

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

## Consensual Decision-Making Among Epistemic Peers

**This is a pre print version of the following article:**

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/1662633> since 2020-05-06T15:41:34Z

*Published version:*

DOI:10.3366/E1742360009000598

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

# Consensual Decision-Making Among Epistemic Peers

Stephan Hartmann<sup>\*</sup>, Carlo Martini<sup>†</sup> and Jan Sprenger<sup>‡</sup>

March 5, 2009

## Abstract

This paper focuses on the question of how to resolve disagreement, and uses the Lehrer-Wagner model as a formal tool for investigating consensual decision-making. The main result consists in a general definition of when agents treat each other as epistemic peers (Kelly 2005; Elga 2007), and a theorem vindicating the “equal weight view” to resolve disagreement among epistemic peers. We apply our findings to an analysis of the impact of social network structures on group deliberation processes, and we demonstrate their stability with the help of numerical simulations.

## 1 Introduction

Imagine the following scenario: the director of a company asks her advisory board to estimate the number of sales a planned new product will achieve on the market. On the basis of this assessment she will decide whether it is worth the risk to develop the new product. All board members have a certain amount of evidence at their disposal; they may, for example, be acquainted with recent marketing studies or know of similar attempts in the past. They may also take more indirect evidence into account, such as the general economic prospects. Each board member will then accumulate all her

---

<sup>\*</sup>Tilburg Center for Logic and Philosophy of Science, Tilburg University, PO Box 90153, 5000 LE, Tilburg (The Netherlands) - <http://www.stephanhartmann.org> - [S.Hartmann@uvt.nl](mailto:S.Hartmann@uvt.nl).

<sup>†</sup>Tilburg Center for Logic and Philosophy of Science, <http://www.martinicarlo.net>, [C.Martini@uvt.nl](mailto:C.Martini@uvt.nl).

<sup>‡</sup>Tilburg Center for Logic and Philosophy of Science, <http://www.laeuferpaar.de>, [J.Sprenger@uvt.nl](mailto:J.Sprenger@uvt.nl).

evidence and arrive at a judgment on the expected number of sales. Note that the assessment of the evidence will generally vary from board member to board member and depend, to a large extent, on subjective factors. Thus, it will not come as a surprise that the initial judgments may differ. But how shall the board proceed? After all, it has to come up with a value (or at least an interval) for the expected number of sales.

We investigate such group decisions from an epistemic perspective. In the above example, the problem consists in updating individual beliefs in the face of disagreement with the other group members. The question is whether the group members attain a *consensus*, i.e. whether they revise their beliefs such that, at the end, they all agree on the expected number of sales. This would resolve the initial disagreement and determine a decision whether or not to develop the product. We assume that the agents are truth-orientated and don't block a consensus for egoistic reasons, e.g. because a certain decision serves their own career. Certainly, this assumption is an idealization, but at least from a normative perspective, it is not unrealistic because many expert panels, political committees etc. are supposed to give *correct* judgments instead of enforcing decisions that serve their private interests.

The question of whether and how to resolve disagreement among epistemic peers (agents that are equally likely to come up with a correct answer) has recently triggered much attention in the literature (Christensen 2007; Elga 2007, 2009; Feldman 2007; Kelly 2005, 2009). The debate exposes two characteristic traits: first, it focuses on whether compromising ("meeting in the middle"), suspending judgment or giving in to the other agents' opinions is the rational decision. Second, the debate makes little use of formal methods or models. In this paper, we provide a general and formally tractable account of when agents take each other to be epistemic peers. Moreover, we study the effects of certain *social network structures* on the decision that the group reaches. The formal framework for our analysis is provided by the *Lehrer-Wagner model* (Lehrer 1976; Lehrer and Wagner 1981; Wagner 1978), which was developed as a normative model for rational group decision-making. Our analysis presents various formal results on how to resolve disagreement, and elucidates the role of experts and network structures in group decision-making.

