This is a preprint of a paper published in Episteme 2(1): 25-38 (2005).

Group knowledge and group rationality: a judgment aggregation perspective

Christian List[1]

In this paper, I introduce the emerging theory of judgment aggregation as a framework for studying institutional design in social epistemology. When a group or collective organization is given an epistemic task, its performance may depend on its 'aggregation procedure', i.e. its mechanism for aggregating the group members' individual beliefs or judgments into corresponding collective beliefs or judgments endorsed by the group as a whole. I argue that a group's aggregation procedure plays an important role in determining whether the group can meet two challenges: the 'rationality challenge' and the 'knowledge challenge'. The rationality challenge arises when a group is required to endorse consistent beliefs or judgments; the knowledge challenge arises when the group's beliefs or judgments are required to track certain truths. My discussion seeks to identify those properties of an aggregation procedure that affect a group's success at meeting each of the two challenges.

1. Introduction

Institutional design has received much attention in the social sciences. Many different institutional structures of societies, organizations or social groups have been investigated with respect to their effects on social decision making. Examples of such institutional structures are constitutions, electoral systems, legislative and judicial procedures, forms of government and other organizational forms. A widely accepted conclusion is that institutions matter. Different institutional structures may lead to different social outcomes even if everything else remains fixed. Some institutional structures may lead to more optimal, stable or rational outcomes than others (for an overview, see Goodin 1996).

Questions about institutional design arise in social epistemology too. Many epistemic tasks are performed not by individuals, but by multi-member groups such as expert panels, committees and organizations. How is the epistemic performance of such groups affected by their institutional structure? The failure of the US intelligence services to draw certain inferences from available information before 9/11, for example, has often been attributed to flaws in their institutional structure, and various institutional reforms have been proposed in response to 9/11 (Goldman 2004). Some institutional structures may facilitate the integration of information held by different individuals, others not.

In this paper, I suggest a formal approach to thinking about institutions in social epistemology, drawing on the newly emerging theory of judgment aggregation. I argue that institutions matter here too. I focus on particular institutional structures that affect a group's epistemic performance: 'aggregation procedures', as defined in the theory of judgment aggregation (e.g. List and Pettit 2002, 2004; Pauly and van Hees 2005; Dietrich 2005; List 2005a,b). Aggregation procedures are mechanisms a multi-member group can use to combine ('aggregate') the individual beliefs or judgments held by the group members into collective beliefs or judgments endorsed by the group as a whole.

---

[1] C. List, Department of Government, London School of Economics and Political Science, London WC2A 2AE, UK, c.list@lse.ac.uk. I am grateful to Alvin Goldman, Leslie Marsh, Philip Pettit, Miriam Solomon and the participants of the 2005 EPISTEME Conference at the University of Massachusetts at Amherst for helpful comments and discussion.

I argue that in designing an aggregation procedure for a group, we are faced with two challenges. Inspired by Goldman (2004), I call these the 'rationality challenge' and the 'knowledge challenge'. The rationality challenge arises when the group's collectively endorsed beliefs or judgments have to be consistent. The knowledge challenge arises when those beliefs or judgments have to track certain truths.

But while Goldman has associated these challenges with two different approaches to social epistemology, I argue that they can be studied within a single approach, namely within the theory of judgment aggregation. I argue that whether a group can meet each of the two challenges depends on the group's aggregation procedure, and I investigate the ways in which aggregation procedures matter.

The paper is structured as follows. I begin with some introductory remarks about social epistemology in section 2 and introduce the concept of an aggregation procedure in section 3. The core of my discussion consists of sections 4 and 5, in which I address the rationality and knowledge challenges, respectively. In section 6, I draw some conclusions.

2. Epistemology: individual and social

Epistemology is the study of the processes by which beliefs and knowledge are acquired and justified. In traditional epistemology, the agents acquiring beliefs or knowledge are individuals, and the relevant processes usually involve only a single individual. Examples of such processes are perception, memory or reasoning (Goldman 2004).

Social epistemology comes in less and more radical forms. In social epistemology of the less radical form, the epistemic agents are still individuals, but the focus is on processes of belief or knowledge acquisition involving social interaction. Examples of such processes are testimony, discourses and information transmission in social networks (Goldman 1999). Social epistemology of this form is an extension of traditional epistemology, distinguished primarily by its recognition that individuals often acquire their beliefs or knowledge not in isolation, but in interaction with others.

In social epistemology of the more radical form, by contrast, certain multi-member groups themselves are taken to be epistemic agents capable of acquiring beliefs or knowledge. As Goldman (2004, p. 12) has noted, "[i]n common parlance … organizations are treated as subjects for knowledge attribution", such as in discussions about what the FBI did or did not know before 9/11. My discussion in this paper concerns social epistemology of this more radical form.

To pursue social epistemology of this form, one has to be prepared to consider groups as epistemic agents over and above their individual members. Many philosophers and individualistically minded social scientists are reluctant to treat groups as agents on a par with individuals. Others may be prepared to treat certain groups as agents, provided some stringent conditions are met (Rovane 1998; Pettit 2003; List and Pettit 2005a,b). In particular, to be an agent, a group must exhibit patterns of behaviour vis-à-vis the outside world that robustly satisfy certain rationality conditions. Many groups fail to exhibit such rational integration in their behaviour. For example, a group of people who happen to be at London's Leicester Square at the same time lacks the required level of integration. On the other hand, a well organized committee or organization with clearly established decision-making procedures might well qualify as sufficiently integrated.

