Judgement Aggregation and Distributed Thinking

Kai Spiekermann

Revision 1, June 2009

Abstract

In recent years, judgement aggregation has emerged as an important area of social choice theory. Judgement aggregation is concerned with aggregating sets of individual judgements over logically connected propositions into a set of collective judgements. It has been shown that even seemingly weak conditions on the aggregation function make it impossible to find functions that produce rational collective judgements from all possible rational individual judgements. This implies that the step from individual judgements to collective judgements requires trade-offs between different desiderata, such as universal domain, rationality, epistemological quality, and unbiasedness. These dilemmas challenge us to decide which conditions we should relax. The typical application for judgement aggregation is the problem of group decision making. Juries and expert committees are the stock examples. However, the relevance of judgement aggregation goes beyond these cases. In this survey I review some core results in the field of judgement aggregation and social epistemology and discuss their implications for the analysis of distributed thinking.

Introduction

Thinking is often taken as an activity exercised by individuals. In recent years, however, it has been acknowledged that thinking can also be a collective process. It is not only individuals who process information, take stances, and make decisions—groups can do this, too. For instance, a court jury needs to gather information, reach collective stances on the information available, and make a decision on the sentence. The same is true for cabinets, expert panels, shipping crews, air-traffic controllers, appointment committees, et cetera. Individuals can differ in their ability to process information rationally and reach correct decisions. Similarly, groups can differ in their success to arrive at correct decisions, and they may arrive at these decisions in a rational or in an irrational way. In this sense groups are engaged in collective thinking.

It is difficult to observe how individuals process complex information and arrive at a decision. For groups, however, this process is more transparent. Psychologists and social scientists can observe how groups deliberate, how they form judgements, and how they finally reach decisions. These collective decision processes can be compared and evaluated. Some processes are obviously epistemically poor and irrational: For instance, a court jury should not throw a coin to decide whether a defendant is guilty; an expert panel should (arguably) not randomly select one expert to make all decisions, a cabinet should not always choose the course of action with the least support, etc. But it is more difficult to determine good collective decision procedures, and we will see that it is often impossible to determine collective decision procedures that meet some seemingly harmless desiderata.

Different strands of literature have discussed the nature of distributed thinking. One strand, inspired by the psychologist Lev Vygotsky (1962, 1978), analyses distributed thinking in relation to distributed cognition. In recent years, this research paradigm was advanced by Edwin Hutchins (1995) and his influential study of "cognition in the wild". Research on distributed cognition was also influenced by Andy Clark's and David Chalmers' concept of the "extended mind" (Clark and Chalmers, 1998; Clark, 1997). This approach emphasizes that the boundaries between mind and world are often difficult to draw, and that organisms reshape their environment to solve the problems they encounter. The mind is a "leaky organ", as Clark (1997, p. 53) puts it, "mingling shamelessly with body and with world".

Another strand of literature that tackles the phenomenon of distributed thinking draws on concepts from social choice theory and social epistemology. Goldman (2004) and List (2005, forthcoming*a*) have pointed out that there are at least two dimensions on which a group's performance as a thinking system can be measured. First, the group can succeed or fail to be rational, where rationality is understood as avoiding logical contradictions in the judgements the group makes. Second, the group can be more or less successful in reaching correct decisions, given the information available. The first dimension poses a "rationality" or "coherence" challenge, the second a "knowledge" or "correspondence" challenge to the group. The rationality challenge can be explored with tools provided by social choice theory, the knowledge challenge with generalizations of the Condorcet Jury Theorem and the information pooling literature.

Social choice theory systematically investigates the processes to aggregate individual information into collective information. The classical problem for social choice theory is the aggregation of preferences. The famous Arrow Theorem shows that there is no aggregation procedure to map the individual preferences to collective preferences that meets some seemingly harmless and arguably normatively desirable conditions (Arrow, 1963). While the aggregation of preferences is of great importance for welfare economics, it is not quite the right framework to address distributed thinking. However, recently the social choice framework has been extended to the more general question of judgement aggregation. Judgement aggregation investigates different procedures to aggregate individual judgements to collective judgements. Again, impossibility results arise, posing challenges for distributed thinking.

Related to the field of judgement aggregation are considerations regarding the epistemic quality of different aggregation procedures. The discussion starts with Condorcet's famous observation that large groups tend to make correct dichotomous choices. But Condorcet's ideas can be extended to other choice situations as well. If information is distributed between agents, and these agents need to arrive at a joint decision based on the information, then one can ask which procedures are best suited to aggregate this information to maximize the probability of a correct decision.

Returning to the two strands of literature mentioned above, it appears that the distributed cognition approach on the one hand, and the social choice and epistemology approach on the other, talk past each other. So far, the exchange between the two approaches has been limited. This paper does not attempt a reconcilliation or propose a new synthesis. The rather more modest goal is to introduce some basic ideas from the field of judgement aggregation and considers the upshot for distributed thinking. The paper is in 4 sections. I start by explaining the problem posed by the “discursive dilemma” and how it pertains to distributed thinking. Section 2 generalizes by approaching problems of judgement aggregation more formally. In section 3 I analyze the epistemic performance of different judgement aggregation procedures. I discuss the relation between rationality, consistency, and distributed thinking in section 4. At this point, I will return to the relation between distributed cognition and social choice theory, and discuss how these two approaches may relate. More specifically, I will argue that judgement aggregation provides a framework for the analysis of distributed thinking, despite charges that it is too reductionist to be of interest.

1 The Discursive Dilemma

A central problem that has triggered much work in the field of judgement aggregation is the so-called “doctrinal paradox” (Kornhauser and Sager, 1986) or, more generally, the “discursive dilemma” (List and Pettit, 2002). I start by describing two examples that illustrate the problem.

Consider three MI5 officers who have to evaluate whether an observed suspect is planning to build a bomb. There are three officers, and they assess the situations by forming judgements on correctness of three propositions:

- The suspect has bought fertilizer (*P*).
- If the suspect has bought fertilizer, it follows that the subject plans to assemble a bomb ($P \rightarrow Q$) .[1]
- The suspect plans to assemble a bomb (*Q*).

These propositions are logically connected. For instance, if an officer believes that the subject has bought fertilizer, and if she also believes that if the subject has bought fertilizer then the subject is building a bomb, then the officer must also hold that the subject plans to build a bomb. If she does not, the officer’s judgements would be inconsistent.

We assume that all officers (individually) hold consistent sets of beliefs, i. e. they do not contradict themselves, and that they make judgements on all propositions at stake. One possible constellation of consistent individual judgements over these three propositions is shown in table 1.