The remainder of this paper is organized as follows: Section 2 introduces the Lehrer-Wagner model, and discusses some conceptual issues about applying it to group decision making. Sections 3 and 4 present our analytical and conceptual results while section 5 explores their stability with the help of numerical simulations. Section 6 summarizes our insights.

## 2 The Lehrer-Wagner Model

The problem of group decision-making can be seen from at least two perspectives: social choice and social epistemology. In social choice, agents submit a set of preferences or judgments which have to be aggregated into a collective preference or a collective judgment. In other words, the problem is to find an aggregation function from the set of individual preferences/judgments which satisfies a number of reasonable constraints, such as strategy-proofness, non-dictatorship, anonymity, etc. The aim of such an aggregation method is not to convince anyone of the adequacy of the group preference, or the correctness of the group judgment – rather, aggregation functions implement a *compromise* between the different individual opinions. In social epistemology, however, a characteristic problem consists in finding a *consensual* group decision – a decision that any agent is convinced of, regardless of which opinion she initially held. The challenge is to model how agents rationally revise their beliefs in the light of deliberation in the group, good arguments for either case, and under which conditions they finally arrive at a joint opinion.

The Lehrer-Wagner model implements consensual decision-making. It was first developed as a descriptive mathematical model by Morris DeGroot (1974). Later, the philosopher Keith Lehrer and the mathematician Carl Wagner presented it as a normative model (Lehrer and Wagner 1981). It describes how agents reach consensus on a particular question from a set of diverging opinions. The problem consists in estimating a particular quantity  $x$ , such as the expected number of product sales, from the individual estimates  $v_i$  of every group member  $i$ . To be a model for rational group deliberation and decision-making, the Lehrer-Wagner model has to incorporate how competent the agents find each other as estimators of  $x$ . Other, more complex information seems to be relevant, too, e.g. an agent's competence in judging the expertise of other agents. For instance, while a group member may be perceived as an outstanding expert in estimating the expected number of sales, she might not be very self-confident and erroneously take the other agents to be more competent than herself. Typically, her deference to other, less competent agents will undermine her standing in the group. By contrast, some group members might be very good at evaluating the cogency of arguments, and recognizing experts in the group, but might not have a clue about what a reasonable estimate would be. Such group members apparently deserve to be taken seriously, despite their lack of expertise in the original subject matter.

Thus, we obtain a hierarchy of social information, namely perceived expertise on the issue under consideration, perceived expertise on other agents' expertise, and so on.<sup>1</sup> Lehrer and Wagner's crucial idea consists in codifying

mutual respect assignments in  $N \times N$  matrices  $W^{(m)}$  ( $N$  denoting the size of the group), where beliefs about expertise of the  $m$ -th order (e.g. who is a competent estimator of  $x$ , who is good at discerning experts, etc.) corresponds to the matrix  $W^{(m)}$ . Notably, the  $W^{(m)}$  need not be equal to each other.

$$W^{(m)} = \begin{pmatrix} w_{11}^{(m)} & w_{12}^{(m)} & \cdots & w_{1N}^{(m)} \\ w_{21}^{(m)} & w_{22}^{(m)} & \cdots & w_{2N}^{(m)} \\ \cdots & \cdots & \cdots & \cdots \\ w_{N1}^{(m)} & w_{N2}^{(m)} & \cdots & w_{NN}^{(m)} \end{pmatrix}$$

An important mathematical constraint is that the values in each row are nonnegative and normalized as to sum up to 1:  $\sum_{j=1}^n w_{ij}^{(m)} = 1$ . Thus, the  $w_{ij}^{(m)}$  represent relative weights which the agents ascribe to themselves and to others when it comes to estimating the unknown value  $x$ , or estimating competence in recognizing experts.<sup>2</sup> – It is then argued that the *rational* consensus of the agents consists in taking account of information of all levels of expertise, and the matrix  $W^\infty$  which aggregates all this information and gives the overall weight assignments will be a *consensus matrix*, i.e. all agents agree about each other's competence. Then,  $W^\infty$  is multiplied with a vector  $v$  that contains the agents' individual estimates of  $x$ , obtaining a *consensus value* for  $v$ .