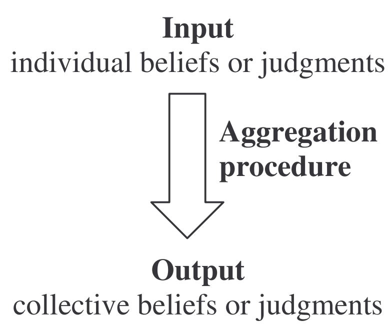
Here I set aside the broader question of whether groups can be fully fledged agents, and focus instead on the narrower question of how they perform as *epistemic* agents, i.e. how they perform at acquiring beliefs or knowledge. Of course, not all groups are capable of forming collectively endorsed beliefs, let alone knowledge. Whether or not they are capable of forming such beliefs depends on their (formal or informal) institutional structure. An example of a group incapable of forming collective beliefs is once again the random crowd at Leicester Square. But if a group's institutional structure allows the group to make certain public declarations, then that group may well count as an epistemic agent capable of acquiring beliefs or even knowledge. An example might be an expert panel or research group that publishes a joint report on some scientific matter, the monetary policy committee of a central bank that makes an economic forecast, or a court that publicly announces its factual judgments relevant to some case.

In short, a necessary condition for epistemic agency in a group is an institutional structure (formal or informal) that allows the group to endorse certain beliefs or judgments as collective ones; and the group's performance as an epistemic agent depends on the details of that institutional structure.

3. The concept of an aggregation procedure

How can we think about a group's institutional structure? Let me introduce the concept of an 'aggregation procedure' to represent (a key part of) a group's institutional structure. As defined in the theory of judgment aggregation (List and Pettit 2002, 2004; List 2005a), an aggregation procedure is a mechanism by which a group can generate collectively endorsed beliefs or judgments on the basis of the group members' individual beliefs or judgments (illustrated in table 1). A simple example is '(propositionwise) majority voting', whereby a group judges a given proposition to be true whenever a majority of group members judges it to be true. Below I discuss several other aggregation procedures.

Table 1: An aggregation procedure



Of course, an aggregation procedure captures only part of a group's institutional structure (which may be quite complex), and there are also multiple ways (both formal and informal ones) in which a group might implement such a procedure.

Nonetheless, as argued below, aggregation procedures are important factors in determining a group's epistemic performance.

In the next section, I ask what properties a group's aggregation procedure must have for the group to meet the rationality challenge, i.e. to generate consistent collective judgments, and in the subsequent section, I ask what properties it must have for the group to meet the knowledge challenge, i.e. to track the truth in its judgments. Both discussions illustrate that a group's performance as an epistemic agent depends on its aggregation procedure.

4. The rationality challenge

Suppose a group has to form collectively endorsed beliefs or judgments on certain propositions. Can it ensure the consistency of these judgments?

4.1. A 'discursive dilemma'

Consider an expert committee that has to prepare a report on the health consequences of air pollution in a big city, especially pollution by particles smaller than 10 microns in diameter. This is an issue on which there has recently been much debate in Europe. The experts have to make judgments on the following propositions:

p: The average particle pollution level exceeds 50μgm^{-3} (micrograms per cubic meter air).

p→q: If the average particle pollution level exceeds 50μgm^{-3}, then residents have a significantly increased risk of respiratory disease.

q: Residents have a significantly increased risk of respiratory disease.

All three propositions are complex factual propositions on which the experts may disagree.[2] Suppose the experts use majority voting as their aggregation procedure, i.e. the collective judgment on each proposition is the majority judgment on that proposition, as defined above. Now suppose the experts' individual judgments are as shown in table 2.

Table 2: A 'discursive dilemma'

	p	p→q	q
Individual 1	True	True	True
Individual 2	True	False	False
Individual 3	False	True	False
Majority	True	True	False

Then a majority of experts judges p to be true, a majority judges p→q to be true, and yet a majority judges q to be false, an inconsistent collective set of judgments. The expert committee fails to meet the rationality challenge in this case.

[2] Propositions p and p→q can be seen as 'premises' for the 'conclusion' q. Determining whether p is true requires an evaluation of air quality measurements; determining whether p→q is true requires an understanding of causal processes in human physiology; finally, determining whether q is true requires a combination of the judgments on p and p→q.

This problem – sometimes called a 'discursive dilemma' – illustrates that, under the initially plausible aggregation procedure of majority voting, a group may not achieve consistent collective judgments even when all group members hold individually consistent judgments (Pettit 2001; List and Pettit 2002, 2004; List 2005a).

Is the present example just an isolated artefact, or can we learn something more general from it?

4.2. An impossibility theorem

Consider again any group of two or more individuals that has to make judgments on a set of non-trivially interconnected propositions, as in the expert committee example.[3] Suppose that each individual holds complete and consistent judgments on these propositions, and that the group judgments are also required to be complete and consistent.[4] One can then prove the following impossibility result.

Theorem (List and Pettit 2002). There exists no aggregation procedure generating complete and consistent collective judgments that satisfies the following three conditions simultaneously:

Universal domain. The procedure accepts as admissible input any logically possible combinations of complete and consistent individual judgments on the propositions.

Anonymity. The judgments of all individuals have equal weight in determining the collective judgments.

Systematicity. The collective judgment on each proposition depends only on the individual judgments on that proposition, and the same pattern of dependence holds for all propositions.

In short, majority voting is not the only aggregation procedure that runs into problems like the one illustrated in table 2 above. Any procedure satisfying universal domain, anonymity and systematicity does so. If these conditions are regarded as indispensable requirements on an aggregation procedure, then one has to conclude that a multi-member group cannot meet the rationality challenge in forming its collective judgments. But this conclusion would be too quick. The impossibility theorem should be seen as characterizing the logical space of aggregation procedures (List and Pettit 2002; List 2005a). In particular, we can characterize different aggregation procedures in terms of which conditions they meet and which they violate.

---

[3] A set of propositions is 'non-trivially interrelated' if it is of one of the following forms (or a superset thereof): (i) it includes k>1 propositions p_1, ..., p_k and either their conjunction 'p_1 and ... and p_k' or their disjunction 'p_1 or p_2 or … or p_k' or both (and the negations of all these propositions); (ii) it includes k>1 propositions p_1, ..., p_k, another proposition q and either the proposition 'q if and only if (p_1 and ... and p_k)' or the proposition 'q if and only if (p_1 or p_2 or … or p_k)' or both (and negations); (iii) it includes propositions p, q and $p \rightarrow q$ (and negations). This definition is given in List (2005).