1 For the example discussed here we can take $\rightarrow$ as the material conditional.

Officer	P	$P \rightarrow Q$	Q
1	true	true	true
2	true	false	false
3	false	true	false
Majority:	**true**	**true**	**false**

Table 1**:** An example of the Discursive Dilemma.

Officer 1 thinks that the suspect has bought fertilizer, that if the subject has bought fertilizer he is planning to build a bomb, and consequently thinks the suspect builds a bomb. Officer 2 also thinks that the suspect has bought fertilizer, but disagrees with the claim that buying fertilizer implies that the suspect builds a bomb, and thinks that the suspect does not build a bomb. Officer 3 disagrees with his two colleagues about whether the suspect has bought fertilizer. He believes that if the suspect had bought fertilizer he would be building a bomb. But since he has not, officer 3 can hold (for whatever reason, as no conclusion follows from the premises) that the suspect does not build a bomb.

The problem in this situation is that the three officers will find it difficult to determine their joint stance as an investigative unit. A majority vote on each proposition yields the results as stated in bottom row of table 1. A majority thinks that the suspect has bought fertilizer, a majority thinks that if the suspect has bought fertilizer he is assembling a bomb, but a majority also thinks that the suspect is not building a bomb. Thus the majority judgements are contradictory. This contradiction instantiates one version of the discursive dilemma.

Consider a second example to demonstrate that the discursive dilemma comes in different forms. Here a team of detectives has to decide whether to bring charges forward against a suspected murderer. The three detectives consider the following propositions:

- The murder weapon is identified (P).
- The suspect had a motive (Q).
- The suspect should be charged (R).
- Charges should be brought forward if and only if the weapon is identified and the suspect had a motive ($P \wedge Q \leftrightarrow R$).

We assume that the three detectives all agree on the last proposition, which one can interpret as a universally agreed doctrine. They disagree, however, on the other three propositions, as table 2 shows.

	P	Q	$P \wedge Q \leftrightarrow R$	R
Detective 1	true	true	true	true
Detective 2	true	false	true	false
Detective 3	false	true	true	false
Majority	**true**	**true**	**true**	**false**

Table 2**:** The discursive dilemma in conjunctive form.

As in the first example, each individual position is consistent, but the majority position is not. Holding P, Q, and $P \wedge Q \leftrightarrow R$ to be true, but R to be false, is a contradiction. The question is: how should the three detectives agree on a joint position?

In these examples we can recognize some features of the discursive dilemma and problems of judgement aggregation more generally. First, the dilemmas described here are fairly realistic in the sense that there are many situations in which groups of people hold judgements over different logically connected propositions and have to form a joint position on these propositions. Second, the examples can easily be extended to groups with more than 3 agents. Third, the individual judgements the agents hold are not unusual or unreasonable. Fourth, the dilemma only arises for certain judgement profiles, but the possibility of their occurrence challenges us to find judgement aggregation procedures that can deal with these situations.

How does the discursive dilemma pertain to distributed thinking? A thinking system understood in a minimal way is a system that takes inputs and produces outputs by processing these inputs. A distributed thinking system can be understood as a group of thinkers who coordinate their thinking activities. Since the thinking is distributed, one can expect every single thinker to do some thinking on their own. However, to function as a thinking system it is necessary to aggregate the information available to the single thinkers and produce a collective output. Judgement aggregation is a model of such a process: Each individual is a single thinker with stances on certain propositions. Since the single thinkers are part of a distributed thinking system, the system must aggregate their stances on the propositions and produce a collective stance. In the same way as we want single thinkers to be rational, we also require a system of distributed thinking to be rational. Judgement aggregation maps out the logical space of possible aggregation procedures and informs us of the options and constraints for distributed thinking.

The notion of "thinking" in the analysis offered here is deliberately minimal. It presupposes only that thinkers assign truth values to each proposition and that thinkers correctly apply propositional logic. In addition, the distributed thinking system must be able follow an aggregation rule. The problems arising from this simplified notion of thinking are neither trivial nor simple, and it is worthwhile to start with simple examples before moving on to more complex analyses. This minimal notion of thinking deliberately omits many other aspects of human thinkers: People can have degrees of beliefs, not just dichotomous judgements. Thinking does not only involve beliefs, but also desires. A complete picture of human thinking would also incorporate intentions, emotions, and consciousness. Nonetheless, I argue that such a rich notion of thinking can be set aside for now. It can be set aside because even the minimal notion of thinking used in this paper raises interesting questions about the rationality and epistemic quality of distributed thinking systems.

2 Impossibility Results And Escape Routes

I now describe the problem of judgement aggregation more generally and explain List and Pettit's (2002) impossibility result. Each individual has a *set of judgements* on a given *agenda*. The agenda contains all propositions in question and their respective negations. For the impossibility result to arise, the agenda must be sufficiently complex, that is it must contain at least two propositions P and Q and either $P \wedge Q$, $P \vee Q$, or $P \rightarrow Q$ (and their negations). An individual set of judgements must be complete (so that for all items on the agenda, it contains either the proposition or its negation), it must be consistent and deductively closed. If these three conditions are met, we call a judgement set *fully rational*. All the individual sets of judgements together form a *judgement profile*. For instance, tables 1 and 2 state specific judgement profiles.

The aim of judgement aggregation is to proceed from judgement profiles to a *collective judgement set*. We assume that the collective judgement set must also be consistent, complete, and deductively closed (this is called the *collective rationality condition*, see e.g. List, forthcoming*a*) and that the aggregation function never fails to produce output. An *aggregation function* has all possible judgement profiles as domain and all possible collective sets of judgements as co-domain, i. e. it maps judgement profiles onto collective sets of judgements. Put differently: an aggregation function takes a judgement profile as input and gives one fully rational collective set of judgements as output.

List and Pettit describe three desiderata that an aggregation function should meet:

Universal Domain.
The aggregation function accepts as input all logically possible judgement profiles, as long as all individual judgement sets are consistent, complete, and deductively closed.

Anonymity.
The aggregation function is not responsive to permutations of judgement sets in the profile. This means that the outcome should not change if we shuffle the agents, but leave everything else unchanged.

Systematicity.
The result of the aggregation function for any proposition depends only on the judgements made on *this* proposition, and the pattern of dependence is the same for all propositions.

Universal Domain is an immediately convincing desideratum: The aggregation function should be able to aggregate all logically possible profiles, as long as all individuals hold fully rational judgements. If the aggregation function did not have a universal domain it would fail to aggregate some judgement profiles that can occur, and there is no good reason to rule out any judgement profiles ex ante.