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{pmatrix}$$

In general, taking into account only first-order respect assignments  $W^{(1)}$  will not directly lead to consensus ( $(W^{(1)}v)_i \neq (W^{(1)}v)_j$ ), nor does it incorporate more complex social information. Lehrer and Wagner (1981) and Wagner (1978) argue that the matrices  $W^{(m)}$  need to be multiplied with each other in order to feed beliefs of different orders into the belief revision process. For instance, the entry  $(W^{(2)}W^{(1)})_{ij}$  is equal to  $\sum_{k=1}^N w_{ik}^{(2)}w_{kj}^{(1)}$ , i.e. agent  $i$ 's overall judgment on the expertise of  $j$  is affected by all agents' judgments on  $j$ 's expertise in the subject matter ( $w_{kj}^{(1)}$ ), weighed by  $i$ 's own opinion on how good those agents are at discerning experts ( $w_{ik}^{(2)}$ ). Thus, by multiplying the  $W^{(m)}$  with each other, we finally take all relevant information into account, and thus arrive at a *rational consensus*. This convergence to consensus could also be understood as a process of group deliberation: after submitting their initial opinion, the agents subsequently take into account the various levels of

expertise, change their positions accordingly until the disagreement is finally resolved (cf. Wagner’s (1978) method of anonymous position papers).

But is a consensus guaranteed? Lehrer and Wagner (1981) show that under very weak constraints, the sequence  $(\prod_{m=1}^n W^{(m)})_{n \in \mathbb{N}}$  converges to a matrix  $W^\infty$  where all rows are identical, i.e. where all agents agree on their relative weights (see also Golub and Jackson 2007). In other words, the agents do not only achieve consensus on the issue under consideration (=all entries of  $W^\infty v$  are equal), but they also agree on the overall relative weight of each group member.<sup>3</sup>

Lehrer and Wagner (1981; see also Lehrer 1976) proposed their model as a *normative* model for consensual decision-making, based on a set of rationality requirements that deliberating agents should abide by, and which are satisfied by the model. Certainly, there is something to the procedure they advocate, in particular the aggregation of beliefs of different levels into an overall weight assignment.<sup>4</sup> For the present purpose, we suspend judgment on both normative and descriptive adequacy. Suffice to say that the model can be applied to a variety of expert panel decisions, such as environmental policy-making, where agents assign relative weights to the predictions of different ecological models, depending on how adequate they believe these models to be (see also Steele et al. 2007).

More precisely, we treat the Lehrer-Wagner model as a helpful exploratory tool to study group decision-making that captures important aspects of group reasoning and opinion change. It might be helpful to compare its value for social epistemology to the Lotka-Volterra model for predator-prey dynamics, or the billiard ball model for colliding molecules, which, though far from being fully adequate, have helped us to gain important general insights into the studied subject. In a similar way, we will now use the Lehrer-Wagner model to formalize, and to generalize, the notion of an epistemic peer, and to study the effects of certain social network structures on group decisions. Moreover, there is a lack of alternatives in the philosophical literature. The following section presents some results on the dynamics of disagreement derived in that framework.

### 3 Consensus among Epistemic Peers

A key notion in the current debate in social epistemology is the *epistemic peer* (Gutting 1982; Kelly 2005). Elga (2007, 484) describes epistemic peers as truth-oriented agents that are equally likely to be mistaken about a certain issue – for instance they might be equally good at evaluating (certain types of) evidence. Then, he proposes and defends the *equal weight view*, namely

that agents who consider themselves to be epistemic peers should assign equal weight to each other. Elga focuses on cases of binary judgment where one of the agents is right and the other is wrong, such as forecasting a winner of a horse race. Still, his considerations equally apply to the case of estimating an unknown quantity, such as our opening example of estimating the market success of a planned new product. If two agents submit estimates  $v_1$  and  $v_2$  and consider themselves to be epistemic peers, the equal weight view amounts to endorsing the linear average  $(v_1 + v_2)/2$  as the rational consensus.