[4] An agent's judgments are 'complete' if, for each proposition-negation pair, the agent judges either the proposition or its negation to be true; they are 'consistent' if the set of propositions judged to be true by the agent is a consistent set in the standard sense of propositional logic. This is a slightly stronger consistency notion than the one in List and Pettit (2002). But when the present consistency notion is used, no additional deductive closure requirement is needed (unlike in List and Pettit 2002).

To find an aggregation procedure that allows a group to meet the rationality challenge, we have to relax at least one of the conditions of the theorem.

4.3. First solution: giving up universal domain

If the amount of disagreement in a particular group is limited or if the group has mechanisms in place for reducing disagreement – such as mechanisms of group deliberation – the group might opt for an aggregation procedure that violates universal domain. For example, a deliberating group that successfully avoids combinations of individual judgments of the kind in table 2 might use majority voting as its aggregation procedure and yet meet the rationality challenge.

But this solution does not work in general. Even in an expert committee whose task is to make judgments on factual matters without conflicts of interest, disagreement may still be significant and pervasive. Although one can study conditions that make the occurrence of judgment combinations of the kind in table 2 less likely (Dryzek and List 2003; List 2002), I here set this issue aside and assume that groups involved in epistemic tasks should normally use aggregation procedures satisfying universal domain.

4.4. Second solution: giving up anonymity

It can be shown that, if we give up anonymity but insist on the other two conditions, the only possible aggregation procedure is a 'dictatorial procedure', whereby the collective judgments are always those of some antecedently fixed group member (the 'dictator') (Pauly and van Hees 2005). Some groups might be prepared to put one individual – say a committee chair – in charge of forming its collective judgments. But this solution conflicts with the idea of a democratically organized group or committee. Moreover, as discussed below, a group organized in this dictatorial way loses out on the epistemic advantages of a democratic structure. (But I also suggest that some groups' epistemic performance may benefit from using an aggregation procedure that gives up anonymity together with systematicity, so as to implement a division of epistemic labour among several individuals.)

4.5. Third solution: giving up systematicity

A potentially promising solution lies in giving up systematicity, i.e. treating different propositions differently in the process of forming collective judgments. In particular, a group may designate some propositions as 'premises' and others as 'conclusions' and assign epistemic priority either to the premises or to the conclusions (for a more extensive discussion of this process, see List 2005a).

If the group assigns priority to the premises, it may use the so-called 'premise-based procedure', whereby the group first makes a collective judgment on each premise by taking a majority vote on that premise and then derives its collective judgments on the conclusions from these collective judgments on the premises. In the expert committee example, propositions p and p→q might be designated as premises (perhaps on the grounds that p and p→q are more basic than q), and proposition q might be designated as a conclusion. The committee might then take majority votes on p and p→q and derive its judgment on q from its judgments on p and p→q.[5]

Alternatively, if the group assigns priority to the conclusions, it may use the so-called 'conclusion-based procedure', whereby the group takes a majority vote only on each conclusion and makes no collective judgments on the premises. In addition to violating systematicity, this aggregation procedure fails to produce complete collective judgments. But sometimes a group is required to make judgments only on conclusions, but not on premises, and in such cases incompleteness in the collective judgments on the premises may be defensible.

The premise- and conclusion-based procedures are not the only aggregation procedures violating systematicity. Further interesting possibilities arise when the group is willing to give up both systematicity and anonymity. The group can then adopt an aggregation procedure that not only assigns priority to the premises, but also implements a division of epistemic labour. Specifically, the group may use the so-called 'distributed premise-based procedure'. Here different individuals specialize on different premises and give their individual judgments only on these premises. Now the group makes a collective judgment on each premise by taking a majority vote on that premise among the relevant 'specialists', and then the group derives its collective judgments on the conclusions from these collective judgments on the premises. This procedure is discussed in greater detail below.

For many epistemic tasks performed by groups, giving up systematicity and using a (regular or distributed) premise-based or conclusion-based procedure may be an attractive way to avoid the impossibility result explained above. Each of these procedures allows a group to meet the rationality challenge. Arguably, a premise-based or distributed premise-based procedure makes the pursuit of epistemic agency at the group level particularly visible. A group using such a procedure may seem to act like a reason-driven agent when it derives its collective judgments on conclusions from its collective judgments on relevant premises.

However, giving up systematicity comes with a price. Aggregation procedures that violate systematicity may be vulnerable to manipulation by prioritizing propositions strategically. For example, in the case of a regular premise-based procedure, the collective judgments may be sensitive to the choice of premises. In the example of table 2, if p and p→q are designated as premises, then all three propositions, p, p→q and q, are collectively judged to be true; if p and q are designated as premises, then p is judged to be true and both q and p→q are judged to be false; finally, if q and p→q are designated as premises, then p→q is judged to be true, and both p and q are judged to be false. Although there seems to be a natural choice of premises in the present example, namely p and p→q, this may not generally be the case, and the outcome of a premise-based procedure may therefore depend as much on the choice of premises as it depends on the individual judgments to be aggregated. In the case of a distributed premise-based procedure, an additional sensitivity to the choice of 'specialists' on each premise arises. Likewise, in the case of the conclusion-based

[5] In the present example, the truth-value of q is not always settled by the truth-values of p and p→q; so the group may need to stengthen its premises in order to make them sufficient to determine its judgment on the conclusion.

procedure, the choice of conclusions obviously matters, since the group makes collective judgments only on these conclusions and on no other propositions.[6]

4.6. Fourth solution: permitting incomplete collective judgments

The first three solutions to the rationality challenge have required giving up one of the three minimal conditions on how individual judgments are aggregated into collective judgments. The present solution preserves these minimal conditions, but weakens the requirements on the collective judgments themselves by permitting incompleteness in these judgments (see also List 2005a).

