Behaviour of masses merely looks as if it was socially organized, even though it is not an instance of a group acting.

How exactly shared or joint intentions ought to be analysed is still the topic of a lively debate. Bratman (1993, 2014) takes shared intentions to be a combination of intentions of individual members that have a content referring to the group as 'we'. These intentions are supplemented by common knowledge conditions. Searle (1995, 2010) and Tuomela (2004, 2005, 2013) suggest that propositional attitudes can have a

particular ‘we-mode’ that is constitutive of groups. Gilbert (1989, 1990, 2009), Pettit (2001, 2007), List (2005), List and Pettit (2011) and Schmid (2009) opt for different versions of accepting a plural subject as the basis for groups. A plural subject thereby does not commit them to a group mind existing completely independently from the group members, but rather that the group members can constitute a plural subject in virtue of some grounding, supervenience or aggregation relation. I remain neutral on the exact view of social groups. All I assume for this paper is that groups are constituted at least in part by some form of shared or joint intention. How that joint intention comes about or ought to be analysed is not important for my purpose. I am thereby not claiming that the particularities of these views might not play a role at all in collective vices or collective self-knowledge, but that I aim to analyse the topic in a way in which these particularities likely will not matter for the bigger picture. All I need for now is that groups can be differentiated from mere collections of individuals and that groups can form intentional states as a group, regardless of how they are constituted according to the different accounts.

I am working with the assumption that some sort of shared or joint intention is required to constitute a group. This leaves many options open, but not all of them. Alexander Bird (2010) has prominently argued that not all groups are formed by joint intention. Some groups merely require a social cohesion that is established without joint intentions. Bird calls these *organic groups* (Bird, 2010, p. 37) and points to a division of labour within these groups as a paradigmatic feature. According to Bird, these organic groups can also generate collective mental states, such as group beliefs – although Bird characterises the groups in functional rather than mentalistic terms. Furthermore, Bird’s account allows for these groups to have a group belief or group knowledge that is not grounded in individual group members or their mental states. The scientific community counts as an organic group that might have social knowledge that is grounded in a scientific article that is lost in an archive and no individual scientist knows about. However, this comes with problems. Lackey (2014) shows that social knowledge that is not grounded in individuals at all clashes with plausible views on the connection between knowledge and action. Groups seem to act through their members. Moreover, actions fall under an epistemic norm that can be fulfilled by knowledge, although the norm itself might be weaker than knowledge. Combining these two ideas poses a challenge to Bird’s account of social knowledge. Social knowledge could make an individual acting for the group rational, even if the individual has no access to that social knowledge at all. For instance, it would be rational for scientists to approve cancer drugs based on scientific research that no one remembers (Lackey, 2014, p. 288). At least at face value that seems to be an odd consequence and something that ought to be avoided. A scientist approving such drugs would be just making a lucky guess and not following any epistemic norm of action. While this is a challenge for Bird’s view, any account of social groups that accepts some sort of grounding relation to the individual group members is in a much better position. Proponents of such an account can reject the existence of group knowledge that is wholly inaccessible by group members. Hence, I will be working with accounts that ground groups and their mental states in the individual members for the rest of the paper. This does not rule out a Bird-style account of groups completely. It still allows some organic groups that are not grounded in anything other than individual group members. Those organic groups

will still be compatible with the account of collective self-knowledge of epistemic vices I am developing in this paper. The important choice for my account is not the joint intention as a basis for social groups itself, but that social groups are grounded in the group members in some form. Forming a joint intention is merely the most popular way to spell out this grounding relation.

The largely neutral stance on theories of collective intentionality and social groups might appear problematic for some accounts of epistemic vice. Broadly motivational accounts of vices and virtues (e.g. Zagzebski (1996), Baehr (2015), Tanesini (2016, 2018)) require specific cognitive states or processes to constitute vices and virtues. If an epistemic virtue requires the right sort of motivation for cognitive contact with the world (Zagzebski, 1996) a corresponding account of collective virtue and vice would require me to discuss how groups can have any motivation that is sufficiently similar to motivation in individuals. This might not be impossible for some views of collective intentionality, but is a task that requires a lot more space than I can offer in this work. Moreover, it is anything but obvious whether every available option for collective intentionality is compatible with a state of collective motivation that is needed for motivation-based accounts of epistemic vices. Fortunately, I can avoid these issues by using a different, non-motivational account of epistemic vice: Cassam's (2019) obstructivism.

### 3 Epistemic vice

Obstructivism is the view that epistemic vices get in the way of knowledge. As such, it is a consequentialist view. Motivations play no role in the obstructivist account. All that matters is whether something stops a subject from acquiring, keeping or sharing knowledge systematically. Quassim Cassam has championed the view in his *Vices of the Mind* (Cassam, 2019) and I work with Cassam's version of obstructivism. Even though Cassam himself only deals with epistemic vices in individuals, the view itself is well suited for my purpose. All that is needed for a subject – individual or otherwise – to potentially be epistemically vicious is the ability to know. Once a subject can know, then the possibility for something to get in the way of knowledge is open. And fortunately, the ability to know is not a big hurdle for the previously discussed accounts of social groups. If collective intentionality is the decisive factor for being a proper group, then the step to collective knowledge is rather small.<sup>1</sup> Hence, it is not surprising that group belief and group knowledge are universally accepted by anyone who does not take the talk of collective intentionality to be merely metaphorical. Moreover, knowledge ascriptions are often taken to be a paradigmatic example of ordinary language supporting the idea of collective intentionality. We assert claims such as 'Google knows...' or 'The Supreme Court knows...'. This is good news for an

---

<sup>1</sup> Even though the step is small, there is a worry about a belief requirement for group knowledge that is worth mentioning. Under some accounts of collective intentionality, especially Gilbert's (1989) joint commitment model, collective belief has features that differ from belief in individuals. Wray (2001) has argued that groups therefore only have acceptance, but not belief proper. That would be a challenge for views that take knowledge to require belief. In response one can either argue that knowledge does not require belief in the first place (Hakli, 2007), or that belief in groups does not need to have the same features as in individuals.

obstructionist account of collective epistemic vices. If groups can know, then something can get in the way of the group's knowing.