We would like to explicate, and to generalize, this informal understanding of an epistemic peer. The Lehrer-Wagner model is suitable to make the notion precise. There is a natural, but restrictive definition:

**Definition 1:** The members of a group of agents  $G = \{1, \dots, N\}$  with weight matrices  $(W^{(m)})_{m \in \mathbb{N}}$  treat each other as *strict epistemic peers* if and only the first-order weight they ascribe to each other equals the weight they ascribe to themselves. In other words,  $w_{ij}^{(1)} = 1/N$ , for all  $i, j \leq N$ .

**Fact 1:** A group of strict epistemic peers settles on a equal weight consensus, regardless of the weight assignments for higher-order expertise (expertise on recognizing experts, etc.).

**Proof:** Let  $(W^{(m)})_{m \in \mathbb{N}}$  be the weight matrices of a group of strict epistemic peers. For any  $N \times N$  matrix  $M$  where the row entries sum up to 1,  $M \cdot W^{(1)} = W^{(1)}$ . Hence, the sequence  $(\prod_{m=1}^n W^{(m)})_{n \in \mathbb{N}}$  converges, and the consensus matrix  $W^\infty$  must be equal to  $W^{(1)}$ .

In other words, for agents who believe that everybody is equally competent in the subject matter, higher-order expertise does not matter: even if we perceived a peculiar group member as good at picking out experts, this would not bring any change since every agent believes that there are no particular experts in the group. Therefore the agents settle on the equal weight view without further ado.

However, such a perfect symmetry might be unrealistic, especially if a larger group of agents is involved. We feel that agents can treat each other as epistemic peers in a meaningful sense even if the definition is relaxed. Let us begin with the simple case of two agents. We propose the rationale that two agents are epistemic peers if and only if the sum of relative weights agent 1 is assigned equals the sum of relative weights agent 2 is assigned: each agent gets as much respect as she disseminates. More concisely:

$$w_{11}^{(m)} + w_{21}^{(m)} = w_{12}^{(m)} + w_{22}^{(m)} \quad \forall m \geq 1.$$

Due to the constraint  $w_{11}^{(m)} + w_{12}^{(m)} = 1$ , the weight matrices have the structure

$$W^{(m)} = \begin{pmatrix} w^{(m)} & 1 - w^{(m)} \\ 1 - w^{(m)} & w^{(m)} \end{pmatrix}.$$

Thus, for  $w^{(m)} \neq 1/2$ , the agents don't consider themselves to be on a par. However, they are *equally responsive to each other*, for each level of expertise as long as they are not dogmatic ( $w_{ij}^{(m)} > 0$  for all  $i, j \in \{1, 2\}$ ). The crucial idea is that if the symmetry is preserved at each level of expertise, the differences on the first-order level will finally wash out. You might think that you are more competent than your neighbor, but if you give some weight to your neighbor with respect to discerning experts, then this difference is smaller than thought if your neighbor believes himself to be more competent on the subject matter. Indeed, it is not difficult to show (proof omitted) that such agents will finally arrive at a consensus: through the symmetry on each level of expertise, the agents recognize each other as strict epistemic peers, leading them to the conclusion that each of them should carry equal weight:

$$\left( \prod_{m=1}^n W^{(m)} \right) \xrightarrow{n \rightarrow \infty} W^\infty = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

We believe that this rationale can be generalized to the case of  $N$  agents. Here is our definition of what it means that agents consider themselves to be epistemic peers:

**Definition 2:** A matrix  $(w_{ij})_{i,j \leq N}$  is *doubly stochastic* if and only if  $\forall i, j \leq N$ :

$$\sum_{k=1}^N w_{ik} = \sum_{l=1}^N w_{lj} = 1.$$

**Definition 3:** The members of a group of agents  $G = \{1, \dots, N\}$  with weight matrices  $(W^{(m)})_{m \in \mathbb{N}}$  treat each other as *epistemic peers* if and only if the following two conditions are satisfied:

**Overall Balance** For each  $m \geq 1$ , the matrix  $W^{(m)} = (w_{ij}^{(m)})_{i,j \leq N}$  is doubly stochastic, with  $w_{ij}^{(m)} < 1$  for all  $i, j \leq N$ .