If a group is prepared to refrain from making a collective judgment on some propositions – namely on those on which there is too much disagreement between the group members – then it may use an aggregation procedure such as the 'unanimity procedure', whereby the group makes a judgment on a proposition if and only if the group members unanimously endorse that judgment. Propositions judged to be true by all members are collectively judged to be true; and ones judged to be false by all members are collectively judged to be false; no collective judgment is made on any other propositions. (Instead of the unanimity procedure, the group might also use 'supermajority voting' with a sufficiently large supermajority threshold.)

Groups operating in a strongly consensual manner may well opt for this solution, but in many cases making no judgment on some propositions is simply not an option. For example, when an expert committee is asked to give advice on a particular issue, it is usually expected to take a determinate stance on that issue.

4.7. Lessons to be drawn

I have shown that aggregation procedures matter with respect to the rationality challenge: a group of individuals that seeks to make collective judgments on a set of non-trivially interconnected propositions can meet the rationality challenge only if it is willing to adopt a procedure that violates one of universal domain, anonymity or systematicity or that produces incomplete collective judgments. Moreover, different aggregation procedures may lead to different collective judgments for the same combination of individual judgments. As an illustration, table 3 shows the collective judgments for the individual judgments in table 2 under different aggregation procedures.

[6] It can be shown that in some important respects, the premise-based procedure is more vulnerable to strategic manipulation than the conclusion-based procedure. See Dietrich and List (2005).

Table 3: Different aggregation procedures applied to the individual judgments in table 2

	p	p→q	q
Majority voting*	True	True	False
Premise-based procedure with p, p→q as premises	True	True	True
Conclusion-based procedure with q as conclusion	No judgment	No judgment	False
Distributed premise-based procedure with individual 1 specializing on p and individual 2 specializing on p→q	True	False	False
Unanimity procedure	No judgment	No judgment	No judgment
Dictatorship of individual 3	False	True	False

* inconsistent

If we were to assess a group's epistemic performance solely on the basis of whether the group meets the rationality challenge, this would give us insufficient grounds for selecting a unique aggregation procedure. As I have illustrated, many different aggregation procedures generate consistent collective judgments, and even if we require completeness in addition to consistency, several possible aggregation procedures remain. To recommend a suitable aggregation procedure that a group can employ for a given epistemic task, the question of whether the group meets the rationality challenge alone is not a sufficient criterion. Goldman (2004) has noted this point in his critique of a pure rationality-based approach to social epistemology.

5. The knowledge challenge

Can a group's collective beliefs or judgments constitute knowledge? Following Nozick (1981), an agent knows that p if four conditions are met. First, p is true. Second, the agent believes that p. Third, if p were true, the agent would believe that p. Fourth, if p were not true, the agent would not believe that p. These conditions can be applied to any epistemic agent, individual or collective. In particular, if a group's instutitional structure allows the group to form collectively endorsed beliefs or judgments, then one can ask whether these beliefs or judgments satisfy Nozick's conditions. (Readers who prefer a different account of knowledge may substitute their preferred account.)

As a simple reliabilist measure of how well an agent satisfies Nozick's third and fourth conditions, I use two conditional probabilities (List 2005a): the probability that the agent believes p to be true given that p is true, and the probability that the agent does not believe p to be true given that p is false. Call these two conditional probabilities the agent's 'positive' and 'negative reliability' on p, respectively.

By considering a group's positive and negative reliability on various propositions under diffferent aggregation procedures and different scenarios, I now show that it is possible for a group to meet the knowledge challenge, but that, once again, the aggregation procedure affects a group's success.

5.1. The first scenario and its lesson: epistemic gains from democratization

Suppose that a group has to make a collective judgment on a single factual proposition, such as proposition p in the expert committee example above. As a baseline scenario (e.g. Grofman, Owen and Feld 1983), suppose that the group members hold individual judgments on proposition p, where two conditions are met. First, each group member has the same positive and negative reliability r on proposition p, where $1 > r > 1/2$ (the 'competence' condition); so individual judgments are noisy but biased towards the truth. Second, the judgments of different group members are mutually independent (the 'independence' condition). (Obviously, it is also important to study scenarios where these conditions are violated, and below I consider some such scenarios.[7])

The group must use an aggregation procedure to make its collective judgment on p based on the group members' individual judgments on p. What is the group's positive and negative reliability on p under different aggregation procedures?

Let me compare three different procedures: first, a dictatorial procedure, where the collective judgment is always determined by the same fixed group member; second, the unanimity procedure, where agreement among all group members is necessary for reaching a collective judgment; and third, majority voting, which perhaps best implements the idea of democratic judgment aggregation (at least in the case of a single proposition).

Under a dictatorial procedure, the group's positive and negative reliability on p equals that of the dictator, which is r by assumption.

Under the unanimity procedure, the group's positive reliability on p equals r^n, which approaches 0 as the group size increases, but its negative reliability on p equals $1-(1-r)^n$, which approaches 1 as the group size increases. This means that the unanimity procedure is good at avoiding false positive judgments, but bad at reaching true positive ones. A determinate collective judgment on p is reached only if all individuals agree on the truth-value of p; if they don't agree, no collective judgment on p is made.

Finally, under majority voting, the group's positive and negative reliability on p approaches 1 as the group size increases. Why does this result hold? Each individual has a probability $r>0.5$ of making a correct judgment on p; by the law of large numbers, the proportion of individuals who make a correct judgment on p approaches $r>0.5$ as the group size increases and thus constitutes a majority with a probability approaching 1. Informally, majority voting allows the group to extract the signal from the group members' judgments, while filtering out the noise. This is the famous 'Condorcet jury theorem'.

---

[7] Cases where different individuals have different levels of reliability are discussed, for example, in Grofman, Owen and Feld (1983) and Borland (1989). Cases where there are dependencies between different individuals' judgments are discussed, for example, in Ladha (1992), Estlund (1994) and Dietrich and List (2004). Cases where individuals express their judgments strategically rather than truthfully are discussed in Austen-Smith and Banks (1996).

