Mandevillian Intelligence: From Individual Vice to Collective Virtue

Paul R. Smart, Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK (ps02v@ecs.soton.ac.uk)

Abstract

Mandevillian intelligence is a specific form of collective intelligence in which individual cognitive shortcomings, limitations and biases play a positive functional role in yielding various forms of collective cognitive success. When this idea is transposed to the epistemological domain, mandevillian intelligence emerges as the idea that individual forms of intellectual vice may, on occasion, support the epistemic performance of some form of multi-agent ensemble, such as a socio-epistemic system, a collective doxastic agent, or an epistemic group agent. As a specific form of collective intelligence, mandevillian intelligence is relevant to a number of debates in social epistemology, especially those that seek to understand how group (or collective) knowledge arises from the interactions between a collection of individual epistemic agents. Beyond this, however, mandevillian intelligence raises issues that are relevant to the research agendas of both virtue epistemology and applied epistemology. From a virtue epistemological perspective, mandevillian intelligence encourages us to adopt a relativistic conception of intellectual vice/virtue, enabling us to see how individual forms of intellectual vice may (sometimes) be relevant to collective forms of intellectual virtue. In addition, mandevillian intelligence is relevant to the nascent sub-discipline of applied epistemology. In particular, mandevillian intelligence forces us see the potential epistemic value of (e.g., technological) interventions that create, maintain or promote individual forms of intellectual vice.

Keywords: social epistemology, virtue epistemology, applied epistemology, collective intelligence, virtue reliabilism, virtue responsibilism, collective cognition.

1. Introduction

The performance of groups, teams and organizations is a common focus of interest for those working in the cognitive (e.g., Hutchins, 1995), computational (e.g., Kearns, 2012) and social (e.g., Kerr and Tindale, 2004) sciences. It is also an increasingly popular focus of interest for those working in epistemology. The sub-discipline of social epistemology, for example, seeks, at least in part, to understand the way in which issues of social structure, social organization, social practice and socio-technical interaction affect the collective generation of epistemically-relevant commodities, such as true belief, knowledge and understanding (Goldman, 2011). Given that many of our most cherished intellectual accomplishments (e.g., advances in scientific knowledge) are typically forged in these crucibles of collective effort, it is vital that we develop a better (meta-epistemological?) understanding of the forces and factors that affect the outcomes of socio-epistemic interactions.

It might be thought that the best way to enhance the epistemic performance of social groups would be to enhance the cognitive or epistemic capabilities of the individual group members. This makes intuitive sense, since it would seem that by improving the cognitive abilities of individual agents (e.g., by enhancing their powers of mnemonic recall or improving their ability to engage in reason-respecting chains of thought) we thereby provide the basis for improved performance at the collective level (i.e., we enhance the problem-solving and decision-making capabilities of the group itself). This intuitive view of the relationship between individual ability and collective performance is challenged by the notion of *mandevillian intelligence*.[1] The term "mandevillian intelligence" names a specific form of collective intelligence in which certain kinds of (individual-level) cognitive and epistemic properties are seen to be causally-relevant to the expression of intelligent behaviour at the collective level. In particular, the properties of interest in mandevillian intelligence are ones that are typically denigrated as cognitive or epistemic 'vices'. These include cognitive processing limitations (e.g., limitations in attentional capacity), cognitive shortcomings (e.g., forgetting) and cognitive biases (e.g., confirmation bias). What makes these properties vices is that they are generally seen to undermine the performance of the individual on a variety of cognitive and epistemic tasks.

With these clarifications in place, mandevillian intelligence can be defined as follows:

> **Mandevillian Intelligence**
> Mandevillian intelligence is a specific form of collective intelligence in which cognitive and epistemic properties that are typically seen as vices at the individual level can, on occasion, play a positive functional role in supporting the emergence of intelligent behaviour at the collective level.

There are few points that are worth noting about this definition. Firstly, mandevillian intelligence is cast as a specific form of collective intelligence. As such, it only applies to situations where the term "collective intelligence" is itself deemed to be applicable. We do not, therefore, have mandevillian intelligence if we observe the presence of cognitive vice (or virtue), but we do not observe collective intelligence.[2] Secondly, mandevillian intelligence is only applicable to those situations in which collective success is causally linked to individual properties that are properly labelled as cognitive or epistemic vices. If the relevant properties fail to qualify as vices (from the perspective of individual cognitive functioning), then we do not have mandevillian intelligence. Finally, it is important to note that there is nothing in the definition of mandevillian intelligence that would lead us to conclude that individual forms of cognitive/epistemic vice will inexorably lead to positive outcomes at the collective level: it is perfectly possible for individual cognitive vices to undermine or enhance collective performance depending on the specific context in which collective cognitive processing occurs.

The value of the concept of mandevillian intelligence is that it forces us to acknowledge the potential role of individual vice in securing collective forms of cognitive success. With the concept of mandevillian intelligence to hand we can thus question the extent to which the seemingly negative

---

[1] Mandevillian intelligence is named after the Anglo-Dutch philosopher and economist, Bernard Mandeville, who wrote about the causal link between private (individual) vice and public (collective) benefits.
[2] It is thus a mistake to regard situations where individual virtues undermine collective intelligence as genuine instances of mandevillian intelligence. This is because the relevant situations are not ones that can be seen as indicative of collective intelligence.

or undesirable cognitive properties of individual agents (broadly construed as cognitive vices) should always be seen as detrimental to group performance. Indeed, mandevillian intelligence encourages us to entertain the possibility that individual cognitive vices may, on occasion, provide a productive route to collective forms of 'cognitive virtue'. In other words, mandevillian intelligence enables us to see how individual vices may sometimes yield group-level properties that enable the group to discover or track the truth.

The main aim of the present chapter is to introduce the notion of mandevillian intelligence to the epistemology community. In particular, I will attempt to sketch an initial vision concerning the potential relevance of mandevillian intelligence to contemporary epistemology. Of course, given that mandevillian intelligence is presented as a specific form of collective intelligence, it should be clear that mandevillian intelligence is of potential relevance to the branch of epistemology known as social epistemology (Goldman and Whitcomb, 2011; Haddock et al, 2010). Beyond this, however, mandevillian intelligence may also strike a chord with popular approaches in individualistic analytic epistemology. This is because issues of vice and virtue, which lie at the heart of mandevillian intelligence, are a major focus of philosophical attention for what is known as virtue epistemology (Greco and Turri, 2012). Mandevillian intelligence seems to be of crucial relevance, here, because it suggests that a virtue-theoretic conception of socially-produced knowledge (or collective knowledge) might need to appeal to the role of intellectual vice as much as it does the role of intellectual virtue. There are also reasons to believe that a consideration of mandevillian intelligence is relevant to the design and evaluation of socio-technical systems, especially those that have, as their primary goal, the generation or maintenance of collective knowledge. With the advent of the World Wide Web, such forms of socio-epistemic 'machinery' are increasingly commonplace. Prominent examples include the online encyclopedia, Wikipedia (see Fallis, 2011), as well as a variety of so-called citizen science systems (Khatib et al, 2011; Lintott et al, 2008). Given the potential relevance of contemporary epistemology to the design of these systems, it seems that mandevillian intelligence may have as much traction with the emerging sub-discipline of applied epistemology (Chase and Coady, forthcoming) as it does with the already well-established sub-disciplines of social epistemology and virtue epistemology.

2. Virtue Epistemology: Reliabilism, Responsibilism and Virtue Relativism

As a means of better understanding the epistemological relevance of mandevillian intelligence, it helps to have a better understanding of the significance of vice and virtue in contemporary epistemology. The main point of interest, here, is what is known as virtue epistemology (Greco and Turri, 2012). Virtue epistemology names an approach to understanding knowledge that is rooted in the intellectual virtues, where these are understood as the properties of an agent that enable the agent to track the truth. The thing that makes something an intellectual virtue from an epistemological perspective is thus its truth conduciveness, or its tendency to lead to a preponderance of true (as opposed to false) beliefs.