Before I can look at collective vices more closely I want to take a step back and look at individual vices under obstructionism in more detail. First, it is important to highlight that the talk of a vice getting in the way of knowledge makes for a good slogan, but only tells part of the story. It overemphasizes vices as something that stops the acquisition of knowledge. While that is indeed one consequence of epistemic vices, it is not the only one and therefore not a sole criterion for epistemic vice. As Cassam rightly argues, epistemic vices can also impact the keeping and sharing of knowledge. The knowledge acquisition cases are the most obvious. A closed-minded person might not listen to a colleague who warns of a dangerous situation. To use Cassam's example, in 2003 Donald Rumsfeld did not listen to military experts who told him that more troops would be needed to quickly and successfully invade Iraq (Cassam, 2019, pp. 1–2). Rumsfeld failed to listen and thereby failed to know, with disastrous consequences.

Cases of vices preventing the preservation of knowledge are more difficult to find, but they do exist. Consider a prejudiced person who tends to forget scientific contributions by women. The person recognises women's role in scientific research while observing them in the lab, so it is not an issue of acquiring knowledge. However, a little while later the prejudiced person forgets the women's contribution, but remembers the contributions by men. In such a scenario the vice does not obstruct the acquisition of knowledge, but it causes a particular loss of knowledge over time. Memory can be selective in other ways as well. Sometimes particular items in memory are not connected to personal successes, or are not available when reflecting on one's own situation. A good example of selective forgetting as a vice can be found in Thomas Shapiro's (2004) remarks on interviewees forgetting their own historical privilege when assessing their successes. Interviewees were fully aware of their extensive parental assistance until the topic changed to their own achievements and how they reached them. Shapiro reports of one interviewee that her "[...] memory seems accurate as she catalogues all sorts of parental wealthfare with matching dollar figures. [...] However, as soon as the conversation turns to how she and her husband acquired assets like their home, cars, and savings account, her attitude changes dramatically." (Shapiro, 2004, p. 75). In this context the interviewees "[...] describe themselves as self-made, conveniently forgetting that they inherited much of what they own" (Shapiro, 2004, p. 76). One explanation here is that the interviewees did not keep the beliefs relating the parental assistance and inheritance to their current success and assets in their accessible memory. Their memory is selective in a way that lost them knowledge. Of course, this is not the only way to read the case. Perhaps they never made the connection in the first place and the vice prevented knowledge acquisition. However, it seems far more likely that, say, at the moment of inheriting a large amount of money they knew that much of their fortune is inheritance based. But they forgot that fact soon after. And if they systematically forget the effects of a privileged social position and upbringing even beyond the inheritance they seem to have an epistemic vice that prevents them from keeping particular knowledge.

Finally, some epistemic vices are primarily about preventing the sharing of knowledge. This can be either on the side of a speaker, or of a hearer, or both. A speaker

vice might prevent the speaker from testifying even when they easily could or it might make the speaker's testimony ineffective. A speaker's testimony is ineffective if a hearer does not come to know on the basis of the testimony. For instance, a speaker's arrogance might be so off-putting that the potential audience is not willing to listen to the speaker (Cassam, 2019, p. 10). An epistemic vice in the hearer might be a prejudice that prevents the hearer from forming a belief on the speaker's say-so. The paradigmatic example is an instance of testimonial injustice (Fricker, 2007). In a common case of testimonial injustice, a hearer takes a speaker to be less credible because of the hearer's prejudice against the speaker. A woman might not be believed by a hearer because she is a woman and the hearer is prejudiced against women.

I have now presented how epistemic vices obstruct the acquisition, keeping, or sharing of knowledge. One important feature of Cassam's account is still missing, however. Obstruction as such is not sufficient for an epistemic vice. Many things can get in the way of knowledge in some sense, but not all of them are vices. Some are merely epistemic defects. Blindness gets in the way of knowledge. The blind cannot acquire some pieces of knowledge. But blindness is not an epistemic vice. Epistemic vices are only obstructions that are blameworthy or otherwise reprehensible (Cassam, 2019, p. 23). Cassam uses 'otherwise reprehensible' to capture cases in which the obstruction is bad and reflects on the subject, but the subject lacks responsibility for the obstruction. Someone born into an environment that promotes behaviour that prevents knowledge acquisition might not be responsible for their vices, but nevertheless has them. Hence, they are not blameworthy, but still reprehensible because the vices reflect badly on the epistemic subject. Obstructions to knowledge that do not reflect badly on the subject at all, like blindness, are not vices.

Cassam remains open on what exactly can obstruct knowledge in the relevant sense. He is committed to character traits, attitudes, and ways of thinking as epistemic vices, but allows that there are more vices than these. For the collective case all three mentioned options seem available, although some are more easily applied to groups than others. I start with Cassam's take on each option before explaining my adaptation for the group case.