Table 4 shows the group's positive and negative reliability on p under majority voting and under a dictatorial procedure, and tables 5 and 6 show, respectively, the group's positive and negative reliability on p under a dictatorial procedure and under the unanimity procedure. In each case, individual group members are assumed to have a positive and negative reliability of r=0.54 on p. In all tables, the group size is on the horizontal axis and the group's reliability on the vertical axis.[8]

Table 4: The group's positive and negative reliability on p majority voting (top curve); dictatorship (bottom curve) (setting r=0.54 as an illustration)

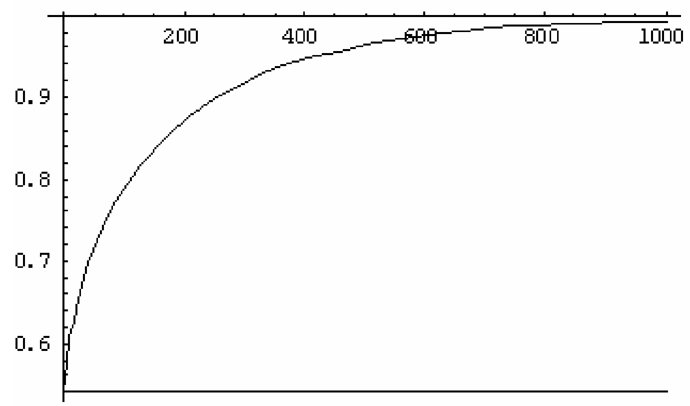
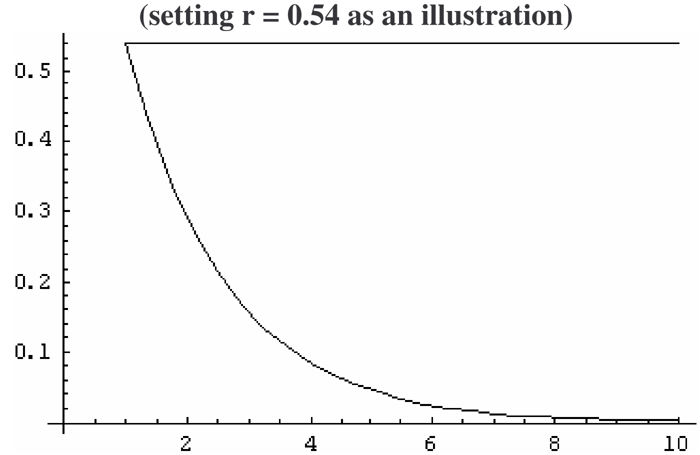
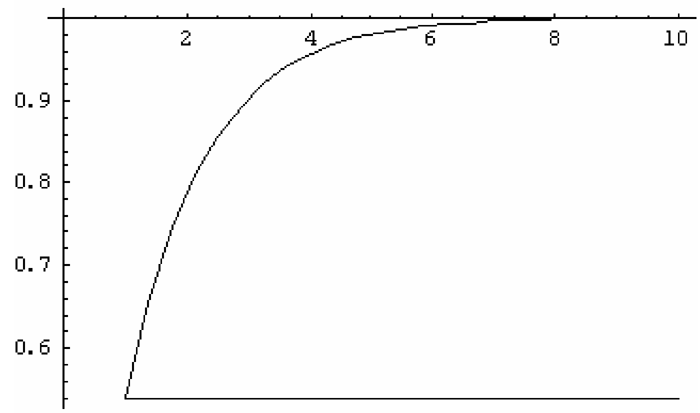


Table 5: The group's positive reliability on p: dictatorship (top curve); unanimity procedure (bottom curve) (setting r = 0.54 as an illustration)



[8] The present curves are the result of averaging between two separate curves for even- and odd-numbered group sizes. (When the group size is an even number, the group's reliability may be lower because of the possibility of majority ties.)

Table 6: The group's negative reliability on p: unanimity procedure (top curve); dictatorship (bottom curve) (setting r = 0.54 as an illustration)



What lessons can be drawn from this first scenario? If individuals are independent, fallible, but biased towards the truth, majority voting outperforms both dictatorial and unanimity procedures in terms of maximizing the group's positive and negative reliability on p. The unanimity procedure is attractive only in those special cases where the group seeks to minimize the risk of making false positive judgments (such as in some jury decisions); a dictatorial procedure fails to pool the information held by different individuals.

Hence, when a group seeks to meet the knowledge challenge, there may be 'epistemic gains from democratization', i.e. from making a collective judgment on a given proposition democratically by using majority voting. More generally, even when individual reliability differs between individuals, a weighted form of majority voting still outperforms a dictatorship by the most reliable individual: each individual's vote simply needs to have a weight proportional to $\log(r/(1-r))$, where r is the individual's reliability on the proposition in question (Ben-Yashar and Nitzan 1997).

5.2. The second scenario and its lesson: epistemic gains from disaggregation

Suppose now that a group has to make a collective judgment not only on a single factual proposition, but on a set of interconnected factual propositions. As an illustration, suppose that there are $k>1$ premises $p_1, ..., p_k$ and a conclusion q, where q is true if and only if the conjunction of $p_1, ..., p_k$ is true. (This structure also allows representing a variant of the expert committee example above. For extensive discussions of the present scenario and other related scenarios, see Bovens and Rabinowicz 2005 and List 2005a,b. Analogous points apply to the case where q is true if and only if the disjunction of $p_1, ..., p_k$ is true.)

In this case of multiple interconnected propositions, individuals cannot generally have the same reliability on all propositions. Suppose, as an illustration, that each individual has the same positive and negative reliability r on each premise $p_1, ..., p_k$ and makes independent judgments on different premises. Then each individual's positive reliability on the conclusion q is r^k, which is below r and often below 0.5 (whenever $r < \sqrt[k]{0.5}$), while his or her negative reliability on q is above r. Here individuals are much worse at detecting the truth of the conclusion than the truth of each premise, but much better at detecting the falsehood of the conclusion than the falsehood of each premise. In the expert committee example, it might be easier to make correct judgments on propositions p and p→q than on proposition q. Of course, other scenarios can also be constructed, but the point remains that individuals typically have different levels of reliability on different propositions (List 2005a).