A point of agreement between all virtue epistemologists concerns the role of intellectual virtue in underwriting claims regarding positive epistemic status. Virtue epistemologists have, however, tended to think of the intellectual virtues themselves in different ways. It is, in fact, this difference that enables us to distinguish between two major theoretical camps in contemporary virtue

epistemology, namely, virtue reliabilism and virtue responsibilism (Baehr, 2006; Battaly, 2008; Greco, 2002). According to virtue reliabilists, the intellectual virtues are to be understood as cognitive abilities or powers. Examples include memory, perception, and various forms of reasoning (e.g., inductive and deductive reasoning). Given that we can regard these abilities or powers as cognitive faculties, virtue reliabilists are sometimes said to embrace a faculty-based conception of intellectual virtue (e.g., Greco, 2002). In contrast to virtue reliabilists, virtue responsibilists see the intellectual virtues as something akin to personality traits. Examples, in this case, include open-mindedness, tenacity, attentiveness, intellectual courage, carefulness, thoroughness, and so on. Given the sense in which these properties correspond to enduring cognitive traits, reflecting (perhaps) an agent's cognitive character,[3] virtue responsibilists are sometimes said to adopt a trait-based conception of intellectual virtue (see Figure 1).

Based on the foregoing characterization of virtue reliabilism and virtue responsibilism, it is perhaps easy to think of vice and virtue as fixed and immutable aspects of an agent's cognitive character. It is, in other words, easy to think that examples of vice (e.g., a poor memory) and virtue (e.g., open-mindedness) are *always* to be seen as vices and virtues, irrespective of the specific context in which these aspects of agent character are evaluated. This seems entirely appropriate; for why assume that something like a poor memory should, in some cases, be regarded as a vice and, in other cases, as a virtue. Similarly, it seems odd to think that a trait such as open-mindedness could ever be anything other than an intellectual virtue: surely it is always better to be open-minded as opposed to dogmatic, particularly if one's goal is to increase one's chances of believing the truth.

---

[3] The notion of cognitive character also surfaces in the context of virtue reliabilist approaches to knowledge (see Greco, 1999). In the current context, cognitive character consists of both faculty virtues and trait virtues (see Figure 1). It is important to note, however, that virtue reliabilists and virtue responsibilists are likely to view the notion of cognitive character in somewhat different ways. From a virtue reliabilist perspective, an agent's cognitive character is to be understood in terms of an "integrated web of stable and reliable belief-forming processes" (Pritchard, 2010, p. 136). For virtue responsibilists, in contrast, the notion of cognitive character is likely to refer to the dispositions of the agent to cognize in particular ways given certain situations.
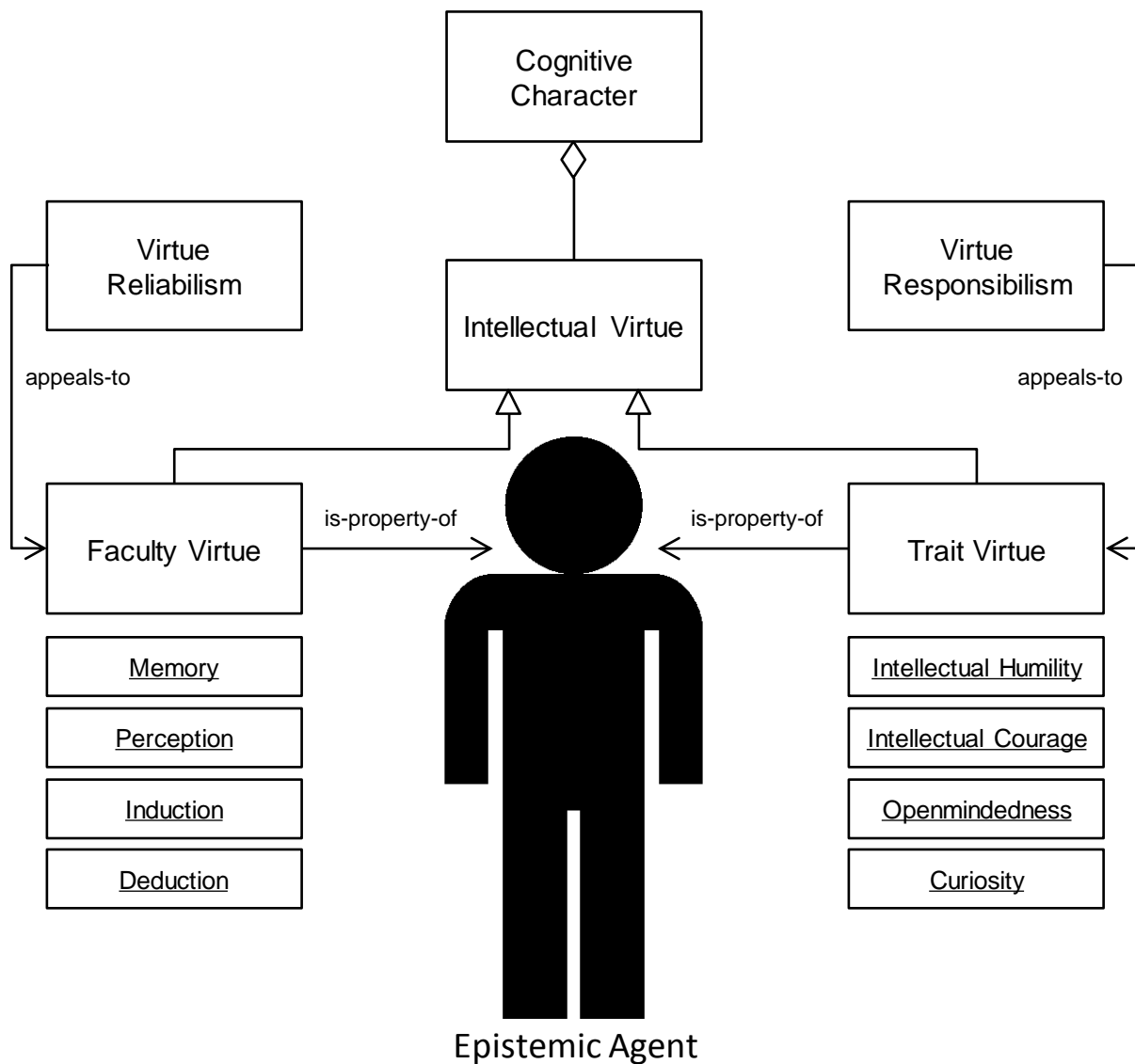
Figure 1: Virtue reliabilism and virtue responsibilism can be seen to appeal to distinct conceptions of intellectual virtue. Both kinds of intellectual virtue, in this case, are seen to form part of what is called an agent's "cognitive character." Triangles in this diagram symbolize taxonomic (or sub-type-of) relationships, while diamonds symbolize compositional (or part-of) relationships.

But now notice something important. Recall that what makes something an intellectual virtue in virtue epistemology is its reliability—the fact that it leads to a preponderance of true rather than false beliefs. Intellectual vices, on the other hand, are properties of an epistemic agent that are generally seen to be unreliable (i.e., they do not result in a preponderance of true beliefs). The reason why we regard open-mindedness as a virtue, therefore, has something to do with the fact that open-mindedness will, in general, yield a greater number of true beliefs, as opposed to a trait like dogmatism, which will yield a preponderance of false beliefs. But what if the reliability of these aspects of agent cognitive character were to vary in a systematic way with different kinds of belief-forming context? In this case, it is at least possible that a vice in one context of evaluation might have a more virtuous feel to it in another context of evaluation.

It is here that we come face-to-face with the notion of what I will call *virtue relativism*. Virtue relativism is the idea that our notions of vice and virtue are relative to specific evaluative contexts.