Character traits are taken to be dispositions to act, think and feel. At least with regards to dispositions to act and think these traits can be applied to groups straightforwardly. Groups can act and insofar as forming intentional states counts as thinking groups also qualify in that regard. That groups can be suitably described in terms of dispositions to act and think is easy to see. We often use such descriptions in predicting a group's behaviour. Even in simple cases like describing a football team's tendencies we ascribe dispositions to act in a particular way to the team. I can predict that a team will play especially defensive when ahead in goals because I believe that the team has a disposition to play defensively in that circumstance. Dispositions to act for groups are nothing out of the ordinary. For a group, a disposition to think is just a special case of the disposition to act. And again, there are common occurrences of ascribing a disposition to think in a particular way. If you plan to deceive a group you will consider the group's dispositions to think and ideally find a way to trigger a disposition to think badly. Ascribing a disposition to feel to a group might be more demanding, but for my purpose dispositions to act and think are sufficient. If groups can have these dispositions and the dispositions are sufficiently stable, then they have

everything that is required for character traits. This also fits our ordinary linguistic practice. Sometimes we say that a group is closed-minded, dogmatic, selfish, etc. This is nothing else than ascribing a set of stable dispositions.

Attitudes in the relevant sense are evaluative perspectives towards an object (Cassam, 2019, p. 81). Importantly, these have to be kept apart from propositional attitudes as usually discussed in philosophy. Attitudes capture a preference or aversion towards an object.<sup>2</sup> ‘Liking’ is a paradigmatic example of such an attitude. To like spinach is to evaluate spinach positively and to seek spinach out in appropriate situations. Attitudes can be held towards all kinds of objects, both concrete and abstract. One can even have an attitude towards a political idea. Other attitudes include dislike, contempt, suspicion or hostility. These are usually taken to be attitudes with different functions in our cognitive make-up. Some have a knowledge function, aiming to fulfil a need for knowledge. Others might have a utilitarian function, evaluating objects with regard to their potential use to fulfil some of our non-epistemic needs (Tanesini, 2016). I might have an attitude towards rice that fulfils a utility function to sustain my bodily functions, but my preference of rice has nothing to do with a knowledge function.<sup>3</sup>

Attitudes come in different strengths. The strength of an attitude can be best understood as the strength of the association between the object and the positive or negative valence associated with the object (Tanesini, 2016). For instance, even a low preference for spinach could be a strong attitude if the association is triggered easily and frequently. If only someone mentioning spinach triggers a positive feeling and an attempt to seek out spinach, then one clearly has a strong attitude with positive valence towards spinach. On the other hand, if a dislike of coffee were only triggered when one is actually drinking coffee, but in no other situations, then that dislike would be a weak attitude.

Cassam categorises attitudes into postures and stances (2019, p. 81). Postures are affective and involuntary. Cassam’s example of contempt includes a low regard for someone and a particular feeling that comes with that low regard. Contempt then comes with behavioural manifestations. If I have the attitude of contempt towards another person I might avoid them, for instance (2019, pp. 81–82). Cassam follows Michelle Mason’s (2003) view that the affective component causes the behavioural manifestations. This picture giving a central role to the affective component of postures is troubling for the ascription of postures to groups. Collective emotions are a fringe position in the debate (e.g. Thonhauser (2018)) that I would rather not commit myself to. However, there are at least two other options available without committing to an account of collective feelings. First, I can bracket postures and limit myself to epistemic vices that can be ascribed to groups more easily – traits and ways of thinking. Second, I can tentatively accept group postures as having no affective component.<sup>4</sup> I work with the second option. The reasoning for this is twofold. First, the ordinary linguistic practice fits with an ascription of postures to groups. Take Cassam’s second example

<sup>2</sup> The concept of an attitude in psychology comes from Allport (1935) and is widely accepted. See for instance Banaji & Heiphetz (2010) or Fazio & Olson (2007).

<sup>3</sup> The list of functions is not exhaustive. For a comprehensive discussion of attitudes and their functions see Maio & Olson (2000).

<sup>4</sup> Another option would be to locate an affective component solely in the individual group members. This would also provide a way to ascribe postures to groups. I bracket this option.



for a posture: arrogance<sup>5</sup> (2019, p. 82). It is a common occurrence in ordinary language to ascribe arrogance to a group. A group can be dismissive of views and perspective of others. One would rightly call out such a group as arrogant without committing to any affective state of the group as such. Second, ascribing such attitudes to groups can lead to predictive success. I can predict how a group acts better if I think of their behaviour as manifesting postures (among other things). If I know a group is arrogant in relation to a particular topic, I can tell how the group will react to relevant views outside of their group. Likely they will not listen and stick to their own perspective. Hence, there are at least *prima facie* reasons to work with group postures even without any commitment to an affective component.

Cassam's second category of attitudes – stances – differ from postures. Stances are voluntary and do not include an affective component. A stance is akin to a policy one adopts when dealing with an object. Stances therefore are easy to fit into the framework of groups. Group members can deliberate and decide to take a particular stance towards an object together as a group.