What is the group's positive and negative reliability on the various propositions under different aggregation procedures? As before, suppose the judgments of different group members are mutually independent.

Majority voting performs well only on those propositions on which individuals have a positive and negative reliability above 0.5. As just argued, individuals may not meet this condition on all propositions. Moreover, majority voting does not generally produce consistent collective judgments (on the probability of majority inconsistencies, see List 2005b). Let me now compare dictatorial, conclusion-based and premise-based procedures.

Under a dictatorial procedure, the group's positive and negative reliability on each proposition equals that of the dictator; in particular, the probability that *all* propositions are judged correctly is r^k, which may be very low, especially when the number of premises k is large.

Under the conclusion-based procedure, unless individuals have a high reliability on each premise, namely $r > \sqrt[k]{0.5}$ (e.g. 0.71 when k=2, or 0.79 when k=3), the group's positive reliability on the conclusion q approaches 0 as the group size increases. Its negative reliability on q approaches 1. Like the unanimity procedure in the single-proposition case, the conclusion-based procedure is good at avoiding false positive judgments on the conclusion, but (typically) bad at reaching true positive ones.

Under the premise-based procedure, the group's positive and negative reliability on every proposition approaches 1 as the group size increases. This result holds because, by the Condorcet jury theorem as stated above, the group's positive and negative reliability on each premise $p_1, ..., p_k$ approaches 1 with increasing group size, and therefore the probability that the group derives a correct judgment on the conclusion also approaches 1 with increasing group size.

As illustration, suppose that there are k=2 premises and individuals have a positive and negative reliability of r=0.54 on each premise. Table 7 shows the group's probability of judging *all* propositions correctly under the premise-based procedure and under a dictatorial procedure. Tables 8 and 9 show, respectively, the group's positive and negative reliability on the conclusion q under a dictatorial procedure and under the conclusion-based procedure.

Table 7: The group's probability of judging all propositions correctly: premise-based procedure (top curve); dictatorship (bottom curve) (setting r = 0.54 as an illustration)

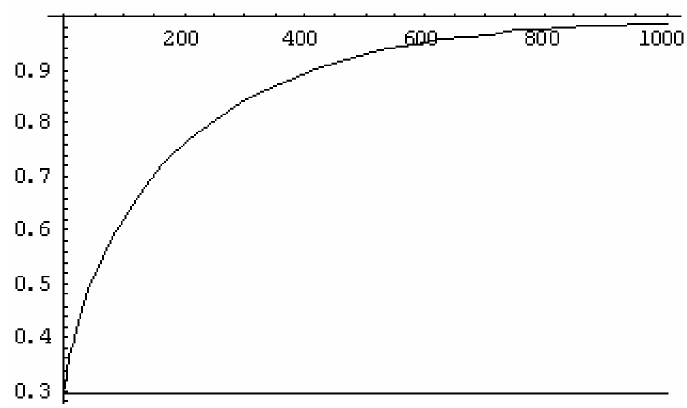


Table 8: The group's positive reliability on the conclusion q dictatorship (top curve); conclusion-based procedure (bottom curve) (setting r = 0.54 as an illustration)

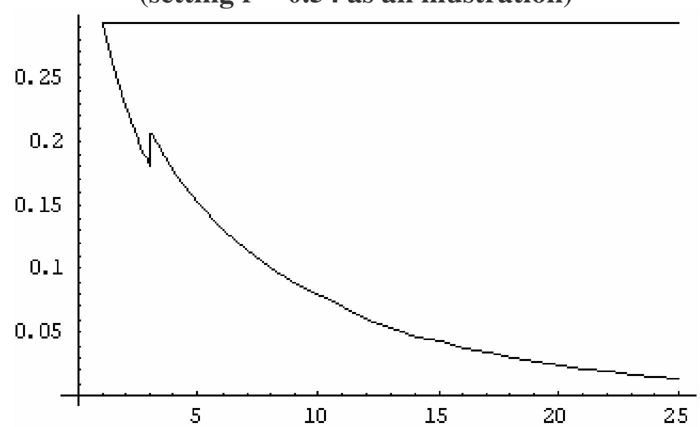
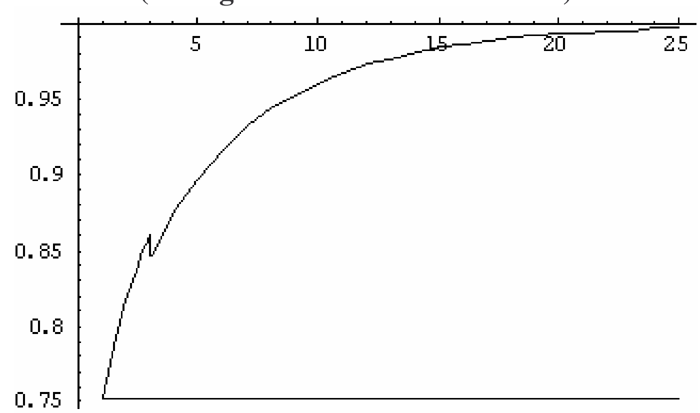


Table 9: The group's negative reliability on the conclusion q conclusion-based procedure (top curve); dictatorship (bottom curve) (setting r = 0.54 as an illustration)



What lessons can be drawn from this second scenario? Under the present assumptions, the premise-based procedure outperforms both dictatorial and conclusion-based procedures in terms of simultaneously maximizing the group's positive and negative reliability on every proposition. Like the unanimity procedure before, the conclusion-based procedure is attractive only when the group seeks to minimize the risk of making false positive judgments on the conclusion; again, a dictatorial procedure is bad at information pooling.

Hence, if a larger epistemic task such as making a judgment on some conclusion can be disaggregated into several smaller epistemic tasks such as making judgments on relevant premises, then there may be 'epistemic gains from disaggregation', i.e. from making collective judgments on that conclusion on the basis of separate collective judgments on those premises. (For a discussion of different scenarios, see List 2005a.)