Our judgements of vice and virtue thus exhibit a degree of context-specificity: what we see as a vice or virtue is largely determined by the sort of context in which such judgements are made. Such issues lie at the heart of mandevillian intelligence. In the case of mandevillian intelligence, the context of evaluation (the context in which we make judgements concerning vice and virtue) relates to whether our focus is on an individual epistemic agent or a larger collection of such agents, for example, a socio-epistemic system (Goldman, 2011) or an epistemic group agent (Palermos, 2015).[4] In particular, when we apply the notion of mandevillian intelligence to virtue epistemology, the claim is that our sense of what is and what is not an intellectual vice varies according to whether our attention is focused on an individual agent (the traditional focus of analytic epistemology) or a collection of such agents (the traditional focus of social epistemology). As a result, a cognitive trait or faculty that has a distinctly vice-like feel to it at the individual level of analysis may, when studied in a more collective context, emerge as something rather more virtuous. In most cases, we can make sense of this by focusing our attention on the way in which a cognitive property (i.e., a trait or faculty) contributes to the reliability of a belief-forming process. Thus, when we consider the reliability of a cognitive trait or faculty at the individual level of analysis, we may observe that a trait such as dogmatism (i.e., an irrational adherence to one's views) is to be regarded as a cognitive vice. We may be perfectly justified in making this sort of judgement, since dogmatism may indeed undermine the ability of an individual agent to believe what is true (i.e., at the individual level dogmatism is to be regarded as a *genuine* intellectual vice). When our perspective shifts, however, and we view things from a more social perspective, we may find that individual forms of dogmatism play a productive role in yielding collective forms of epistemic success (see Section 3.2). Indeed, it may very well be the case that a degree of dogmatism on the part of individual agents is *necessary* to ensure the reliability of the larger socio-epistemic system with respect to the generation or discovery of true beliefs.[5] It is in precisely this sense that we can see the notion of mandevillian intelligence as appealing to a particular form of virtue relativism.

3. My Vice, Our Virtue

Claims regarding mandevillian intelligence would, if true, seem to have a number of implications for contemporary epistemology. Clearly, given the centrality of vice and virtue to virtue epistemology, we would expect the notion of mandevillian intelligence to inform philosophical debates concerning the nature of knowledge. This seems particularly so when it comes to virtue-theoretic conceptions of group (or collective) knowledge (Palermos and Pritchard, 2013; Palermos, 2015). In this case, a collection of individual agents is deemed to function as a form of doxastic machinery, generating belief states that (hopefully) succeed in tracking the truth. Looking at such socio-epistemic systems from the perspective of virtue epistemology, we might be inclined to view collective knowledge as grounded in the intellectual virtues of the individual agents that comprise the system. Such a view, however, seems to be called into question if we accept the notion of mandevillian intelligence (see Section 4.3).

---

[4] Epistemic group agents are defined as "groups of individuals who exist and gain knowledge in virtue of a shared common cognitive character that primarily consists of a distributed cognitive ability" (Palermos and Pritchard, 2013, p. 115).

[5] This does not mean that all individuals within the collective ensemble are required to be dogmatic. In many cases, a mixture of individuals possessing both positive and negative traits is likely to be required. Support for this claim is provided by empirical research showing that agent diversity is sometimes crucial to collective forms of cognitive success (Hong and Page, 2004; Muldoon, 2013; Page, 2007).

The implications of mandevillian intelligence may also be felt in the context of what is called applied epistemology (Chase and Coady, forthcoming). In this case, the objective is often to evaluate the epistemic impact of different forms of social or technological intervention. Inasmuch as we embrace the notion of mandevillian intelligence, it seems we may need to accept that what is epistemically good for the community (in the form of a specific social or technological intervention) may not always be so good for the epistemic standing of the community members (see Section 4.1).

There are thus a variety of reasons to think that mandevillian intelligence is of broad relevance to a number of areas in contemporary epistemology. But why should we buy into the notion of mandevillian intelligence in the first place? What evidence is there to suggest that the notion of mandevillian intelligence denotes a real phenomenon that is worthy of further epistemological consideration. In this section, I will review a number of strands of research that provide an initial response to this question.

3.1. Collective Intelligence and Collective Search

Mandevillian intelligence, recall, is defined as a specific form of collective intelligence (see Section 1). This is important, because from an empirical perspective it is common to see collective intelligence as a form of collective search through a complex space of (e.g.) doxastic possibilities (e.g., Mason, 2013). Consider, for example, the case of collective (or team) sensemaking (Klein et al, 2010; Smart and Sycara, 2013). As a team of individuals strives to make sense of some ambiguous body of information, they will each start to form beliefs about the focal object, event or situation. Such beliefs can be seen as tracing trajectories through a multi-dimensional space of interpretational possibilities, each of which can be judged in terms of its relative distance to a particular point in the search (or problem) space that represents the global optimum, i.e., the set of beliefs that corresponds to the most accurate interpretation of the target state-of-affairs. The performance of the team members, in this case, can be judged in terms of their ability to discover (and, in some cases, converge) on the global optimum. A similar characterization could be made in respect of the process of scientific discovery. Here, a community of scientists are typically engaged in a collective effort to generate ideas and evaluate them with respect to the available evidence. The hope is that as the community explores this space of ideational possibilities they will (eventually) discover regions or points within the space that mark out major advances in scientific knowledge and understanding.

Relative to this vision of collective intelligence as a form of collective search, we can view mandevillian intelligence as a form of collective intelligence that is apparent whenever individual vices play a positive functional role in supporting a collective search effort. In other words, we confront a case of mandevillian intelligence whenever individual vices contribute to the ability of a community to discover optimal solutions (e.g., doxastic states that closely approximate the nature of reality). These vices, we can assume, work to shape the flow of information and influence within an agent community in such a way as to alter the dynamics of the collective search process, enhancing the probability that the community (as a whole) will discover the best available solution. Transposing all this to the epistemic domain, we can see mandevillian intelligence as a form of collective intelligence in which individual vices enhance the reliability of a community to discover (and

sometimes converge on) those beliefs (i.e., a region of some possibly multidimensional doxastic space) that provides the community with an appropriate cognitive grip on reality.[6]

The value of thinking about mandevillian intelligence in this sort of way is that it helps to establish contact with an important body of empirical work concerning the optimal organization of multi-agent ensembles as a means of improving various aspects of the collective search process. Such work appears under a variety of headings based on the complexity (i.e., simple vs. complex) and nature (i.e., biological vs. computational) of the focal group of agents. Work involving relatively simply agents thus includes work on swarm intelligence (Bonabeau et al, 1999; Garnier et al, 2007; Kennedy et al, 2001) (biological agents) and particle swarm optimization (Poli et al, 2007) (computational agents), whereas work involving more complex agents includes work on human subjects (Mason et al, 2008) (biological agents) and cognitive architectures (Reitter and Lebiere, 2012) (computational agents).

One of the key insights to emerge from work in these areas is the importance of balancing what is called exploration and exploitation (March, 1991). In general, the agent community needs to ensure that it engages in a sufficient degree of exploration of the search space, avoiding the temptation to converge on those regions of the search space marked out by early discoveries. The problem is that, especially in complex search spaces characterized by rugged solution landscapes, precipitant forms of inter-agent communication can sometimes lead to premature convergence on sub-optimal solutions, thereby blocking the search for the global optimum. This is a particular problem in situations where early successes are of little value in terms of illuminating the path to the best outcome. It is for this reason that attempts to improve the efficiency of inter-agent communication by, for example, increasing the density of the agent communication network, tend to compromise performance on collective cognitive tasks (Hutchins, 1991; Lazer and Friedman, 2007; Mason et al, 2008; Smart et al, 2010; Zollman, 2010). The problem is that certain kinds of communication (or social) network structure (e.g., fully-connected networks) provide rapid feedback (possibly to all agents) about the best solutions on offer. This has the unfortunate consequence of causing all agents to converge on regions of the search space that are possibly far removed from the actual location of the global optimum (see Lazer and Friedman, 2007, for more details).