Anonymity is also a rather convincing desideratum for many aggregation problems. The intuitive appeal behind anonymity is that it ensures the equal treatment of all judgement sets, no matter who holds them. For example, anonymity rules out that the aggregation function always follows the judgement set of one individual, that is it rules out 'aggregation' by letting one agent be the dictator.

The systematicity condition is more contested. Note that it contains the weaker independence condition (see. e.g. Dietrich, 2007):

Independence.
The result of the aggregation function for any proposition depends only on the judgements made on *this* proposition.

The intuitive plausibility of independence is easy to argue for (even though it is also not uncontested). Independence ensures that the collective judgement on a proposition is influenced only by individual positions on that specific proposition. If we consider a proposition P, changes in the profile regarding any other proposition should not influence the collective judgement on P.

Systematicity is more demanding than independence because it also demands that the same pattern of individual judgements on any proposition should lead to the same collective judgement on these propositions. More precisely, for any two propositions P, Q: if all individuals have the same judgements on P and on Q, then the collective results for P and Q must not differ. The intuitive ideal behind this condition is to treat all propositions equally. Systematicity rules out, for instance, the requirement of different qualified majorities for different propositions.

List and Pettit state and prove a theorem of judgement aggregation:

Theorem (List and Pettit 2002):

> There exists no judgment aggregation function generating complete, consistent and deductively closed collective sets of judgments which satisfies unanimity, anonymity and systematicity.

This impossibility result has kicked off the research into questions of judgement aggregation and has led to a flourishing, often technically advanced literature (see List and Puppe, 2009, for a survey). The theorem is important because it systematizes the special case of the discursive dilemma and shows that any form of judgement aggregation over a sufficiently complex agenda fails to meet all the described desiderata together. This poses a challenge for the aggregation of judgements: Either judgement aggregation fails (for some profiles), or one has to argue that at least one of the desiderata can and should be relaxed in order to avoid the impossibility result.[2]

Returning to the examples of the discursive dilemma above, I will now discuss four procedures to arrive at collective judgements: the majority vote on each proposition, the premise- and the conclusion-based procedure, and a dictatorship. The majority vote was already mentioned in the introduction of the discursive dilemma. If the collective votes on all propositions with simple majority, the group may end up with an inconsistent judgement set. This is unsatisfactory, and several ways to avoid this result have been proposed. The majority vote on all propositions satisfies universal domain, anonymity and systematicity, but fails to produce fully rational judgement sets for all logically possible judgement profiles.

The premise-based procedure divides the propositions on the agenda into two sets: the premises and the conclusion(s). A majority vote is taken on each premise, and the premises adapted by these votes determine the remaining propositions, i. e. the conclusions, by deductive closure. For the discursive dilemma stated in table 1, P and $P \rightarrow Q$ can be taken as premises, Q as the conclusion. The majority adopts both premises, and deduces that Q must also be true. It therefore reaches the collective judgement set $\{P, P \rightarrow Q, Q\}$.[3] More

---

2 The literature on judgement aggregation has produced many refinements and extensions to List's and Pettit's 2002 result, which cannot be described in detail here. Most important is perhaps Pauly and van Hees's (2006) generalizations, and further more general results in Dietrich and List (2007). The general structure of these additions is to discuss other, often weaker or differently constructed desiderata and prove impossibility (and sometimes possibility) results for aggregation functions. A very clear framework for judgement aggregation in general logic is provided by Dietrich (2007).

3 It is not always the case that the propositions can be neatly divided into premises and conclusions. In addition, the premises do not necessarily determine the truth value(s) of the conclusion(s). For instance, if

loosely speaking, the premise based procedure means: vote on the premises, deduce the conclusion. The premise based procedure usually produces fully rational collective judgement sets[4], but it violates systematicity because the collective judgement on the conclusion does not only depend on the individual judgements regarding the conclusion.

In the MI5 example, the worry with the premise based procedure is that it overrules the majority vote. The second and perhaps more obvious procedure to the aggregation problem is to disregard the majority vote on the premises and only vote on the conclusion. This is the conclusion based procedure. For table 1 it leads the collective to adopt not-*Q*. Note that the collective does not take any view on the premises according to the conclusion-based view. Therefore, the conclusion-based procedure fails to produce complete collective judgement sets.

Another procedure to avoid collective inconsistency is to nominate a dictator, that is a person whose individual judgement set fully determines the collective judgement set. For instance, one could stipulate that the group always adopts the judgements of individual 1. Since the individual judgement sets are complete, consistent, and deductively closed, the "collective" judgement set will be, too. A dictatorship is a blatant violation of the anonymity condition, because a reshuffling of individuals (in particular, changing the dictator) may change the outcome.

Table 3 compares the four aggregation procedures. None of the procedures meets all the desiderata and the requirement of collective rationality (completeness, consistency, and deductive closure) together. List's and Pettit's theorem shows that there is in fact *no* aggregation procedure that can meet all these desiderata together. It is therefore necessary to engage in a normative debate as to which desiderata should be sacrificed, or at least relaxed, to find a working aggregation procedure.

---

the votes on the premises had resulted in $\{\neg P, P \rightarrow Q\}$, the conclusion Q would not be determined by deductive closure because both Q and $\neg Q$ are consistent with the judgements on the premises.

4 Except for those cases described in note 3.

Procedure	Universal Domain	Anonymity	Systematicity	Collective Rationality
Majority rule	+	+	+	–
Dictator	+	–	+	+
Premise-based	+	+	–	+
Conclusion-based	+	+	+*	–

Table 3**:** Aggregation procedures in comparison (* systematicity holds for the conclusions).

The desiderata under discussion are motivated by a broadly "democratic" set of values (see for instance List, 2006). Universal domain can be normatively attractive from a democratic perspective because a democratically governed group should not rule out rational individual judgements ex ante. Anonymity can be attractive because it ensures that every member of the group has the same level of influence over the collective result. Systematicity ensures an equal treatment of all propositions, so that the aggregation procedure does not have an ex ante bias to define some propositions as "special" or "more important". List and Pettit (2002) discuss several options to relax one of the three desiderata, or one of the three rationality conditions completeness, consistency, and deductive closure. Relaxing collective consistency and deductive closure is unattractive, because it results in irrational collective judgement sets. Other options are more attractive, depending on the circumstances. Relaxing universal domain is plausible when the individuals tend to have judgement profiles that are "well-behaved", that is do not give rise to the discursive dilemma. Relaxing anonymity may in particular be justified when the competence in the group is unevenly distributed (List, 2006). Relaxing systematicity is perhaps the most attractive move, because the idea that the collective judgements on different propositions do not influence each other appears implausible for a set of logically connected propositions in the first place. Even more implausible, systematicity also demands that all propositions are treated exactly equal in that regard. If a group deliberates on a number of dependent propositions, it should not be ruled out ex ante that the change of individual opinions on a proposition Φ can change the collective judgement on another proposition Ψ , even if the individual judgements on Ψ have not changed. Neither should it be ruled out that the same pattern of individual judgements for Φ and for Ψ can lead to different collective judgements on Φ and Ψ .