5.3. The third scenario and its lesson: epistemic gains from distribution

When an epistemic task is complex in that it requires making judgments on several propositions, different individuals may have different levels of expertise on different propositions. An individual may lack the temporal, computational and informational resources to become sufficiently reliable on every proposition. If we take this problem into account, can we improve on the premise-based procedure?

Suppose, as before, that a group has to make collective judgments on $k>1$ premises p_1, ..., p_k and a conclusion q, where q is true if and only if the conjunction of p_1, ..., p_k is true. Instead of requiring every group member to make a judgment on every premise, we might partition the group into k subgroups (for simplicity, of approximately equal size), where the members of each subgroup specialize on one premise and make a judgment on that premise alone. Instead of a using a regular premise-based procedure as in the previous scenario, the group might now use a distributed premise-based procedure: the collective judgment on each premise is made by taking a majority vote within the subgroup specializing on that premise, and the collective judgment on the conclusion is then derived from these collective judgments on the premises.

When does the distributed premise-based procedure outperform the regular premise-based procedure at maximizing the group's probability of making correct judgments on the propositions?

Intuitively, there are two effects here that pull in opposite directions. First, there may be 'epistemic gains from specialization': individuals may become more reliable on the proposition on which they specialize. But, second, there may also be 'epistemic losses from lower numbers': each subgroup voting on a particular proposition is smaller than the original group (it is only approximately 1/k the size of original group when there are k premises), which may reduce the benefits from majoritarian judgment aggregation on that proposition.

Whether or not the distributed premise-based procedure outperforms the regular premise-based procedure depends on which of these two opposite effects is stronger. Obviously, if there were no epistemic gains from specialization, then the distributed premise-based procedure would suffer only from losses from lower numbers on each premise and would therefore perform worse than the regular premise-based procedure. On the other hand, if the epistemic losses from lower numbers were relatively small compared to the epistemic gains from specialization, then the distributed premise-based procedure would outperform the regular one. The following result holds:

Theorem (List 2003). For any group size n (divisible by k), there exists an individual (positive and negative) reliability level r* > r such that the following holds: if, by specializing on some proposition p, individuals achieve a reliability above r* on p, then the majority judgment on p in a subgroup of n/k specialists (each with reliability r* on p) is more reliable than the majority judgment on p in the original group of n non-specialists (each with reliability r on p).

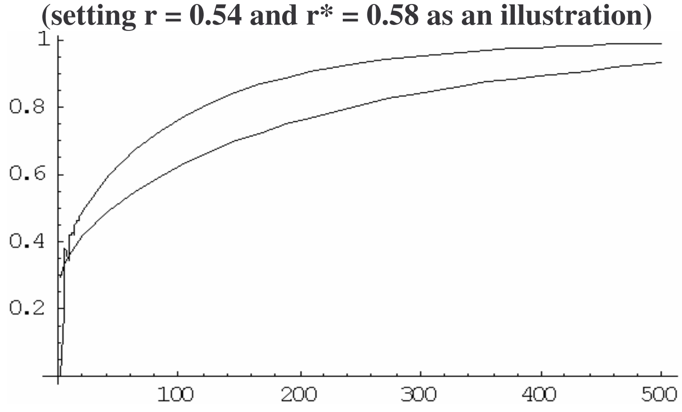
Hence, if by specializing on one premise, individuals achieve a reliability above r* on that premise, then the distributed premise-based procedure outperforms the regular premise-based procedure. How great must the reliability increase from r to r* be to have this effect? Strikingly, a small reliability increase typically suffices. Table 10 shows some sample calculations. For example, when there are k=2 premises, if the original individual reliability was r=0.52, then a reliability above r*=0.5281 after specialization suffices; it it was r=0.6, then a reliability above r*=0.6393 after specialization suffices.

Table 10: Reliability increase from r to r* required to outweigh the loss from lower numbers

	k = 2, n = 50			k = 3, n = 51			k = 4, n = 52		
r =	0.52	0.6	0.75	0.52	0.6	0.75	0.52	0.6	0.75
r* =	0.5281	0.6393	0.8315	0.5343	0.6682	0.8776	0.5394	0.6915	0.9098

Table 11 shows the group's probability of judging *all* propositions correctly under regular and distributed premise-based procedures, where there are k=2 premises and where individuals have positive and negative reliabilities of r=0.54 and r*=0.58 before and after specialization, respectively.

Table 11: The group's probability of judging all propositions correctly: distributed (top curve) and regular premise-based procedure (bottom curve) (setting r = 0.54 and r* = 0.58 as an illustration)



What lessons can be drawn from this third scenario? Even when there are only relatively modest gains from specialization, the distributed premise-based procedure may outperform the regular premise-based procedure in terms of maximizing the group's positive and negative reliability on every proposition.

Hence there may be 'epistemic gains from distribution': if a group has to perform a complex epistemic task, the group may benefit from subdividing the task into several smaller tasks and distributing these smaller tasks across multiple subgroups.

Such division of epistemic labour is also the mechanism underlying the successes of 'collectively distributed cognition', as recently discussed in the philosophy of science. For example, Knorr Cetina (1999) provides a case study of distributed cognition in science. Investigating the research practices in high-energy physics at the European Center for Nuclear Research (CERN), Knorr Cetina observes that experiments, which lead to research reports and papers, involve many researchers and technicians, using complex technical devices, with a substantial division of labour, expertise, and authority (for a critical discussion, see also Giere 2002).[9] Such research practices rely on mechanisms similar to those represented, in a stylized form, by the distributed premise-based procedure.

In conclusion, when a group is faced with a complex epistemic task, it is possible for the group to meet the knowledge challenge, but the group's aggregation procedure plays an important role in determining its success.