One way of addressing this unfortunate state-of-affairs is to reduce the opportunities that agents have to communicate information during the search process. By reducing the opportunities for communication, we alter the temporal dynamics of information flow and influence, thereby retarding the rate at which information can propagate within the agent community and reducing the possibility of premature forms of cognitive convergence. One way of restricting agent communication is, of course, to alter the physical structure of the agent communication network itself. By reducing the number of channels for inter-agent communication, changes to network structure often succeed in ensuring that a felicitous degree of cognitive (e.g., doxastic) diversity is maintained within the agent community, and it is for this reason that we can view alterations to network structure as a form of *diversity maintenance mechanism*, i.e., as a mechanism that works to maintain cognitive diversity within a community of problem-solving agents.

---

[6] In practice, of course, such forms of reliability will always be judged relative to a set of resource constraints relating to (e.g.) time, cost, money, effort and so on.

Changes to the physical structure of a communication network are not, however, the only way to alter the flow of information and influence within a community of agents. We can also look to agent-level characteristics as a means of gating and routing information through the collective ensemble. Individual agents, in this case, can be seen as working to control the time-variant "effective connectivity" of the communication network (see Friston, 2011) (i.e., the set of connections that are actively engaged in the transfer of information between the nodal elements—in this case, agents). As a means of illustrating this, imagine that we bring together a group of (human) individuals to perform a collective sensemaking task. Imagine, also, that these individuals have no prior experience of working together and are thus unknown to each other. The level of trust between the agents is, we may assume, relatively low, and each agent will, as a result, be initially reluctant to wholeheartedly embrace the information supplied by other agents (at least during the early stages of the task). The result is that the *effective* influence of communicated information, in terms of its ability to promote a shift in agent-level cognitive states, is limited. In fact, it may very well be the case that the informational influence is so low as to transform what is (at the structural level) a fully-connected network into something that, at least from the perspective of effective connectivity, looks a lot more austere (e.g., a disconnected network). It is in this sense that we can see a role for inter-agent distrust (an agent-level property) in shaping the effective connectivity of an agent communication network. Crucially, we can see this agent-mediated transformation of effective connectivity as a form of diversity maintenance mechanism that helps (just as do changes to the structural connectivity of the network) to prevent premature convergence on sub-optimal solutions.

In the case of human epistemic communities there are a variety of psycho-cognitive factors that are poised to play the same sort of functional role as that played by inter-agent (i.e., social) distrust. These include the tendency of individuals to hoard information, a lack of willingness to cooperate with others, a vulnerability to copying/transmission errors, and a steadfast (and seemingly irrational) commitment to one's own ideas and methods. What is important to note, here, is that these factors are often cast in a somewhat negative light: they are often seen as shortcomings, flaws or vulnerabilities that need to be addressed by all manner of social, pedagogical and technological interventions. It is here, then, that we begin to creep up on the notion of mandevillian intelligence. In particular, we can begin to see at least one of the ways in which a set of vice-like characteristics pitched at the level of individual agents could make a positive contribute to the cognitive processing capabilities of a collective ensemble. Relative to our intuitions regarding the importance and value of cooperation, sharing, trust and accurate information transmission, it might be all too easy to take a rather dim view of distrust, hoarding, deception and social withdrawal. What the notion of mandevillian intelligence gives us, however, is an alternative way of looking at these individual, agent-level properties: it enables us to see agent-level characteristics as playing an important (i.e., explanatorily salient) role in helping to align the structure of the collective search effort with the demands imposed by the nature of the search problem.

3.2. Dogmatism

While a failure to trust others is hardly the most celebrated of human characteristics, it is not something that is typically regarded as an intellectual vice, at least by virtue epistemologists. Intellectual vices, we have seen, are those properties of an agent's cognitive character that undermine the extent to which an agent's beliefs align themselves with facts about the world. A vice is thus something that hinders an agent's ability to believe the truth. Given that it is far from clear

that the kinds of properties we have been talking about thus far—i.e., distrust, hoarding and deception—really do undermine the truth-tracking capabilities of individual agents, there seems little reason for virtue epistemologists to be convinced about the relevance of mandevillian intelligence to epistemological theorizing. In order to make the case for mandevillian intelligence we therefore need to focus on agent-level properties that are widely recognized as intellectual vices and show how such properties can work to the good of an epistemic community.

With this in mind, consider the results of an important study by Zollman (2010). Zollman was interested in the factors that work to facilitate cognitive diversity within a scientific community, enabling the community to more fully explore a space of possibilities before converging on a particular outcome. One way of achieving such diversity is, as we have seen, to manipulate the structural configuration of the agent communication network and thereby constrain the opportunities that agents have to exchange information (and thus influence one another). But what if we are unable to change the structure of the network? In this situation it seems as though there is nothing to prevent precipitant forms of information sharing, and we might thus expect the agent community to prematurely converge on some sub-optimal part of the search space.

Although restricting the flow of information (e.g., the dissemination of research results) constitutes one means of avoiding premature convergence in this situation, Zollman identifies another way. If we are stuck with a community of well-connected scientists, diversity can still be maintained, he suggests, if (at least some of) the scientists exhibit a dogmatic adherence to their own ideas and methods. By being dogmatic, a scientist is essentially insulating themselves from the influence exerted by conflicting evidence. The result, Zollman suggests, is that the scientific community manages to maintain the sort of cognitive diversity that is required to ensure that scientific consensus, when it does finally emerge, is successful in settling on the truth.

By appealing to the role of dogmatism in maintaining a sufficient degree of cognitive diversity, Zollman (2010) is clearly identifying a positive functional role for aspects of human cognitive character that are typically seen in a somewhat negative light. From an individual perspective it seems highly unlikely that an irrational adherence to one's own ideas could be of epistemic value, especially when ideas with seemingly better evidential support are on offer. Nevertheless, such forms of individual vice can, it seems, work to the good of an epistemic community by reducing the impact of misleading evidence and sustaining the search for new ideas, methods and information. The upshot is that a moderate amount of individual vice—in this case, intellectual dogmatism— seems to play an important role in striking an effective balance between exploration and exploitation. By exhibiting a steadfast commitment to one's own ideas, individual epistemic agents (e.g., scientists) look to be behaving in a decidedly sub-optimal manner. However, such shortcomings are only really apparent at the *individual* level of analysis. At the collective level, individual vice seems to play a much more positive role in facilitating collective forms of cognitive and epistemic success.

There are, in fact, a variety of ways in which intellectual dogmatism might work to the epistemic good of a community. Most obviously, a resistance to social and informational influence helps to maintain a degree of cognitive diversity within a community of problem solvers. We can thus see dogmatism as something of an intellectual safeguard against pernicious forms of cognitive convergence and consensus, especially in situations where it is important to avoid the premature

abandonment of potentially promising lines of enquiry. Another way in which dogmatism might earn its collective cognitive keep is via the 'social scaffolding' of intellectual activities. Consider, for example, how the maintenance of marginal theories and ideas might help to concentrate and reinforce the empirical and deliberative efforts of those who embrace the majority view. Here, the presence of minority views (even if those views are incorrect!) helps to motivate others to develop their own views and make them even more convincing (see Nemeth and Wachtler, 1983). In the absence of such opposition, it is perhaps unlikely that there would be sufficient motivation to engage in these additional forms of intellectual effort.[7] Finally, note that by preserving some degree of diversity within an epistemic community, dogmatism plays a potentially important role in opening the door to further diversity. In situations where all the members of a community have settled on a particular view, it may be difficult for those with a dissenting voice to pluck up sufficient courage to challenge the status quo. Here, the presence of one or two dogmatic individuals may establish a baseline level of dissent that helps to free the timid champions of competing views from the pressures of social conformity.