The third kind of vice proposed by Cassam are ways of thinking. These apply to groups without much adjustment needed. When Cassam talks of ways of thinking he is primarily interested in belief-forming processes. Some belief-forming processes are knowledge conducive, whereas other such processes are not. Wishful thinking is the paradigmatic example of a belief-forming process that is not knowledge conducive. Wishful thinking lacks the necessary contact with the world to generate knowledge. Groups can form beliefs by wishful thinking just as easily as individuals can. How exactly such a belief is formed by wishful thinking depends on the account of belief formation of choice for groups. Here the specific differences of, say, Gilbert's (1989, 1990, 2009) account and List and Pettit (2011) make a difference. In Gilbert beliefs are primarily formed by joint commitment of the group members, in List and Pettit group beliefs are determined by an aggregation function that relates individual beliefs to group belief. However, regardless which account one accepts, they all allow for group beliefs that do not amount to knowledge because their production lacks reliability. This provides room for belief-forming processes that get in the way of knowledge. A Gilbert group can jointly accept a belief for reasons that have nothing to do with the facts (cf. Lackey (2016, 2021), Schwengerer (2021)), and a List and Pettit group can aggregate beliefs that were already individually based on wishful thinking. In both versions we end up with a flawed way of thinking – a kind of wishful thinking on the group level.

Overall Cassam's obstructivism identifies epistemic vice with a character trait, an attitude, or a way of thinking that systematically obstructs the gaining, keeping or sharing of knowledge. And all these forms of epistemic vices seem to apply to groups without too much difficulty. Hence, I can ascribe epistemic vices to groups. At this point the groundwork is done and I can shift to the question of collective self-knowledge and the detection of epistemic vices in a group by the group itself. This seems to be a prerequisite for the group in order to be able to combat their own vices intentionally.

---

<sup>5</sup> Arrogance can be both character trait and posture for Cassam. For instance, it can be a posture if it relates to a specific object like a particular topic or subject matter, but is not manifesting in the more general behaviour of a person.



## 4 Collective self-knowledge of collective vices

Theories of collective self-knowledge are rare. The closest thing can be found in Schmid (2014, 2016) and Schwengerer (2022). Both aim primarily at establishing collective self-knowledge as a phenomenon and try to describe the properties of such collective self-knowledge in comparison with self-knowledge in individuals. However, neither provides more than a sketch towards an account of collective self-knowledge. Schwengerer (2022) suggests that the most promising account will likely be a transparency account, taking on an insight of Gareth Evans (1982) about the relation of self-knowledge to the attitudes known. Evans suggests for individuals that the same procedure that generates my belief that *p* also generates the belief that I believe that *p*. Applying this idea to collective self-knowledge has the advantage of avoiding any commitment to a distinct kind of group level introspection. All that a transparency account requires are procedures that can generate first-order attitudes, and that requirement is already met by all accounts of collective intentionality I considered earlier. Even if such a transparency account can be built, it is – at least on its own – not the right account for detecting epistemic vices in groups. Schwengerer (2022) focuses on self-knowledge of propositional attitudes and primarily self-knowledge of beliefs. Epistemic vices in the obstructivist picture are not propositional attitudes like beliefs. They are character traits, attitudes in the psychologist's sense, or ways of thinking. An account that only explains how groups can know their intentional states will not do the trick. The path forward therefore has to start with looking for an account of self-knowledge of virtues and vices.

The majority of discussions on self-knowledge are focused on mental states. The main puzzle about self-knowledge is taken to be the apparently peculiar access one has to one's own mental states, and the apparent security and privilege that access comes with. Hence, the desiderata for theories of self-knowledge are often the peculiar and privileged access (e.g. Byrne (2005) or Fernández (2013)), or at least the appearance of that peculiar and privileged access (e.g. Carruthers (2011)). One exception is Cassam (2014), who builds his account of self-knowledge explicitly with knowledge of character traits, attitudes and aptitudes in mind. He does not ignore self-knowledge of mental states, but argues that self-knowledge in this broader sense ought to be part of our focus as well. This broader sense of self-knowledge is *substantial self-knowledge*, whereas the narrower, especially secure self-knowledge of mental states is labelled *trivial self-knowledge*<sup>6</sup> (Cassam, 2014).

The difference between substantial and trivial self-knowledge is a matter of degree. Paradigmatic cases of trivial self-knowledge are rather easy to acquire. I believe that I believe it is 2022. It takes little effort to form that self-belief. Moreover, while not infallible, I seem to be highly reliable in forming a belief about my belief that it is 2022. There is little interference with my detection of the first-order belief, regardless how the mechanism detecting that belief functions in detail. And because of my high reliability it seems usually inappropriate to challenge my second-order belief. If I say 'I believe it is 2022', then it seems infelicitous for you to respond 'No, you do not believe that'.

<sup>6</sup> Though Cassam willingly admits that explaining the particular feature of trivial self-knowledge is not trivial or easy at all and that some instances of self-knowledge of belief or desire can fall into substantial self-knowledge.