6. Concluding remarks

I have explained several key concepts and results from the theory of judgment aggregation in order to suggest a formal approach to thinking about institutions in social epistemology. Within this framework, I have discussed the rationality and knowledge challenges that groups as epistemic agents face. I have argued that, rather than pointing towards two different approaches to social epistemology, the two challenges should be seen as two important problems that can be addressed within a single approach. In relation to both challenges, a group's aggregation procedure, and thus its institutional structure, matters.

With regard to the rationality challenge, I have discussed an impossibility theorem, which allows us to characterize the logical space of aggregation procedures under which a group can meet the rationality challenge. No aggregation procedure generating complete and consistent collective judgments can simultaneously satisfy universal domain, anonymity and systematicity. To find an aggregation procedure that allows a group to meet the rationality challenge, it is therefore necessary to relax one of universal domain, anonymity or systematicity, or to permit incomplete collective judgments. Which relaxation is most defensible depends on the group and epistemic task in question.

With regard to the knowledge challenge, I have identified three effects that are relevant to the design of a good aggregation procedure: there may be epistemic gains from democratization, disaggregation and distribution. Again, the applicability and magnitude of each effect depends on the group and epistemic task in question, and

---

[9] Knorr Cetina also investigates research practices in molecular biology, but argues that, in that field, research is more individualized than in high energy physics and individual researchers remain the relevant epistemic agents here.

there may not exist a 'one size fits all' aggregation procedure which is best for all groups and all epistemic tasks. But the fact that a group may sometimes benefit from the identified effects reinforces the importance of institutional design in social epistemology.

Overall, the present results give a fairly optimistic picture of a group's capacity to perform as an epistemic agent. Yet there is also an abundance of work in philosophy and economics that focuses on failures of collective agency. (Consider, for example, the large literature on the impossibility results in social choice theory.) Clearly, the details of my results depend on various assumptions and may change with changes in these assumptions. But my aim has not primarily been to defend a particular set of results on how groups perform as epistemic agents; rather, it has been to illustrate the usefulness of the theory of judgment aggregation as a framework for studying institutional design in social epistemology.

References

Austen-Smith, David, and Banks, Jeffrey (1996). Information Aggregation, Rationality and the Condorcet Jury Theorem. *American Political Science Review* 90: 34-45.

Ben-Yashar, Ruth, and Nitzan, Shmuel (1997). The optimal decision rule for fixed-size committees in dichotomous choice situations: the general result. *International Economic Review* 38: 175-186.

Borland, Philip (1989). Majority systems and the Condorcet jury theorem. *Statistician* 38: 181-189.

Bovens, Luc, and Rabinowicz, Wlodek (2005). Democratic Answers to Complex Questions: an Epistemic Perspective. *Synthese*, forthcoming.

Dietrich, Franz (2005). Judgment aggregation: (im)possibility theorems. *Journal of Economic Theory*, forthcoming.

Dietrich, Franz, and List, Christian (2004). A model of jury decisions where all jurors have the same evidence. *Synthese* 142: 175-202.

Dietrich, Franz, and List, Christian (2005). Strategy-proof judgment aggregation. Working paper, LSE.

Dryzek, John, and List, Christian (2003). Social Choice Theory and Deliberative Democracy: A Reconciliation. British Journal of Political Science 33(1): 1-28.

Estlund, David (1994). Opinion leaders, independence and Condorcet's jury theorem. *Theory and Decision* 36: 131-162.

Giere, Ronald (2002). Distributed Cognition in Epistemic Cultures. *Philosophy of Science* 69: 637-644.

Goldman, Alvin (1999). *Knowledge in a Social World*. Oxford: Oxford University Press.

Goldman, Alvin (2004). Group Knowledge versus Group Rationality: Two Approaches to Social Epistemology. *Episteme, A Journal of Social Epistemology* 1: 11-22.

Goodin, Robert E. (ed.) (1996). *The Theory of Institutional Design*. Cambridge: Cambridge University Press.

Grofman, Bernard, Owen, Guillermo, and Feld, Scott (1983). Thirteen theorems in search of the truth. *Theory and Decision* 15: 261-278.

Knorr Cetina, Karin (1999). *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge/MA: Harvard University Press.

Ladha, Krishna (1992). The Condorcet jury theorem, free speech and correlated votes. *American Journal of Political Science* 36: 617-634.

List, Christian (2002). Two Concepts of Agreement. *The Good Society* 11(1): 72-79.

List, Christian (2003). Distributed Cognition: A Perspective from Social Choice Theory. Presentation, Australian National University, London School of Economics, University of Liverpool.

List, Christian (2005a). The Discursive Dilemma and Public Reason. *Ethics*, forthcoming.

List, Christian (2005b). The Probability of Inconsistencies in Complex Collective Decisions. *Social Choice and Welfare* 24(1): 3-32.

List, Christian, and Pettit, Philip (2002). Aggregating Sets of Judgments: An Impossibility Result. *Economics and Philosophy* 18: 89-110.

List, Christian, and Pettit, Philip (2004). Aggregating Sets of Judgments: Two Impossibility Results Compared. *Synthese* 140(1-2): 207-235.

List, Christian, and Pettit, Philip (2005a). On the Many as One. *Philosophy and Public Affairs*, forthcoming.

List, Christian, and Pettit, Philip (2005b). Group Agency and Supervenience. Working paper, LSE.

Nozick, Robert (1981). *Philosophical Explanations*. Cambridge/MA: Harvard University Press.

Pauly, Marc, and van Hees, Martin (2005). Logical Constraints on Judgment Aggregation. *Journal of Philosophical Logic*, forthcoming.

Pettit, Philip (2001) Deliberative Democracy and the Discursive Dilemma. *Philosophical Issues* (supplement to *Nous*) 11: 268-99.

Pettit, Philip (2003). Groups with Minds of their Own. In Fred Schmitt (ed.). *Socializing Metaphysics*. New York: Rowan and Littlefield.

Rovane, Carol (1998). *The Bounds of Agency: An essay in Revisionary Metaphysics*. Princeton: Princeton University Press.