Note that in many of these cases there is no reason for the dogmatist to actually believe the truth—it may very well be the case that what the dogmatist believes to be the truth is, in fact, false. This does not mean, however, that dogmatic individuals cannot (on occasion) play a useful and productive role in enabling the community as a whole to succeed in their truth-seeking efforts.[8]

3.3. Cognitive Bias

Dogmatism provides one means by which cognitive diversity could be maintained within an epistemic community. Another route to cognitive diversity may have its origins in our susceptibility to cognitive bias. Consider, for example, Solomon's (1992) analysis of the plate tectonics revolution in geology. Solomon suggests that the eventual success of theories relating to continental drift owed a great deal to a set of cognitive heuristics and biases that are often seen as a threat to rational forms of thought and action. Contrary to what might be expected, Solomon argues it was the departure from normative standards of rationality that, in fact, made possible the intellectual success of the larger geological community:

> ...during the geological revolution, the phenomena of bias and belief perseverance were responsible for much of the distribution of research commitment. This distribution facilitated the various directions of research effort that went into the geological revolution. So the cognitive heuristics of availability, salience and representativeness had desirable epistemic effects in the geological community. (Solomon, 1992, p. 452)

Crucially, for our purposes, the kinds of factors that Solomon draws attention to in accounting for the success of the geological community are factors that are typically denigrated as forms of

---

[7] Something along these lines was recognized by Karl Popper (1994). With respect to dogmatism, Popper (1994) notes that: "...a limited amount of dogmatism is necessary for progress. Without a serious struggle for survival in which the old theories are tenaciously defended, none of the competing theories can show their mettle—that is, their explanatory power and their truth content" (p. 16).

[8] Note how all this this encourages us to embrace a form of *epistemic tolerance* with respect to those who we would otherwise seek to censure, denounce or condemn (Robert Clowes, personal communication, October 7, 2016). Inasmuch as we accept the possibility the mandevillian intelligence, we are perhaps able to see the (collective) epistemic value of those who are widely recognized as the purveyors of false beliefs.

intellectual vice, or at least as factors that often work to undermine the epistemic integrity of individual agents. Here, then, we encounter an additional reason to think that the notion of mandevillian intelligence has a substantive impact on contemporary epistemological debates relating to both individual analytic epistemology (most notably, virtue epistemology) and social epistemology. Whereas our earlier discussion of dogmatism focused on a form of vice that lies in the philosophical cross-hairs of virtue responsibilist theorizing, we are now presented with a vice that seems to be of greater interest and relevance to virtue reliabilism. In particular, a susceptibility to cognitive bias seems to represent something of a failing with regard to the proper functioning of our cognitive machinery—it is, at least, something that seems to pose a threat to our cognitively-grounded truth-tracking capabilities. The value of Solomon's contribution, in this respect, is that it helps to highlight the explanatory significance of degraded forms of individual cognitive ability in underwriting the epistemic accomplishments of a larger epistemic community.

Another significant contribution to this debate comes from a consideration of a specific kind of cognitive bias, namely, confirmation bias (Nickerson, 1998). Confirmation bias is almost universally seen as a cognitive vice, and this is reflected in the many (ongoing) attempts to develop technological interventions that aim to minimize or eliminate the bias (e.g., Convertino et al, 2008). A more positive appraisal of confirmation bias is, however, possible if we see confirmation bias as (again) a form of diversity maintenance mechanism. Thus while confirmation bias might appear (genuinely so) as an individual shortcoming (an individual vice), the bias can also, at least in some circumstances, work to ensure the optimal distribution of available cognitive resources within a community of epistemic agents. In order to help us understand this, imagine that we have a group of individuals who are trying to make sense of some complex state-of-affairs. Providing that we have a sufficient degree of cognitive diversity (represented as differences in beliefs, opinions, theories, approaches, or whatever) between the agents at the outset of the problem-solving endeavour, then we can assume that each agent will (under the influence of confirmation bias) attempt to marshal support for their own argumentative positions (see Mercier and Sperber, 2011). The result is that each agent, operating under the influence of confirmation bias, will devote considerable effort to processing a specific subset of the information (e.g., empirical data) that is available within the relevant problem space. This helps to promote the distribution of collective cognitive resources with respect to the target problem. Given that cognitive assets (as well as other resources such as time) are invariably limited, it helps if each individual agent engages in a thorough exploration of a specific and distinct part of the problem space rather than have *all* agents attempt to cognitively engage with the entire space. Not only is this latter strategy likely to be infeasible, it is also likely to be highly ineffective. This is because each agent will only be able to process information at a rather superficial level. By contrast, when each agent is attempting to bolster support for their own argumentative position or preferred interpretation, they will be encouraged to restrict their attention to a much more limited body of information (specifically, that body of information that is consistent with their own particular point of view) and process this information to a much greater depth than would otherwise be the case.

When we look at this state-of-affairs from an epistemological standpoint, the behaviour of the individual agents looks far from optimal. This is particularly so when we look at the situation through the lens of virtue epistemology. In this case, the biased nature of each of the individual agents seems to undermine their status as genuine knowers. A susceptibility to confirmation bias thus seems to undermine the reliability of the individual agent in terms of their ability to track the truth, and it is

therefore difficult, under a virtue-theoretic conception of knowledge, to see the individual agent as the appropriate target of knowledge ascriptions. Note that even if a particular agent, in this situation, turns out to believe the truth, it is far from clear that a virtue reliabilist conception of knowledge will allow us to make any sort of positive statement about the epistemic standing of the agent. In fact, given the nature of the belief forming processes employed by the agents, it seems more a question of luck (as opposed to cognitive ability) that any individual agent will hit on the correct answer.[9] Individual epistemic successes, if there are any, will thus be seen to violate an epistemic anti-luck condition (Pritchard, 2009, 2012), and this is almost universally regarded as inimical to knowledge attribution. The result is that we seem to confront a curious state-of-affairs in which a susceptibility to cognitive bias undermines positive epistemic standing in the case of individual agents, but the same susceptibility to bias underwrites the intellectually virtuous character of the larger, multi-agent ensemble. It is after all plausible that the reliability of the agent community (i.e., the ability of the community to discover the truth) is rooted in the fact that the relevant search space has been explored to the greatest extent possible, and such forms of exploratory effort seem to be best enabled by ensuring an appropriate distribution of individual-level cognitive resources. We might therefore judge whether a community can be said to know something on the basis of whether the community's doxastic outputs stem from the exercise of a form of collective cognitive ability—a form of ability that, in this case, appears to be grounded in the biased nature of the individual members' belief-forming processes.