Trivial self-knowledge also seems to be generated in a peculiar way, such that asking for evidence for that self-knowledge is inappropriate. If you ask me ‘How do you know that you believe it is 2022?’ I will be puzzled and cannot offer you evidence. I just know! Trivial self-knowledge appears to be special in a way that manifests in our ordinary linguistic practice. Paradigmatic cases of substantial self-knowledge, however, have very different properties. Self-knowledge of a character trait such as open-mindedness is nothing like knowing that I believe it is 2022.<sup>7</sup> To know my character takes effort. It is difficult. Of course, I would like to be open-minded. Open-mindedness is a trait I value highly. But that desire to see myself as open-minded itself threatens my ability to evaluate my own character. I want to see myself in a good light and I tend to interpret my behaviour accordingly. I might deceive myself and take closed-minded behaviour to be open-minded behaviour simply because I do not want to see myself as a closed-minded person. It might be difficult to identify traits that do not fit with my own self-conception (Cassam, 2014, p. 30). Partially because of this relation to my self-conception the process of acquiring substantial self-knowledge will be much less reliable. There are more factors in play that interfere with a correct assessment of my character. Knowing one’s own character traits is difficult. The challenge of identifying one’s own traits also leads to differences in everyday linguistic practice. It is much more appropriate to challenge a self-ascription of a character trait than to challenge a self-ascription of a mental state. Whereas challenging my claim that I believe that it is 2022 was infelicitous, challenging my claim that I am open-minded seems perfectly fine in many occasions. I might have interpreted my own behaviour in a biased way and my friend now calls me out on this bias. Similarly, asking for evidence seems more appropriate for claims of substantial self-knowledge. ‘How do you know that you are open-minded?’ does not seem like an inappropriate question. And it can be answered properly by listing past instances of manifesting open-mindedness as evidence. Paradigmatic cases of substantial self-knowledge clearly have different features than paradigmatic cases of trivial self-knowledge. Cassam nevertheless takes these kinds of self-knowledge to be only different in degree with in-between cases in which self-knowledge is more difficult to achieve than the most trivial self-knowledge, but less difficult than the most substantial self-knowledge. Self-knowledge about one’s desire to move to a different city might be more difficult than knowledge that one believes it is 2022, but not quite as difficult as knowing whether one is open-minded.<sup>8</sup>

The most important part of Cassam’s discussion of substantial self-knowledge for the prospect of building my account of collective self-knowledge of epistemic vices is the evidential nature of substantial self-knowledge. Substantial self-knowledge is based on evidence. This is well suited for an application in collective self-knowledge, because an evidence-based account avoids any commitment to a distinct form of

---

<sup>7</sup> Some substantial self-knowledge is still knowledge of one’s own mental states, but those mental states must play a central role in the subject’s life and values. Their detection functions more like the detection of character traits than that of detecting my belief that I believe it is 2022. A good example is Lawlor’s (2009) case of self-knowledge about wanting another child, which is knowledge of a desire but nevertheless substantial self-knowledge that can be difficult to acquire. I will come back to that example later. For a discussion of how some knowledge of one’s beliefs and desires can be substantial see Cassam (2014, p. 33).

<sup>8</sup> Cassam has more relevant features to consider. See Cassam (2014, pp. 30–32).

collective introspection.<sup>9</sup> If substantial collective self-knowledge is based on evidence, then it functions like any other group belief that is based on evidence. The only difference is that the content of the relevant substantial self-knowledge will be about a group's character traits, attitudes, or ways of thinking, and the evidence has to provide a reliable path to knowledge of those traits, attitudes or ways of thinking. A group knowing their own epistemic vices is not all that different from a group knowing anything else.

I am now at a crucial point in my discussion. I have argued with Cassam (2014) that collective self-knowledge of epistemic vices is substantial self-knowledge and therefore based on evidence. But how can a group form substantial self-knowledge about their epistemic vices exactly? What is the evidence that allows a group to infer that the group has a particular vice?

In one way the answer to this question seems easy: it is behavioural evidence. All views on collective intentionality share that they reject a group mind that is independent from the minds and actions of individuals. Groups are grounded in their members in one way or another. Hence, what evidence is accessible by a group has to be based on what evidence is accessible to group members. And the only evidence that is accessible by a significant number of group members is behavioural evidence. Of course, the mental states of an individual member can also be accessed by that member, but without those states being articulated or otherwise presented in behaviour to the rest of the group it seems unlikely that the individual's mental states can become evidence for the group. This is also supported by the individual accounts of collective intentionality mentioned earlier. For instance, common knowledge conditions in Bratman (1993, 2014), Gilbert (1989, 1990, 2009) and Tuomela (2004, 2005, 2013) about individual's mental states or about one's willingness to jointly commit to something are usually satisfied by communication or behavioural clues. A mental state of an individual that does not manifest in behaviour in any way is unlikely to be part of common knowledge between agents. My claim that the evidential base for substantial self-knowledge in groups is behavioural is therefore compatible with being largely neutral on the question of collective intentionality. However, merely proposing that the evidence is behavioural evidence does not make for a sufficient explanation. It is insufficient for two reasons: first, a crude behavioural basis is insufficient for individual substantial self-knowledge, so it is difficult to see why it should be sufficient for an analysis of collective substantial self-knowledge; and second, without more elaboration it is unclear how a behavioural basis can explain differences in the access that a group has to its own vices compared to the access that someone outside the group has to the group's vices. It might turn out that there are little to no such differences, but at least *prima facie* a group seems to have peculiar paths available to assess themselves that are not – or at least not usually – open for outsiders.

Let me begin with the first reason and use that as a starting point to elaborate on behavioural evidence in more detail. In individual substantial self-knowledge behaviourism is rejected as an account of the evidential basis because it cannot explain differences between various cases of substantial self-knowledge. Moreover, it is phenomenologically inadequate. It does not capture how one experiences the acquisition

---

<sup>9</sup> Remember that this was also already part of Schwengerer's (2022) reasoning for a transparency account.

of substantial self-knowledge in many cases. Both explanatory and phenomenological inadequacy of behaviourism can be illustrated with Krista Lawlor's (2009) example of finding out that one wants to have another child. In her example a young woman, Katherine, comes to know that she wants another child by an inference from evidence.<sup>10</sup> However, it is not only behavioural evidence that is used for such an inference. Katherine uses *internal promptings* to figure out whether she wants another child. Internal promptings are a mixed bunch of states that are generated in response to experienced and imagined situations. Katherine imagines herself in a situation with another child and observes the sensations, images and sentence-like thoughts that follow in that imaginary situation. She then uses these states generated by the imaginary prompt as part of her inference basis. She comes to know that she wants another child because in many different everyday experiences and in even more imagined circumstances her reactions are best explained as being caused by a desire for another child. Importantly, these imagined responses in her exercise of internal promptings are not behavioural and therefore cannot be captured by behaviourism. As Cassam correctly states: "You can't lump together all substantial self-knowledge and dismiss it with the remark that it's all based on behavioural evidence. No doubt some of it is based on behavioural evidence, but a lot of it isn't" (Cassam, 2014, p. 173).