When considering distributed thinking systems, the background set of values to decide on an aggregation procedure does not necessarily have to be "democratic". But the properties one would like to see in a judgement aggregation function for distributed thinking may be similar. Universal domain is desirable from a distributed thinking perspective because the thinking process should not break down for certain inputs. Whether anonymity and systematicity are normatively desirable properties of a distributed thinking system is less clear. Anonymity is attractive if every thinking unit in the system of distributed thinking should be treated equally.[5] In the same vein, systematicity may be

5 Also, relaxing anonymity does not yield particularly attractive aggregation procedures. In a very closely related setup, Pauly and van Hees (2006) show that the only aggregation procedure that meets all other desiderata is a dictatorship.

important when all propositions on the agenda should be treated equally.

Even if we relax one or more of the desiderata, we still need to say how we relax these desiderata and which aggregation functions we want to use. One important criterion for an aggregation function is that it meets the “knowledge challenge”. This means that the aggregation function should be good at pooling the individual information that is distributed among individuals to reach correct outcomes. To explore the knowledge challenge in greater detail, I discuss the truth tracking performance of different aggregation functions.

3 An Epistemic Perspective

When voting on the truth or falsity of a single proposition, the Condorcet Jury Theorem shows that large groups can be almost always correct, as long as each member of the group is just slightly better than random at identifying the correct choice. Assume there is one correct state of the world, which is either that Φ or not-Φ is correct (or the better alternative). The competence assumption of the Condorcet Jury Theorem states that all individuals have a competence greater than 0.5. The competence of an individual is the probability to choose the correct alternative. With a competence greater than 0.5 the individuals are better than random in making the correct judgement between two alternatives.

The Condorcet Jury Theorem tells us: If all individuals have the same level of competence greater than 0.5, if their votes are independent[6], and if they do not misrepresent their personal judgements for strategic reasons, then large groups will almost certainly choose the correct alternative in a majority vote.[7] The pooling of the individual competence in the vote renders the group much more competent than each single individual.

Let there be n ndividuals (with n being odd to avoid ties), and let the probability of all the different individuals 1 to n be p , with $p > 0.5$. The probability of a group to choose the correct alternative is (Grofman, Owen and Feld, 1983):

$$P^{CJT}(n,p) = \sum_{h=(n+1)/2}^{n} \binom{n}{h} p^h (1-p)^{n-h}. \quad (1)$$

Table 4 shows the group competence for some levels of individual competence and different group sizes. One can see that even for relatively small levels of competence like 0.55, large groups reach a group competence of almost 100%. Therefore, if the conditions of the Condorcet Jury Theorem hold, groups can be excellent “truth trackers” in dichotomous choice situations.

6 More precisely, if the votes are probabilistically independent, conditional on the truth value of Φ.

7 The joint assumption of competence and independence rarely holds in practice. Weaker versions of the theorem have been proved. Dietrich (2008) points out that it is not possible to (statistically) justify both the independence and the competence assumptions and discusses less demanding assumptions and their implications.

n | p	0.501	0.51	0.55	0.6	0.7	0.8	0.9
11	0.503	0.527	0.633	0.753	0.922	0.988	≈ 1
101	0.508	0.580	0.844	0.979	≈ 1	≈ 1	≈ 1
1001	0.525	0.737	0.999	≈ 1	≈ 1	≈ 1	≈ 1

Table 4**:** Group competence according to formula 1.

The graph in figure 1 shows how the group competence develops for different group sizes and different values of *p*. One can see how larger groups quickly approach high competence if p >0 .5, but approach a group competence of 0 for $p < 0.5$.

*** Figure 1 about here ***

The Condorcet Jury Theorem is the starting point to analyse richer collective decision problems. For the problems of judgement aggregation discussed above, each single proposition is a dichotomous choice problem, but the judgement aggregation problem as a whole is more complex. We have seen that there are different aggregation procedures, each with advantages and drawbacks. One possible normative criterion to decide for one aggregation procedure is to consider its epistemic performance, that is its ability to "track the truth". Here I focus primarily on a comparison between the conclusion and the premise based procedure, in line with discussions in Bovens and Rabinowicz (2006) and List (2006).

If a group follows the conclusion based procedure, it simply votes on the conclusion, and disregards the premises. If the group follows the premise based procedure, it votes on the premises and derives the conclusion by deductive closure. This will lead to different epistemic performances. I will show the diverging epistemic performances by discussing the detective example as stated in table 2 above. The three proposition P , Q, R and their respective negations are on the agenda. In addition, all individuals accept $(P \wedge Q) \leftrightarrow R$ as true[8], and we assume it is true as a matter of fact. Therefore, the world can be in 4 different states:

[8] Assuming that the normative proposition R refers to a fact.

S1	P	Q	R
S2	$\neg P$	Q	$\neg R$
S3	P	$\neg Q$	$\neg R$
S4	$\neg P$	$\neg Q$	$\neg R$

Less technically, all propositions can be true, or one of the premises and the conclusion are false, or both premises and the conclusion are false. These are also the only logically possible complete, consistent and deductively closed judgement sets, since we accept $(P \wedge Q) \leftrightarrow R$ as a background assumption.

For now, let us assume that a decision is epistemically correct if and only if it produces the correct stance on conclusion R (we discuss the idea that the stances on the premises should also be true below). In a premise based procedure one votes only on the premises and deduces the conclusion. Therefore, the correct conclusion can be reached with different collective judgements on the premises:

State	**Conclusion**	**Premise judgements with correct conclusion**
S1	R	$\{P, Q\}$
S2	$\neg R$	$\{P, \neg Q\}, \{\neg P, Q\}, \{\neg P, \neg Q\}$
S3	$\neg R$	$\{P, \neg Q\}, \{\neg P, Q\}, \{\neg P, \neg Q\}$
S4	$\neg R$	$\{P, \neg Q\}, \{\neg P, Q\}, \{\neg P, \neg Q\}$

The point to note here is that the premise based procedure can lead to the right conclusion even if one or both collective judgements on the premises are wrong. One can therefore be right for the wrong reasons. For instance, an agent can have the judgements P and $\neg Q$ and therefore $\neg R$, even though the world is in state where $\neg P$ and Q are true. The agent is right to hold $\neg R$, but for the wrong reasons. If one wants the group to be right for "the right reasons" (Bovens and Rabinowicz, 2006, p. 138f.), one should only consider cases where the collective judgements on both premises are correct, not only the conclusion derived from them.