3.4. Impure Motives

In addition to cognitive traits (e.g., intellectual dogmatism) and cognitive faculties (e.g., cognitive bias), agent motivations have sometimes been shown to yield mandevillian-like effects. Perhaps the clearest demonstration of this is apparent in the philosophy of science literature. Here, there has been a concerted effort to understand the ways in which socio-cultural factors, such as economic rewards and credit assignment policies, influence decisions regarding (e.g.) what scientific topics are worth investigating. A particular focus of attention has been the factors that prevent scientists from converging on the same area of investigation. This has been deemed to constitute a problem, since if we assume that all scientists are perfectly rational and care only about the epistemic payoffs of their endeavours (which seems like the virtuous thing do), then there seems little reason to believe that scientists will choose to distribute their scientific efforts across a range of different focus areas. The result, as noted by Thagard (1993), is a potentially ruinous retreat from novelty:

> ...if all scientists made identical judgements about the quality of available theories and the value of possible research programs, science would become homogenous. Novel ideas and potentially acceptable new theories would never be developed sufficiently to the point where they would in fact become rationally acceptable by all. (p. 65)

In order to achieve an effective division of labour it therefore seems that the interests of scientists need to be guided by more than just epistemic criteria. It is here that an appeal to sullied motives comes into play. For if we reject the image of scientists as epistemically pure-minded individuals and instead adopt the (surely far more plausible) image of scientists as preoccupied with a range of non-

---

[9] Much, of course, will depend on where an agent happens to focus their intellectual efforts at the outset of the search process; however, this may be down to luck as much as anything else.

epistemic concerns (e.g., economic rewards, academic prestige and social status), then it seems natural to assume that scientists will avoid attempting to hit epistemic pay dirt in those parts of the intellectual landscape that are already the focus of significant scientific mining operations. It is this general idea that underlies claims about the value of sullied motives in enabling a community to achieve its epistemic ends (see Kitcher, 1993). Weisberg and Muldoon (2009) thus suggest that:

> ...scientist's micromotives can look epistemically impure or shortsighted, yet these motives can actually help the community as a whole make rapid progress toward finding out the truth. Thus a core tenant [*sic.*] of strategic models about the division of cognitive labor is that what is epistemically good for individuals may differ from what is epistemically good for the community. (pp. 226–227)

There is clearly something of a mandevillian ring to all this. But where should we seek to locate research on agent motivations within the theoretical framework of virtue epistemology? The best answer to that question, to my mind, is to see work on agent motivation (and perhaps emotion[10]) as appealing to a virtue responsibilist conception of knowledge. The reason for this, I suggest, is that agent motivations are better conceived as a form of cognitive trait rather than a form of cognitive faculty. Motivations are, at least, closer to the sort of characteristics targeted by virtue responsibilists than the sort of characteristics targeted by virtue reliabilists. Indeed, virtue responsibilists like Zagzebski (1996) often appeal to the role of agent motivations in guiding evaluations of epistemic standing. For Zagzebski (1996), all intellectual virtues are to be understood as acquired traits of character that involve both a motivational component and a reliability component:

> An act of intellectual virtue *A* is an act that arises from the motivational component of *A*, is something a person with virtue *A* would (probably) do in the circumstances, is successful in achieving the end of the *A* motivation, and is such that the agent acquires a true belief...through these features of the act. (p. 270)

There is, however, a potentially significant difference between agent motivations and agent cognitive traits. Motivations, I suggest, are the sorts of things that are deeply affected by a complex nexus of social, cultural, economic, legislative and organizational influences. It is thus relatively easy to see how the motivations of individual scientists could be shaped and sustained as a result of the delicate interplay between socio-cultural factors, the structure of economic rewards, and the ethical norms and values associated with scientific practice. It is here that issues of mandevillian intelligence start to dovetail with issues of social policy and the socio-economic scaffolding of epistemic virtue. For inasmuch as we accept the idea that individual vice can, on occasion, play a productive role in serving the collective epistemic good, then we will need to consider to what extent socio-economic and socio-political interventions should be judged *solely* on the basis of their potential to exacerbate individual forms of intellectual vice. It seems that, at least in some cases, an intervention might earn its keep solely on the basis of performance outcomes that are *only* visible at the collective or social level of analysis. Inasmuch as we accept this to be the case, we should perhaps ask ourselves whether certain kinds of socio-economic, socio-political, and (to complete the picture) socio-technical interventions should be adopted on the grounds that they trade collective forms of epistemic success for individual forms of intellectual vice.

---

[10] Morton (2014) discusses a range of epistemic emotions within a broadly mandevillian framework.

4. Implications

4.1. The Evaluation of Epistemic Technology

In considering the way in which vice and virtue impact aspects of collective performance, it is natural for our attention to settle on the properties of individual human agents, for example, their cognitive capabilities. In an epistemological context, this focus is not surprising; for it is precisely these sorts of properties (i.e., the elements of cognitive character) that are the primary target of virtue epistemological theorizing. In many cases, however, epistemic goods (e.g., knowledge, true belief and understanding) are the product of processes that involve some form of technological resource. Palermos (2011) thus discusses the way in which the production of astronomical knowledge is tied to the skilful use of a variety of technological artefacts (e.g., telescopes). Our evaluations of positive epistemic status, in this case, are grounded, Palermos argues, in the way in which scientific instruments are used to drive belief-forming processes. In attributing a scientist with knowledge, we thus appeal to the explanatory role of a cognitive ability that (in most real-world cases) involves the technological resource as an intrinsic part of an extended cognitive process (see Palermos, 2011, 2015). Such forms of epistemically-relevant bio-technological coupling are, of course, a common feature of contemporary scientific practice. In the absence of a proper grasp of how to use an array of increasingly sophisticated technological artefacts and systems, it is unlikely that any scientist would be able to reliably track the truth in their respective domain of interest.

The upshot of all this is to highlight the importance of technologies relative to our individual and collective doxastic endeavours. In view of this importance, it seems appropriate to subject technologies to a degree of epistemic evaluation; i.e., it seems that we ought to judge the merits of a particular technology based on the extent to which it enhances the epistemic standing of its users. The problem, of course, is that with the notion of mandevillian intelligence now at hand, it seems entirely possible that the epistemic impact of a technology may differ depending on whether our attention is focused on the individual or the collective level of analysis. We can thus imagine a case in which a technology works against the epistemic interests of the individual but nevertheless confers an epistemic advantage to the larger community.

To help make this a little clearer, consider the case of personalized search, as supported by popular Internet search engines such as Google Search. Personalized search helps to tailor search results to the interests and concerns of a particular Internet user based on their previous history of search engine use. The search engine is, in this case, acting as a form of bio-external attentional filter, directing a user's attention to those regions of the online environment that are deemed to be of greatest interest and relevance. The danger, of course, is that this mechanism works to selectively alter the accessibility of online information in a manner that is aligned with a set of pre-existing interests, beliefs, and (perhaps) prejudices. As a result, we encounter the notorious problem of "filter bubbles" (Pariser, 2011). These undoubtedly limit a user's access to epistemically-relevant information, but they do so in such a way that (at least in the worst case) ignores a range of important epistemic desiderata, such as reliability, objectivity, credibility, scope and truth (see Miller and Record, 2013). The concern, then, is that personalized search exerts something of a pernicious influence on our epistemic standing. It is for this reason that epistemologists such as Simpson (2012) and Miller and Record (2013) advocate changes in user behaviour and corporate policy, potentially backed up by government intervention, as a means of reducing the negative impact of personalized search engine technology on our (individual) epistemic capabilities.

It is here that the notion of mandevillian intelligence helps to reshape (or at least refocus) the nature of the epistemological debate. For even if we accept that personalized search is, in general, injurious to an individual's epistemic health, this does not mean that the technology is bereft of *any* kind of epistemic benefit. In particular, it is far from clear that the epistemic consequences of search engine technology for a community of Internet users is exactly the same as that for the individual members of the community itself. We might, for example, claim that the use of personalized search plays a productive role in maintaining a degree of cognitive diversity within a community of Internet users. Such diversity, we have seen, is often crucial in terms of enabling some collective, multi-agent ensemble to discover, resolve or otherwise track the truth in some domain of interest. In essence, what the notion of mandevillian intelligence gives us is a means of avoiding a rush towards premature judgements concerning the epistemic or veritistic (see Goldman, 1999, 2002) value of a technology. Just because a particular technology, such as personalized search, turns out to have little or no benefit at the level of individual agents, this does not mean that it has no benefits *tout court*.[11] The concept of mandevillian intelligence can thus be seen as a means of helping epistemologists remain open-minded[12] about the epistemic significance of a technology and the kinds of interventions (e.g., government regulation) that might be required to ensure its 'virtuous' operation.