But if individual substantial self-knowledge is not all behavioural, then how can substantial self-knowledge in groups be anything like the individual case? Cassam himself focuses on individuals and hence does not provide an answer. There is no group mind independent from the individuals that could perform a ritual of internal promptings. Fortunately, there does not need to be. All that I require is a distinction between different forms of behavioural evidence. Not every behaviour that occurs within a group is behaviour of the group. I distinguish between *external behaviour* of the group and *internal behaviour* in the group. External behaviour captures behaviour performed by the group as a collective, or by individual group members as representing the group. A member represents the group in an action, for instance, when they perform an action in order to fulfil the group's aim or intention. They are acting as part of the group, not as an individual. If a philosophy department sends a student an admission letter, the physical action of mailing the letter has to be performed by an individual department member, but that member is acting as part of the department, not merely as an individual. In such a case ordinary language has no problem ascribing the action to the group. The philosophy department sent an admission letter. This kind of group behaviour is analogous to behavioural evidence considered in individual self-knowledge. Limiting an account of collective substantial self-knowledge to external behaviour lacks just as much nuance as a crude behavioural evidence basis for individual substantial self-knowledge. The picture needs a different kind of evidence that mirrors internal promptings. This is internal behaviour. By internal behaviour I want to capture behaviour of group members that leads to decisions and intentional states of the group. It is behaviour that one might also describe as a group's deliberation process. How exactly a group deliberates varies from group to group. Some

---

<sup>10</sup> This falls into substantial self-knowledge even though it is knowledge about a desire because of the central role the desire plays in Katherine's life and values.

have strict rules and internal hierarchies structuring deliberation. Some include various voting practices that aggregate individual opinions into a group position. Others might be more freeform – a back and forth of giving reasons and arguments between the group members until an acceptable compromise is found. Regardless how the details might look like for a particular group, this group internal process is not as such already behaviour of the group. It precedes group behaviour. The group is making up its mind in virtue of the group members interacting. Of course, there is still a sense in which ordinary languages ascribes the deliberation to the group, but here even ordinary language users describe the process differently than they would do with external behaviour. Mentalistic language is used to describe the group's deliberation. One might say 'The department is considering my application' or 'The department is thinking about admitting me to the program' without such a sentence being inappropriate. Ordinary language makes a distinction between the internal and the external behaviour of a group.

My suggestion is that internal behaviour is of a kind that can facilitate the group analogue to internal promptings. The important part of internal prompting in individuals was the ability to put oneself into imaginary situations that generate responses to those situations. These responses can be concrete behaviour, but for the most part they are imaginary behaviour, sensations, or even more complex thoughts coming to one's mind. Many of those responses therefore are not strictly behavioural. Nevertheless, all these responses are available as evidence for the individual in acquiring substantial self-knowledge. I propose that something very similar can occur in groups on the level of internal behaviour, of deliberation within the group. Group members can deliberate on imagined situations. The group can perform internal promptings by its members deliberating on such an imagined situation. They can discuss how a group ought to respond in a situation, even if that situation is not currently present.

Let me illustrate this with an example: take the structured group of Microsoft. Suppose Microsoft wants to find out whether the group wants to keep expanding. Microsoft wants to find out whether the group has a desire to expand, which seems to be a group mental state. This might not qualify as paradigmatic substantial self-knowledge, but is suitable as a first example. One way in which Microsoft can find out about its desire is by deliberating on possible scenarios that allow the company to expand. The operative members of decision making might come together and consider a range of cases that might happen in the future. For instance, a group member puts forwards a scenario in which a smaller independent developer of an Office alternative has financial troubles and offers to sell stocks. The operative members of Microsoft now deliberate on what they would do in such a scenario. Would they buy the stocks? If so, how many would they buy? How much control would they want to have over the rival company? Would they integrate it into Microsoft's department developing the next version of its Office suite? The group can play through a range of such scenarios, from small opportunities with low cost to giant corporate takeovers of billion-dollar companies. All this deliberation on possible scenarios generates responses that are available for the group as an inferential basis to assess whether Microsoft wants to expand. If for almost all scenarios the group members conclude that Microsoft should seize the opportunity and invest whenever possible, then the group has excellent evidence that Microsoft wants to expand – Microsoft has the relevant desire. And the process to acquire that

piece of self-knowledge looks remarkably similar to Katherine figuring out whether she wants another child. Microsoft can perform internal promptings.

Microsoft's internal promptings are still behavioural in a sense. The deliberation within the group is observable behaviour of members. But it is internal behaviour – behaviour that is within the group but not by the group. Internal behaviour can play the same role as mental processes do for internal promptings in individuals. Moreover, the concept of internal behaviour can also explain how a group's generation of substantial self-knowledge is special. It is not special in the same sense that trivial self-knowledge is sometimes taken to be non-evidential. The internal prompting in groups is observable and could be in principle evidence for anyone. However, the evidential access is special because the group is in a particularly good position to perform internal promptings. The group can intentionally start internal promptings via its group members. And the group is in an excellent position to observe those internal promptings and responses to them because the group members themselves are performing the prompting. Outsiders might be fortunate enough to observe the group deliberation, but the group itself is always present when the group members deliberate. The group's access is not in principle privileged, but in practice because internal prompting is nothing but group deliberation.