I now turn to the conclusion based procedure. There are two distinct ways for individuals to deal with a conclusion based system. Either each single individual takes their judgements on the premises and derives the conclusion. This is the way Bovens and Rabinowicz propose. The conclusion based procedure leads to a correct judgement on the conclusion if and only if a majority of individuals has the correct assessment of the conclusion. However, they may well have come to that assessment for the wrong reasons. For instance, if the correct assessment of the conclusion is that *R* is false, one can arrive at that conclusion from three different judgement sets on the two premises: $\{P,\neg Q\},\{\neg P,Q\},\{\neg P,\neg Q\}$. Only one set of judgement can be the right one, but all lead to the correct judgement on the conclusion. Alternatively, the agents completely disregard their judgements on the premises and make judgements only on the conclusion. In this case, the decision problem is collapsed into a decision on a single proposition, and the standard Condorcet Jury Theorem formula (1) applies. This way is unattractive from an epistemic standpoint because it completely disregards the information the individuals have on the premises.

Bovens and Rabinowicz calculate the probabilities for the group to make the right judgement on the conclusion, They consider four cases:

1. The use of the premise based procedure where all correct conclusions are counted.
2. The use of the premise based procedure where only judgements based on the right reasons are counted as correct.
3. The use of the conclusion based procedure where all correct conclusions are counted.
4. The use of the conclusion based procedure where conclusions are counted as correct if a majority of individuals has reached the correct judgement on the conclusion for the right reason.

These calculations depend on parameters. In addition to the individual competence *p* and the group size *n*, it also matters how likely the different states S1 to S4 are, which is determined by the prior probabilities of *P* , *Q*, and *R* . Let $\pi(P)$ be the prior probability that *P* is true; $\pi(Q)$ be the prior probability that *Q* is true. This in turn determines $\pi(R)=\pi(P)\pi(Q)$.

Figure 2 shows the results for the group competence, dependent on the individual competence *p*, for *n* = 101 , $\pi(P)=\pi(Q)=0.5$ and $\pi(R)=0.25$. The two solid curves are the results for the premise based procedure, the two dashed curves for the conclusion based procedure. pbp is the result for the premise based procedure, pbp-rr for the premise based procedure when only results with the right reasons are counted as correct. Similarly, cbp shows the result for the conclusion based procedure, cbp-rr the conclusion based procedure with the correctness for the rights reason criterion.

*** Figure 2 about here ***

First, consider the results for $p \geq 0.5$, that is the results with the (usually) more plausible assumption that individuals tend to be at least as good as a coin toss in making their decisions. For the premise based procedure, the group competence is 0.5 for $p = 0.5$, and then quickly approaches 1 for larger p . If being right for the right reason matters, the group competence starts from a lower level, but still approaches 1 quickly. The conclusion based procedure starts from a higher level (0.75), but is quickly outperformed by the premise based procedure at a competence level of around 0.55. Interestingly, which procedure performs better depends on the level of p. Unsurprisingly, the procedures that care for being right for the right reasons have lower levels of collective competence. The premise based procedure for the right reasons performs better than the conclusion based procedure for the right reasons.

Now consider the results for competence levels lower than 0.5. First, there is a range of p for which the conclusion-based procedure fares better than the premise based procedure. Second, the reliability of the premise based procedure dips for values that are close but below 0.5 for p.[9] Third, if individuals are incompetent, they are very unlikely to be right for the right reasons. Overall, results for competence levels of $p < 0.5$ are of less interest because it is implausible that individuals are systematically worse than a toin coss.

Figure 3 shows the results for the same group size, but with different prior probabilities, namely $\pi(P) = \pi(Q) = 0.8$ and consequently $\pi(R) = 0.64$. For these parameter values, the premise based procedure does better for all values $p > 0.5$. One can see that the performance of the two procedures depends on the prior probabilities. Both procedures perform worse around $p = 0.5$ compared to figure 2, but the conclusion based procedure is still stronger in an area below 0.5.

*** Figure 3 about here ***

List (2006, n. 25) criticizes the approach taken by Bovens and Rabinowicz because they do not distinguish between positive and negative reliability.[10] Positive reliability is the probability that the group correctly identifies R as true, negative reliability the probability that the group correctly identifies R as false. Different decision problems require different attention to the two reliabilities. Bovens and Rabinowicz simply calculate the probability that the group is correct. This may be misleading. Intuitively, this can be seen in figure 2 by considering the performances of the different aggregation procedures with $p = 0.5$, i. e.

9 This feature of the premise based procedure has been overlooked by Bovens and Rabinowicz (see figure 6, where this dip is missing). The reason for this dip is quite easy to grasp intuitively: For very low p, the premise based procedure is reliably wrong on both premises. If the world is in state S1 or S4, it will produce the wrong judgement on R, but if the world is in S2 or S3, it will produce the right outcome (though for the wrong reason, swapping the true and the false premise). As p approaches the watershed of 0.5, the procedure is less reliable false. It is still very unlikely that it is correct about both premises, but it is occasionally correct on one of them. Being sometimes right on one conclusion produces better results if the world is in S4, but worse results if the world is in either S2 or S3 (and it does not matter for S1). Since the world is more often in either S2 or S3 than in S4, the performance of the premise based procedure dips for p close to but lower than 0.5.

10 List also operates with asymmetrical individual competence, that is individuals have different competence for correctly judging true and false propositions.

when the individual competence is no better than a coin toss. An unbiased procedure should then be able to pick out the right result in half of the cases. But the conclusion based procedure is doing much better. This is because the conclusion based procedure has a bias towards assuming that R is false. Since figure 2 is drawn for $\pi(R) = 0.25$, this bias plays to the advantage of the conclusion based formula. This result is due to a high negative reliability and the fact that for the given prior probabilities the conclusion is more often false than true. However, this comes at the cost of a low positive reliability. Figure 3 shows that the premise based procedure is also biased for other parameter setting. Here both procedures show a bias that is to their disadvantage.

The upshot of the epistemic analysis is that different procedures for aggregating judgements have different qualities to "track the truth". These considerations show that a formal analysis of Goldman's "knowledge challenge" can help to decide which aggregation rule to use. For the example analysed here, the premise based procedure performs well in most situations where individuals are competent. With regard to distributed thinking more generally, it is worthwhile exploring with formal models how different systems of distributed thinking lead to different epistemic success.