4.2.   The Engineering of Socio-Epistemic Systems

Inasmuch as we accept the notion of mandevillian intelligence, it seems that the epistemic consequences of a particular technology or agent property may vary according to whether our attention is focused at the individual or collective level. In particular, mandevillian intelligence suggests that a technology (e.g., personal filtering) or agent property (e.g., dogmatism) that looks to harm the epistemic standing of an individual agent could also play a positive role in securing epistemic benefits at the level of social groups and socio-technical systems. This insight has important implications for how we evaluate the properties of agents and the technologies they use. It also, however, alters our view as to how we might engineer socio-technical systems so as to maximize their epistemic potential. In particular, if we accept the claim that individual cognitive vices can (on occasion) underwrite collective epistemic accomplishments, then it becomes possible to entertain the (otherwise outlandish) idea of a technology being used to exploit, accentuate, or even establish(!) an individual form of intellectual vice. This, of course, is largely counter to the way we think about the technology design process. In most cases, the aim of technology design is to enhance the cognitive and epistemic functioning of the individual by addressing cognitive limitations and reducing the incidence of cognitive bias. In fact, it seems odd (not to mention ethically questionable) to consider the prospect of designing a technology so as to capitalize on the availability of a form of individual cognitive vice: in general, it seems much more appropriate to focus one's efforts on minimizing cognitive bias rather than exploiting, accentuating or even creating

---

[11] This issue is of particular importance when one considers the criticisms levelled at major technology providers (e.g., Facebook and Google) by a number of political leaders (http://www.bbc.co.uk/news/technology-37798762). The main value of mandevillian intelligence, in this respect, is that helps us see the current political debate as epistemologically deficient. In particular, it is unclear whether the criticisms of major technology vendors are justified in the absence of a clearer understanding of how personalized search mechanisms affect epistemic outcomes at *both* the individual and collective levels.

[12] The concept of mandevillian intelligence is, in this context, functioning as a form of 'cognitive scaffold' that helps to support a specific form of (meta-epistemological) intellectual virtue: one that is often deemed relevant to issues of positive epistemic standing in the case of individual agents (see Zagzebski, 1996).

it. But what if we were to encounter a form of cognitive bias that, while injurious to the epistemic interests of the individual, was nevertheless able to play a productive role in ensuring epistemic success at the collective level. In this case, we can surely ask ourselves to what extent our technology design efforts should attempt to exploit the bias, as opposed to always attempting to eradicate it.

What all this amounts to, I suggest, is an approach to technology development that factors in the potential contribution of individual cognitive biases and processing limitations as a means of promoting, supporting, or enabling collective forms of cognitive and epistemic success. The general idea is that we should see the cognitive vices of the individual as a form of exploitable resource that can work to the overall cognitive and epistemic good of an agent community. It is in this sense that we can perhaps see individual cognitive vices as part of a suite of resources (a sort of 'vice-kit') that system developers can exploit as part of their efforts to engineer socio-epistemic systems.

4.3.   The Virtue-Theoretic Conception of Collective Knowledge

Finally, we should ask ourselves to what extent the notion of mandevillian intelligence threatens to undermine or at least destabilize virtue-theoretic approaches to understanding knowledge. Inasmuch as we view vice and virtue as apt to change whenever we switch our epistemological gaze between the individual and the collective levels, does this call into question the explanatory appeal of intellectual virtue in accounting for the nature of knowledge? Such may indeed be the case if we accept that collective forms of knowledge are produced by the exercise of an individual's 'vicious' intellectual character. If, for example, the true beliefs ascribed to an epistemic group agent are seen to result from the exercise of what, at the individual level, is regarded as an intellectual vice, then how can virtue epistemology claim to provide a *generic* account of knowledge, one that applies as much to collective forms of knowledge as it does to the sorts of knowledge ascribed to individual agents?

In order to help us appreciate what is at stake, here, consider how we might seek to develop a virtue theoretic conception of so-called collective doxastic agents (Goldman, 2011) or epistemic group agents (Palermos and Pritchard, 2013; Palermos, 2015). In general, analyses of intellectual virtue have focused on individual agents and paid little attention to how notions of intellectual virtue might be applied to groups of individuals. An exception is an important body of work by Palermos and Pritchard (2013). They suggest that we can understand attributions of group-level knowledge (i.e., the claim that a group of agents, such as a scientific team, knows that *p*) from a virtue theoretic perspective. In particular, they suggest that we should see group-level or collective knowledge as resulting from the exercise of cognitive abilities that are ascribed to the group as a whole:

> ...the collective cognitive success of believing the truth of some (scientific) proposition will...be significantly creditable to the group's cognitive agency (i.e., the set of organismic cognitive faculties of its individual members). (p. 115)

It is here that we encounter a potential problem. This is revealed by the fact that what Palermos and Pritchard (2013) are seeking to do is use an epistemic condition that was formulated in the context of individualistic analytic virtue epistemology in order to pin down the epistemic standing of a collective epistemic entity. The danger, of course, is that it is all too easy to fall into the trap of assuming that individual forms of intellectual virtue are necessary to account for collective knowledge. In the case of mandevillian intelligence, recall, it is individual vice that lies at the root of

collective forms of epistemic success. While Palermos and Pritchard stop short of claiming that the cognitive faculties of the group members need be of the virtuous variety, it is clearly tempting, given their appeal to a virtue-theoretic epistemic condition, to see the epistemic status of groups as grounded in the *same sort* of cognitive abilities (i.e., intellectual virtues) as those that determine the validity of epistemic attributions made in the case of individual agents. It should, however, be clear that this kind of claim is called into question by the notion of mandevillian intelligence. In particular, we should not assume that *collective* forms of intellectual virtue are necessarily rooted in the exercise of *individual* forms of intellectual virtue. Instead, the thing that makes some collective cognitive trait or cognitive faculty a virtue (i.e., the thing that makes a collective cognitive trait or faculty reliable) may very well turn out to be a cognitive trait or faculty that, at the individual level, has a much more vice-like feel to it.

The extent to which all of this presents a problem for virtue epistemology is, for the time being at least, unclear. Inasmuch as epistemologists accept the notion of mandevillian intelligence, it seems they should avoid committing to the idea that collective knowledge is founded on individual forms of intellectual virtue. Even if this sort of claim has not (as yet) been made by the epistemological community, it is perhaps important to be aware of some of the theoretical pitfalls that lie on the philosophical road ahead.

5. Conclusion

Mandevillian intelligence is a form of collective intelligence in which individual cognitive shortcomings, limitations and biases play a positive functional role in yielding various forms of collective cognitive success. Transposed to the epistemological domain, mandevillian intelligence emerges as the idea that individual forms of intellectual vice can, on occasion, play a productive role in improving the epistemic performance of a multi-agent ensemble (e.g., a socio-epistemic system, a collective doxastic agent or an epistemic group agent). This, it should be clear, marks out an important area of the epistemological terrain that is of common interest to those working in virtue epistemology (Greco and Turri, 2012), social epistemology (Goldman and Whitcomb, 2011) and applied epistemology (Chase and Coady, forthcoming). In particular, the notion of mandevillian intelligence forces us to question the extent to which individual cognitive shortcomings should always be denigrated as forms of intellectual vice. While such shortcomings may work against the epistemic interests of the individual agent, they may also play a crucial role in enabling an epistemic community to track the truth in a reliable manner. This has potentially profound implications for the way we think about the epistemic value of a broad array of social and technological interventions. For example, it is now unclear whether we should view a technology (e.g., personalized search) as of limited epistemic value simply because it undermines the epistemic standing of individual agents: in some cases, it seems that a specific technology may work to the collective epistemic good of a community while simultaneously jeopardizing the epistemic standing of the community's constituent members. This sort of conflict, with epistemic enhancement at one level implying epistemic diminishment at the other, may very well serve as a point of tension when it comes to the theoretical merger of virtue epistemology with social epistemology (see Palermos and Pritchard, 2013). It also, no doubt, serves as a potent form of dialectical fodder for those who are concerned with the application of epistemological theorizing to real-world situations (see Chase and Coady, forthcoming).