I have now provided an account of collective substantial self-knowledge via internal promptings in group deliberation. So far, I have only given an example of detecting a desire this way, but the aim is to detect group vices. How does the proposed structure function for vices? It does so rather straightforwardly in many cases. Consider the vice of closed-mindedness. An individual that is closed-minded has a need for closure and is reluctant to take alternative possibilities seriously (Bar-Joseph & Kruglanski, 2003). That description can apply to a group without a problem. A group that is closed-minded can attempt to detect that vice by observing their external behaviour and by performing internal promptings. Consider a small investment firm consisting of a tight-knit group without much hierarchical structure. In reviewing their own behaviour to optimize profits the group wants to find out if they are closed-minded, a trait that would hurt them financially in the long run. As evidence they can consider past decisions by the group and review their past decision-making process. However, they can also generate evidence by internal prompting. They can imagine a situation in which data conflicting with their predictions comes in, or scenarios in which rival companies publish predictions that conflict with the group's own calculations. The group then deliberates on how they would and should react in each imagined scenario. This deliberation generates evidence for the group that can be used to assess whether the group is behaving closed-mindedly. Hence, the group can in a next step look together at that evidence to evaluate themselves. If in most prompted scenarios they would default to give the new conflicting alternatives to their own predictions little weight, then they have a good indicator that the group might be closed-minded. If that is the conclusion they come to – and that conclusion is correct – then the group acquired substantial self-knowledge about their closed-mindedness. Now they are in a prime position to combat that vice head on. And similar stories can be given for all kinds of vices, regardless of them being character traits, attitudes, or ways of thinking. Internal promptings are performed by the group members deliberating together, these

promptings generate evidence,<sup>11</sup> and that evidence can be used as an evidence basis to infer the group's epistemic vices. This is how collective self-knowledge of epistemic vices is possible.

To complete the picture, take the case of the epistemic vice of wishful thinking. A group that wants to find out whether it is merely wishful thinking that something will happen can perform an internal prompting routine with scenarios in which the payoffs for different results change. Would the group still believe that *p* in a scenario in which *p* was bad for the group? In this way the group might find out that the beneficial effects of *p* had a significant impact on the group believing that *p*. Hence, the group can infer that it formed a belief in a vicious way of thinking by letting the group's wishes impact the belief formation. The internal prompting story for acquiring knowledge about the group's epistemic vices generalises for all kinds of epistemic vices in groups.

I want to end with a note of caution. So far, I have been largely optimistic about the possibility of collective self-knowledge of vices. I provided a framework based on Cassam (2014, 2019) that allows me to ascribe epistemic vices to groups and that gives me room to explain how groups can infer their own epistemic vices from behaviour, with an emphasis on internal behaviour in the joint deliberation of group members. Sometimes groups can detect their own vices in that way. However, vices are not always easy to detect by oneself. They often conflict with the self-conception of an individual or group. No one wants to see themselves in a bad light. Biases might always prevent one from detecting one's own vices. Moreover, some vices can be *stealthy* because of their particular nature (Cassam, 2019, p. 145). Stealthy vices are those that undermine their own detection. They are regular companions in individuals. The arrogant person has difficulties recognising their own arrogance because they are overconfident in their own belief-forming processes. The closed-minded one who is not open to alternative possibilities is thereby also closed to the possibility of being closed-minded. If these vices are applicable to groups in a straightforward way, as I have been suggesting earlier, then the very same structural problems of stealthy vices will show up in groups. A group's self-knowledge of epistemic vices is just one potential path towards detecting them. Sometimes it will be successful. At other times it will not and outside assistance is needed.

A sceptic might worry here that collective self-knowledge of vices is an undesirable path because it can be so difficult and unreliable. Both individuals and groups are wrong about themselves frequently, so outside testimony is the better way to detect vices. I think this sceptical conclusion goes too far for two reasons. First, even if collective self-knowledge is not the best way to detect vices, in some circumstances it might be the only option available. Moreover, even if it is not the only option, it seems reasonable to use all the tools available to detect the group's vices and virtues. As long as the group is aware of potential difficulties with collective self-knowledge it can be an additional source of knowledge. Second, wholly relying on outside testimony would be a bad idea as well. Sometimes outsiders can be wrong about a group's virtues and vices – even intentionally wrong. A malicious agent might tell a group that they are epistemically virtuous when they are not, or that they are epistemically vicious when they are not. Both have undesirable consequences. A group might change its behaviour

---

<sup>11</sup> In the form of internal behaviour.



to the worse if they accept testimony about the group's virtues and vices without checking themselves at all. The group needs collective self-knowledge to compare to testimony from outside the group and then consider the total evidence. Otherwise the group could be too easily manipulated from outside the group.<sup>12</sup> Hence, even with all its difficulties it is important and worthwhile for groups to aim for self-knowledge of epistemic vices.