4 Distributed and Consistent Thinking

Distributed thinking can proceed in different ways. One way to conceptualize a distributed thinking system is to imagine a system where distributed non-thinking parts are connected in such a way that the whole assembly is a thinking system. A computer may be a distributed thinking system in that weak sense. Each single transistor could be seen as a non-thinking part, while the computer arranges these non-thinking parts in such a way that it can think, where thinking is taken as being able to solve logical problems. This notion of a thinking system is too weak because *any* thinking system is distributed in that sense. Brains, for instance, could be seen a distributed thinking system made of neurons.

The definition becomes more interesting if we assume that a distributed thinking system consists of several *thinking* sub-units. This definition is better because it rules out single computers and (perhaps) single brains, but includes relevant cases like groups of several agents, networked systems, et cetera. The interesting aspect of a distributed thinking system defined like that is the potential tension between the individual and the collective thinking. Oftentimes this tension is productive. We talk (rather vaguely) of "swarm intelligence" or "collective intelligence", and we sometimes experience how group deliberation can lead to better, more informed results than decisions by single individuals. But this tension can also lead to breakdowns of "collective intelligence", when no agreement can be reached, when the outcomes are inconsistent, or just plain wrong.

I have argued that judgement aggregation provides a useful framework for the analysis of distributed thinking. However, two anonymous referees argued that the judgement aggregation framework does not connect with the concept of distributed thinking for at least three reasons. First, judgement aggregation is *not dynamic*, in contrast to cognitive distribution, which emphasizes the dynamic interaction between the thinking units. Second, judgement aggregation does not engage with a central feature of the distributed cognition framework: the fact that *minds and world interact*, and that organisms reshape their environment in order to solve cognitive problems. This claim is often referred to by claiming that cognition happens “in the wild”. Third, the judgement aggregation framework allegedly attempts to reduce distributed thinking to the thinking of sub-units, and does not appreciate that distributed thinking arises because *higher level structures emerge*.

My response is as follows. I largely concur with the first claim regarding the discursive dilemma, but point out that research in judgement aggregation raises interesting question about the possible dynamics that avoid the described impossibility results. This answer is connected with the second claim. While judgement aggregation *per se* does not address the interaction between minds and environment, it does raise questions as to how agents restructure their decision environment in order to avoid paradoxes like the discursive dilemma. Finally, I maintain that the validity of the third claim depends on the notion of emergence employed. In a weak sense, judgement aggregation and social choice theory support the claim that distributed thinking systems have emergent properties. I will now address each objection in greater detail.

Judgement aggregation, at least in the simple versions discussed here, does not incorporate a dynamic change of judgements through an interaction of individual and group judgements.[11] But the question of dynamics is raised *indirectly* by the impossibility results mentioned above, since the impossibility results pose the question how the breakdown of the aggregation process is avoided in practice. The discursive dilemma, for example, only arises for some of the many possible judgement profiles. It is therefore conceivable that a dynamic process, especially a process of deliberation, reduces or eradicates those profiles that lead to impossibility results. It is well known that the Arrow paradox can be avoided if the preference profile has certain structural properties, thereby relaxing the universal domain axiom (Dryzek and List, 2003; Black, 1948). Similar results hold for judgement aggregation. In case of the discursive dilemma, a suitable restriction of the universal domain axiom avoids the impossibility result (List and Pettit, 2002). Empirical observations support the claim that deliberation leads to fewer occurences of the discursive dilemma (List et al., 2006).[12] Thus, a dynamic process like deliberation may mitigate the occurence of the impossibility result, and the framework of judgement aggregation raises interesting questions about the nature of the dynamic processes to avoid a breakdown of collective rationality. I therefore claim that even a static analysis in terms of judgement aggregation provides the debate on distributed thinking with useful concepts to analyse the dynamic processes.

The charge that judgement aggregation fails to scrutinize cognition “in the wild” can also be addressed by considering the escape routes to avoid impossibility results. Hutchins (1995) discusses several ways of how groups can structure their own decision making to simplify it, among them hierarchy and consensus (p. 256–259). Clark also emphasizes the importance of “broader social and institutional contexts of action” (p. 186). List and Pettit show that if the individuals agree on a unidimensional alignment of the problem (similar to

11 However, research on how the judgement aggregation framework pertains to the change of judgements is undertaken. See List, forthcoming*b*.

12 In addition,Bonnefon (2007) reports that individuals change their preference for the conclusion and premise based procedure changes with the nature of the decision.

a left-right dimension in party politics), the dilemma can be avoided, even though individuals can disagree on their judgements. In addition, the dilemma disappears when the decision is delegated to specialists for each proposition (a form of "local dictatorship"), or when deliberation leads to a convergence of judgements. Thus, even though judgement aggregation does not directly explore decision group thinking "in the wild", the discussion of escape routes is very much concerned with the the dynamic interaction of individuals and their potential to restructure the decision problem.

Finally, I turn to the charge that judgement aggregation is reductionist and fails to do justice to the emergent properties of distributed thinking systems. This charge hinges on the notion of emergence and reduction used. It is true that judgement aggregation is interested in aggregating individual to collective judgements. But the central result of the judgement aggregation research programme is that the aggregation is non-trivial, and that the group judgements cannot just be derived by summing up and counting the individual judgements. To underline this point, I use William Wimsatt's work on emergence and reduction. Wimsatt (1997) proposes a weak working definition for emergence: a system has emergent properties if it fails to be aggregative. For Wimsatt, the ideal aggregative system is invariant with regard to changes of like-for-like components, it scales linearly in size, the system properties are invariant with regard to a decomposition or reaggregation of the system, and there is no positive or negative interaction among the parts of the system. For instance, a heap of sugar is aggregative with regard to its mass. I can exchange one gramm of sugar for another gramm and its mass remains the same. If I add 1 gramm of sugar, the total mass increases by 1 gramm. If I divide the heap of sugar in two piles, the two piles each have half the mass of the original heap. If I put the heaps together again, I obtain the same mass. Finally, if I had two different types of sugar (brown and white sugar, say), this would not lead to positive or negative interactions in terms of the mass of the two types. Most systems are not entirely aggregative. For Wimsatt, the less aggregative a system is with regard to its properties, the more emergent properties it has.

Since the results presented above show that judgements on logically connected propositions cannot always be aggregated, given the stated axioms, such a system has emergent properties in Wimsatt's weak sense. One central result of the judgement aggregation research programme is that the sentence "A collective judgement of a group on a set of logically connected propositions is nothing but the aggregation of individual judgements" is not trivially true, since the aggregation encounters impossibility results. The results from judgement aggregation thus casts doubt on a simple "nothing but" reduction of group judgements, and weak emergence in Wimsatt's sense is embraced. For Wimsatt, "[a]n emergent property is—roughly—a system property which is dependent on the mode of the organization of the system's parts" (1997, p. S373, italics omitted). In this sense, the process of judgement aggregation has at least weak emergent properties. Whether this weak notion of emergence is enough to be of interest for the distributed cognition framework is a further question I leave to others. But I agree with Poirier and Chicoisne (2008) that the borders of distributed cognition are fuzzy.