All of this should help to convince us of the relevance of mandevillian intelligence to contemporary epistemology. Clearly, much more work needs to be done to evaluate the notion of mandevillian intelligence. It will, in particular, be important to develop a better understanding of the sorts of situations in which individual vice contributes to collective virtue. It will also be important to undertake studies that reveal something about the nature of the processes involved in translating individual vice into collective virtue. Even in the absence of these studies, however, the notion of mandevillian intelligence can play an important role in highlighting a host of issues that lie at the interface of the epistemological sub-disciplines of virtue, social and applied epistemology. It is for this reason, I suggest, that the notion of mandevillian intelligence is worthy of further consideration by the epistemological community.

Acknowledgements

References

Baehr J (2006) Character, reliability and virtue epistemology. *The Philosophical Quarterly* 56(223):193–212

Battaly H (2008) Virtue epistemology. *Philosophy Compass* 3(4):639–663

Bonabeau E, Dorigo M, Theraulaz G (1999) *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, New York, New York, USA

Chase J, Coady D (forthcoming) *The Routledge Handbook of Applied Epistemology*. Routledge, New York, New York, USA

Convertino G, Billman D, Pirolli P, Massar J, Shrager J (2008) The CACHE study: Group effects in computer-supported collaborative analysis. *Computer Supported Cooperative Work* 17(4):353–393

Fallis D (2011) Wikipistemology. In: Goldman AI, Whitcomb D (eds) *Social Epistemology: Essential Readings*, Oxford University Press, New York, New York, USA

Friston KJ (2011) Functional and effective connectivity: A review. *Brain Connectivity* 1(1):13–36

Garnier S, Gautrais J, Theraulaz G (2007) The biological principles of swarm intelligence. *Swarm Intelligence* 1(1):3–31

Goldman AI (1999) *Knowledge in a Social World*. Oxford University Press, Oxford, UK

Goldman AI (2002) Précis of knowledge in a social world. *Philosophy and Phenomenological Research* 64(1):185–190

Goldman AI (2011) A guide to social epistemology. In: Goldman AI, Whitcomb D (eds) *Social Epistemology: Essential Readings*, Oxford University Press, New York, New York, USA

Goldman AI, Whitcomb D (eds) (2011) *Social Epistemology: Essential Readings*. Oxford University Press, New York, New York, USA

Greco J (1999) Agent reliabilism. *Noûs* 33(s13):273–296

Greco J (2002) Virtues in epistemology. In: Moser PK (ed) *The Oxford Handbook of Epistemology*, Oxford University Press, New York, New York, USA

Greco J, Turri J (eds) (2012) *Virtue Epistemology: Contemporary Readings*. MIT Press, Cambridge, Massachusetts, USA

Haddock A, Millar A, Pritchard D (eds) (2010) *Social Epistemology*. Oxford University Press, Oxford, UK

Hong L, Page SE (2004) Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences* 101(46):16,385–16,389

Hutchins E (1991) The social organization of distributed cognition. In: Resnick L, Levine J, Teasley S (eds) *Perspectives on Socially Shared Cognition*, The American Psychological Association, Washington DC, USA

Hutchins E (1995) *Cognition in the Wild*. MIT Press, Cambridge, Massachusetts, USA

Kearns M (2012) Experiments in social computation. *Communications of the ACM* 55(10):56–67

Kennedy J, Eberhart RC, Shi Y (2001) *Swarm Intelligence*. Morgan Kaufmann Publishers, San Francisco, California, USA

Kerr NL, Tindale RS (2004) Group performance and decision making. *Annual Review of Psychology* 55:623–655

Khatib F, Cooper S, Tyka MD, Xu K, Makedon I, Popović Z, Baker D, Foldit Players (2011) Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences* 108(47):18,949–18,953

Kitcher P (1993) *The Advancement of Science*. Oxford University Press, Oxford, UK

Klein G, Wiggins S, Dominguez CO (2010) Team sensemaking. *Theoretical Issues in Ergonomics Science* 11(4):304–320

Lazer D, Friedman A (2007) The network structure of exploration and exploitation. *Administrative Science Quarterly* 52(4):667–694

Lintott CJ, Schawinski K, Slosar A, Land K, Bamford S, Thomas D, Raddick MJ, Nichol RC, Szalay A, Andreescu D, Murray P, van den Berg J (2008) Galaxy Zoo: Morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society* 389(3):1179–1189

March JG (1991) Exploration and exploitation in organizational learning. *Organization Science* 2(1):71–87

Mason W (2013) Collective search as human computation. In: Michelucci P (ed) *Handbook of Human Computation*, Springer, New York, New York, USA

Mason WA, Jones A, Goldstone RL (2008) Propagation of innovations in networked groups. *Journal of Experimental Psychology: General* 137(3):422– 433

Mercier H, Sperber D (2011) Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences* 34(2):57–111

Miller B, Record I (2013) Justified belief in a digital age: On the epistemic implications of secret Internet technologies. *Episteme* 10(02):117–134

Morton A (2014) Shared knowledge from individual vice: The role of unworthy epistemic emotions. *Philosophical Inquiries* 2(2):163–172

Muldoon R (2013) Diversity and the division of cognitive labor. *Philosophy Compass* 8(2):117–125

Nemeth CJ, Wachtler J (1983) Creative problem solving as a result of majority vs minority influence. *European Journal of Social Psychology* 13(1):45–55

Nickerson R (1998) Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* 2(2):175–220

Page SE (2007) *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools and Societies*. Princeton University Press, Princeton, New Jersey, USA

Palermos O, Pritchard D (2013) Extended knowledge and social epistemology. *Social Epistemology Review and Reply Collective* 2(8):105–120

Palermos SO (2011) Belief-forming processes, extended. *Review of Philosophy and Psychology* 2(4):741–765

Palermos SO (2015) Active externalism, virtue reliabilism and scientific knowledge. *Synthese* 192(9):2955–2986

Pariser E (2011) *The Filter Bubble: What the Internet is Hiding from You*. Penguin, London, England, UK

Poli R, Kennedy J, Blackwell T (2007) Particle swarm optimization. *Swarm Intelligence* 1(1):33–57

Popper K (1994) *The Myth of the Framework: In Defence of Science and Rationality*. Routledge, Abingdon, Oxon, UK

Pritchard D (2009) *Knowledge*. Palgrave Macmillan, Basingstoke, England, UK

Pritchard D (2010) Cognitive ability and the extended cognition thesis. *Synthese* 175(1):133–151

Pritchard D (2012) Anti-luck virtue epistemology. *The Journal of Philosophy* 109(3):247–279

Reitter D, Lebiere C (2012) Social cognition: Memory decay and adaptive information filtering for robust information maintenance. In: *26th AAAI Conference on Artificial Intelligence*, Toronto, Canada

Simpson TW (2012) Evaluating Google as an epistemic tool. *Metaphilosophy* 43(4):426–445

Smart PR, Sycara K (2013) Collective sensemaking and military coalitions. *Intelligent Systems* 28(1):50–56

Smart PR, Huynh TD, Braines D, Shadbolt NR (2010) Dynamic networks and distributed problem-solving. In: *Knowledge Systems for Coalition Operations (KSCO'10)*, Vancouver, British Columbia, Canada

Solomon M (1992) Scientific rationality and human reasoning. *Philosophy of Science* 59(3):439–455

Thagard P (1993) Societies of minds: Science as distributed computing. *Studies in History and Philosophy of Science Part A* 24(1):49–67

Weisberg M, Muldoon R (2009) Epistemic landscapes and the division of cognitive labor. *Philosophy of Science* 76(2):225–252

Zagzebski LT (1996) *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge University Press, Cambridge, UK

Zollman KJ (2010) The epistemic benefit of transient diversity. *Erkenntnis* 72(1):17–35