## 5 Conclusion

I have argued for an account of collective self-knowledge of epistemic vices based on an obstructivist notion of vices and an inferential account of substantial self-knowledge. Groups can detect their own epistemic vices based on an inference from evidence. That evidence is broadly behavioural, wherein part of that basis is external behaviour that the group as such performs, and the other part is internal behaviour. The internal behaviour in question is the group's deliberation process in which the relevant group members deliberate on what the group would do or believe in response to imagined scenarios. Internal prompts are performed by the group members deliberating together, these prompts generate evidence, and that evidence can be used as an evidence basis to infer the group's epistemic vices. This path towards collective self-knowledge of vices is difficult and can fail for a variety of reasons. Sometimes the group's self-conception might get in the way, sometimes the vices themselves might prevent their detection. Nevertheless, collective self-knowledge of group vices is possible.

**Acknowledgements** Work on this paper has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 462399384. Open Access funding provided by Projekt DEAL. A version of this paper has been presented in a research seminar at the University of Duisburg-Essen led by Raphael van Riel and Thomas Spitzley.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

## Declarations

**Conflict of interest** This manuscript has no conflict of interest.

**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/>.

---

<sup>12</sup> Winokur (Forthcoming) calls this sort of vulnerability in interpersonal reasoning without self-knowledge an instance of being taken as an epistemic hostage. For Winokur some sort of epistemic security and privilege is needed to prevent that. Even though the concept of an epistemic hostage is developed for individuals it also applies to the group case.

## References

- Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *Handbook of social psychology* (pp. 798–844). Clark University Press.
- Baehr, J. (2015). *Cultivating good minds: A philosophical and practical guide to educating for intellectual virtues*. Retrieved from <https://intellectualvirtues.org/why-should-we-educate-for-intellectual-virtues-2-2/>
- Banaji, M. R., & Heiphetz, L. (2010). Attitudes. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (pp. 353–393). Wiley.
- Bar-Joseph, U., & Kruglanski, A. W. (2003). Intelligence failure and need for cognitive closure: On the psychology of the yom kippur surprise. *Political Psychology*, pp. 75–99.
- Bird, A. (2010). Social knowing: The social sense of “scientific knowledge.” *Philosophical Perspectives*, 24(1), 23–56.
- Bratman, M. E. (1993). Shared intention. *Ethics*, 104(1), 97–113.
- Bratman, M. E. (2014). *Shared agency: A planning theory of acting together*. Oxford University Press.
- Byrne, A. (2005). Introspection. *Philosophical Topics*, 33(1), 79–104.
- Carruthers, P. (2011). *The opacity of mind: An integrative theory of self-knowledge*. Oxford University Press.
- Cassam, Q. (2014). *Self-knowledge for humans*. Oxford University Press.
- Cassam, Q. (2019). *Vices of the mind: From the intellectual to the political*. OUP.
- Evans, G. (1982). *The varieties of reference*. (J. McDowell, Ed.) Oxford: Oxford University Press.
- Fazio, R. H., & Olson, M. A. (2007). Attitudes: Foundations, functions and consequences. In M. A. Hogg & J. Cooper (Eds.), *The Sage handbook of social psychology* (pp. 123–145). SAGE.
- Fernández, J. (2013). *Transparent minds: A study of self-knowledge*. Oxford University Press.
- Fricker, M. (2007). *Epistemic injustice*. Oxford University Press.
- Gilbert, M. (1989). *On social facts*. 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.
- Hakli, R. (2007). On the possibility of group knowledge without belief. *Social Epistemology*, 21(3), 249–266.
- Lackey, J. (2014). Socially extended knowledge. *Philosophical Issues*, 24, 282–298.
- Lackey, J. (2016). What is justified group belief? *Philosophical Review*, 125(3), 341–396.
- Lackey, J. (2021). *The epistemology of groups*. Oxford University Press.
- Lawlor, K. (2009). Knowing what one wants. *Philosophy and Phenomenological Research*, 79(1), 47–75.
- 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 University Press.
- Maio, G. R., & Olson, J. M. (2000). *Why we evaluate: Functions of attitudes*. Lawrence Erlbaum.
- Mason, M. (2003). Contempt as a moral attitude. *Ethics*, 113(2), 234–272.
- Pettit, P. (2001). Collective intentions. In N. Naffine, R. Owens, & J. Williams (Eds.), *Intention in law and philosophy* (pp. 241–254). Ashgate.
- Pettit, P. (2007). Rationality, reasoning and group agency. *Dialectica*, 61(4), 495–519.
- Schmid, H. B. (2009). *Plural action: Essays in philosophy and social science*. 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 University Press.
- Schwengerer, L. (2021). Defending joint acceptance accounts of justification. *Episteme*. <https://doi.org/10.1017/epi.2020.55>
- Schwengerer, L. (2022). Towards collective self-knowledge. *Erkenntnis*, 87(3), 1153–1173.
- Searle, J. (1995). *The Construction of Social Reality*. The Free Press.
- Searle, J. (2010). *Making the social world: The structure of human civilization*. Oxford University Press.
- Shapiro, T. (2004). *The hidden cost of being African American: How wealth perpetuates inequality*. Oxford University Press.
- Tanesini, A. (2016). Teaching virtue: Changing attitudes. *LOGOS & EPISTEME*, 7(4), 503–527.
- Tanesini, A. (2018). Epistemic vice and motivation. *Metaphilosophy*, 49(3), 350–367.

- Thonhauser, G. (2018). Shared emotions: A steinian proposal. *Phenomenology and the Cognitive Sciences*, 17(5), 997–1015.
- 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 University Press.
- 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.
- Winokur, B. (Forthcoming). Self-knowledge and interpersonal reasoning. *Dialectica*.
- Wray, B. K. (2001). Collective belief and acceptance. *Synthese*, 129(3), 319–333.
- Zagzebski, L. (1996). *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge University Press.

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