Judgement aggregation as a field (in the simple treatments as discussed above) shows that even very simple reasoning processes run into difficulties when trying to turn rational individual judgements on logically connected propositions to rational collective judgements. If these problems arise even for the fairly simple problems like the discursive dilemma, one can anticipate similar and more difficult problems once one moves to more complicated settings. The basic lesson from the discursive dilemma is that the decision on the best processes applied in distributed thinking involves trade-offs between different properties of the reasoning process. Some processes are clearly worse than others, but when it comes to the best processes, different considerations need to be weighed against each other.

One possible consideration is the epistemic success of the procedures, i. e. the ability of the distributed thinking system to “track the truth”. It is interesting to note that, for instance, in the comparison of the premise and the conclusion based procedure, it depends on the context of the decision problem which procedure performs best. However, if we introduce the additional requirement that the procedure must reach the correct decision for the right reasons, then the premise based procedure is the clear winner in the example discussed. Being right for the right reasons can also be important if the group has to justify its decisions, or if the reasoning the group applies will be adopted or imitated in future reasoning processes.

Many extensions of the simple examples discussed in this paper are possible. One should explore more complex decision problems, different logical dependencies, cases with incomplete judgement sets, the heterogeneous competence levels, or settings where certain types of judgement errors are worse than others. Most of these questions have already been addressed in the literature on judgement aggregation and information pooling. The emerging literature on distributed thinking can benefit from the analytical and normative debates in these areas.

Appendix

Bovens and Rabinowicz (2006) calculate the probabilities of the group being correct, conditional on the state. They define M^{pbp} as the proposition ‘The premise based procedure yields the correct result’ and calculate probabilities conditional on all 4 states:

$$\begin{aligned}
&P(M^{pbp} \mid \mathrm{S1}) = P^{CJT}(n,p)^2, \\
&P(M^{pbp} \mid \mathrm{S2}) = Pr^{CJT}(n,p)^2 + P^{CJT}(n,p)(1-P^{CJT}(n,p)) + (1-P^{CJT}(n,p))^2, \\
&P(M^{pbp} \mid \mathrm{S3}) = P(M^{pbp} \mid \mathrm{S2}), \\
&P(M^{pbp} \mid \mathrm{S4}) = P^{CJT}(n,p)^2 + 2P^{CJT}(n,p)(1-P^{CJT}(n,p)).
\end{aligned} \qquad (2)$$

Note that one can arrive at the correct result even though some or even both collective judgements on the premises are wrong. Given the logical dependency between the propositions, we know that $\pi(R) = \pi(P)\pi(Q)$. Summing up the conditional probabilities of being correct with the premise based procedure, weighted by the probabilities that the different states obtain yields:

$$P(M^{pbp}) = P(M^{pbp} \mid \mathrm{S1})\pi(P)\pi(Q) + P(M^{pbp} \mid \mathrm{S2})(1-\pi(P))\pi(Q)$$
$$+P(M^{pbp} \mid \mathrm{S3})\pi(P)(1-\pi(Q)) + P(M^{pbp} \mid \mathrm{S4})(1-\pi(P))(1-\pi(Q)). \quad (3)$$

Following Bovens and Rabinowicz's exposition for the conclusion based procedure, let V be the proposition that a single voter determines the conclusion correctly, and $P(V)$ the probability the voter does so. Since each single voter applies deductive closure, we obtain the following probabilities for each single voter to be correct on the conclusion, based on their competence p:

$$P(V \mid \mathrm{S1}) = p^2$$
$$P(V \mid \mathrm{S2}) = P(V \mid S3) = p^2 + p(1-p) + (1-p)^2 \quad (4)$$
$$P(V \mid \mathrm{S4}) = p^2 + 2p(1-p).$$

Each individual can reach the correct conclusion by being correct on both premises (probability p^2) but one can also be correct, even if one is wrong on one or even both of the premises. Let M^{cbp} denote the proposition that the conclusion based procedure yields the correct result. Conditional on the state, we can apply equation 1 to calculate the probability of a correct majority vote on the conclusion:

$$P(M^{cbp} \mid \mathrm{S}i) = P^{CJT}(n, P(V \mid \mathrm{S}i)). \quad (5)$$

Summing up the probabilities weighted by the prior probabilities of the different states yields:

$$P(M^{cbp}) = P(M^{cbp} \mid S1)\pi(P)\pi(Q) + P(M^{cbp} \mid S2)(1-\pi(P))\pi(Q)$$
$$+P(M^{cbp} \mid S3)\pi(P)(1-\pi(Q)) + P(M^{cbp} \mid S4)(1-\pi(P))(1-\pi(Q)). \quad (6)$$

The results for the premise based procedure in (6) and the conclusion based procedure in (11) are based on the assumption that it does not matter whether the correct result is deduced from correct or incorrect judgements on the premises. If we want to be right "for the right reasons", the cases where incorrect judgements lead to correct outcomes need to be removed. Let M^{pbp-rr} denote the proposition that the group has arrived at the right judgement for the right reasons. This yields:

$$P(M^{pbp-rr}) = P^{CJT}(n, p)^2 \quad (7)$$

Similarly, for the conclusion based procedure one want to consider the probability that a majority of voters is correct for the right reasons:

$$P(M^{cbp-rr}) = P^{CJT}(n, p^2).$$

Acknowledgements

An earlier version of this paper was presented at the "Distributed Thinking" workshop at the University of Hertfordshire, organized by the Distributed Language Group. I am grateful for many constructive comments. I owe special thanks to Hemdat Lerman for several helpful discussions and extensive comments. The paper benefitted from the reviews provided by three anonymous referees. All remaining errors are my own.

References

Arrow, Kenneth J (1963) Social choice and individual values. 2nd. rev. ed. Wileyt, New York/London.

Black, Duncan (1948) On the Rationale of Group Decision-Making. Journal of Political Economy 56(1):23–34.

Bonnefon, Jean-Francois (2007) How Do Individuals Solve the Doctrinal Paradox in Collective Decisions. Psychological Science 18(9):753–755.

Bovens, Luc and Wlodek Rabinowicz (2006) Democratic Answers to Complex Questions - an Epistemic Perpective. Synthese 150:131–153.

Clark, Andy (1997) Being There: Putting Brain, Body, and World Together Again. MIT Press, Cambridge (MA) and London.

Clark, Andy and David Chalmers (1998) The Extended Mind. Analysis 58(1):7–19.

Dietrich, Franz (2007) A generalised model of judgment aggregation. Social Choice and Welfare 28(4):529–565.

Dietrich, Franz (2008) The premises of Condorcet's jury theorem are not simultaneously justified. Episteme 58(1):56–73.

Dietrich, Franz and Christian List (2007) Arrow's theorem in judgment aggregation. Social Choice and Welfare 29(1):19–33.

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

Goldman, Alvin I (2004) Group Knowledge versus Group Rationality: Two Approaches to Social Epistemology. Episteme 1:11–22.

Grofman, Bernard, Guillermo Owen and Scott L. Feld (1983) Thirteen Theorems in Search of the Truth. Theory and Decision 15:261–278.

Hutchins, Edwin (1995) Cognition in the Wild. MIT Press, Cambridge (MA).

Kornhauser, L. A. and L. G. Sager (1986) Unpacking the Court. Yale Law Journal 96:82–117.

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

List, Christian (2006) The Discursive Dilemma and Public Reason. Ethics 116(2):362–402.

List, Christian (forthcoming*a*) Collective wisdom: a judgment aggregation perspective. In: Hélène Landemore and Jon Elster (ed.) Collective Wisdom: Principles and Mechanisms.
URL: *http: //personal.lse.ac.uk/list/pdf-files/CollectiveWisdom19Oct.pdf*

List, Christian (forthcoming*b*) Group communication and the transformation of judgments: an impossibility result. Journal of Political Philosophy.

List, Christian and Clemens Puppe (2009) Judgment aggregation: a survey. In: Paul Anand, Clemens Puppe and Prasanta Pattanaik (eds) Oxford Handbook of Rational and Social Choice, Oxford University Press, Oxford.

List, Christian and Philip Pettit (2002) Aggregating Sets of Judgments: An Impossibility Result. Economics and Philosophy 18(1):89–110.

List, Christian, Robert Luskin, James Fishkin and Iain McLean (2006) Deliberation, Single-Peakedness, and the Possibility of Meaningful Democracy: Evidence from Deliberative Polls. manuscript. Available via *http://personal.lse.ac.uk/list/PDF-files/DeliberationPaper.pdf*

Pauly, Marc and Martin van Hees (2006) Logical Constraints on Judgement Aggregation. Journal of Philosophical Logic 35(6):569–585.

Pigozzi, Gabriella (2006) Belief merging and the discursive dilemma: an argument-based account to paradoxes of judgment aggregation. Synthese 152(2):285–298.

Poirier, Pierre and Guillaume Chicoisne (2008) A Framework for Thinking About Distributed Cognition. In: Itiel E. Dror and Stevan R. Harnad (eds.) Cognition Distributed, John Benjamins, Amsterdam.

Vygotsky, Lev S (1962) Thought and Language. MIT Press, Cambridge, MA.

Vygotsky, Lev S (1978) Mind in Society. Harvard University Press, Cambrigde, MA.

Wimsatt, William C (1997) Aggregativity: Reductive Heuristics for Finding Emergence. Philosophy of Science 64:S372–S384.

Figures [also delivered in high quality encapsulated postscript format for production]

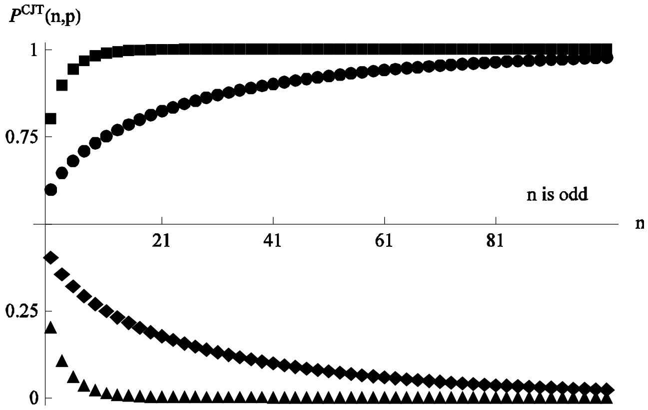

Figure 1**:** Group competence $P^{CJT}(n, p)$ as a function of group size *n*, for different levels of individual competence *p*.

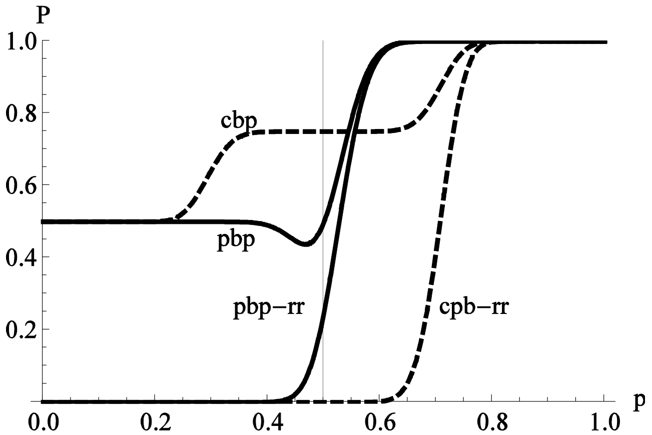

Figure 2**:** Results for the premise (pbp) and conclusion based procedure (cbp) with *n* = 101, $\pi(P) = \pi(Q) = 0.5$. rr signifies the results is for the “rights reasons” constraint. The x-axis shows individual, the y-axis collective competence.

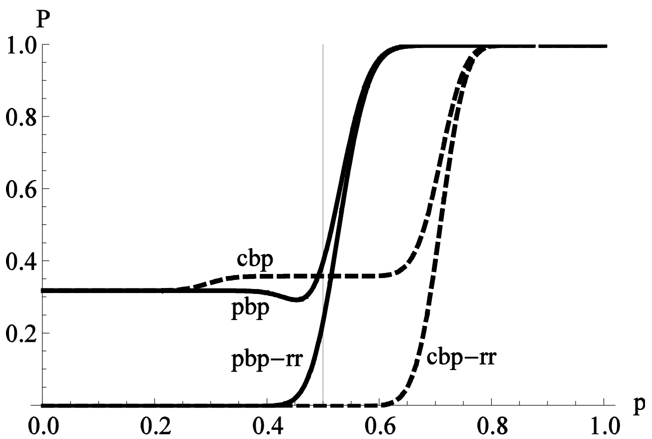

Figure 3**:** Results for the premise (pbp) and conclusion based procedure (cbp) with *n* = 101, $\pi(P) = \pi(Q) = 0.8$. rr signifies the results is for the “rights reasons” constraint. The x-axis shows individual, the y-axis collective competence.