**Connectivity** There is no subgroup  $S \subsetneq G$  that satisfies overall balance for any  $m \geq 1$ .

The requirement that  $W^{(m)}$  be doubly stochastic generalizes the intuition that the agents are eventually on a par with each other, but allows for a more liberal understanding: for instance, agent  $i$  might be considered as



incompetent by agent  $j$ , but as an outstanding expert by agent  $k$  who is, in turn, appreciated by  $j$ . When the agents are in overall respect balance, such differences should, or so we argue, finally cancel each other out. In particular, there are no dogmatic, stubborn agents that give all weight to themselves while neglecting the rest ( $w_{ii}^{(m)} < 1$  for all  $i \in G$ ). The net of respect assignments which connects the agents is woven such that on a whole, nobody has a worse standing than one of his peers: there are no privileged agents or opinion leaders. Thus, not only agents who have identical doxastic attitudes towards each other, but also agents with more complex attitudes may treat each other as epistemic peers as long as the structural requirement of overall balance is satisfied.

Finally, the condition of connectivity rules out the possibility that the group is divided into several subgroups that ignore each other. Such subgroups may be in internal respect balance, but certainly, we are then not dealing with a *single* group of epistemic peers.

It might be objected that agents who assign different weights to other group members do not treat each other as epistemic peers *by definition*, invalidating our explication. We believe this objection to be misguided: Firstly, such an understanding of epistemic peers is too narrow to apply to real cases. Secondly, it renders the link to the equal weight view trivial (see Fact 1) because there is only one weight assignment that allows the agents to be epistemic peers. We suggest that readers who are concerned that our definition is too liberal understand our project as an inquiry into which kinds of respect assignments lead to an equal weight outcome. That said, we maintain overall balance as the core idea of treating each others as epistemic peers, and we justify our explication by means of connecting it to the equal weight view. Indeed, Elga’s conclusion for the case of strict epistemic peers transfers to the more general case, too: each agent will finally ascribe equal weight to all her peers.

**Theorem 1:** Let  $G = \{1, \dots, N\}$  be a group of epistemic peers with weight matrices  $(W^{(m)})_{m \in \mathbb{N}}$  such that

$$\sum_{m=1}^{\infty} \min_{i,j \leq N} w_{ij}^{(m)} = \infty. \quad (\text{“minimal respect”}) \quad (1)$$

Then the agents finally settle on a consensus, and the consensus matrix  $W^\infty$  satisfies  $w_{ij}^\infty = 1/N$ ,  $\forall i, j \leq N$ .

**Proof:** See appendix.

Hence, theorem 1 vindicates arithmetic averaging as a rational procedure for a group of agents where the respect assignments balance each other.

Notably, the respect assignments can be quite different for each level of expertise, but still, the equal weight view results. Taking into account various levels of expertise, the agents move towards recognizing each other as *strict* epistemic peers. The only technical requirement that we had to make was (1) – it captures the intuition that each agent has some minimal respect for each member of the group, at least when higher-order expertise is asked. In particular, (1) is satisfied if there is a *subsequence* of weight matrices whose entries are bounded away from zero.

The significance of our result might be contested on the grounds that linear averaging does not correspond to the equal weight view, because this view only prescribes that all standpoints obtain equal weight and that there is no reason to “stick to one’s guns”, but not a particular (linear) method how to resolve disagreement. However, when we are estimating a numerical value, averaging various individually reasonable estimators is a good method to minimize the estimation variance, and to lower the chance of being strongly mistaken. This speaks for linear averaging as a default method to implement the equal weight view when the disagreement is about a real number.<sup>5</sup> Now, we apply our results to the study of subgroups of epistemic peers in social networks.

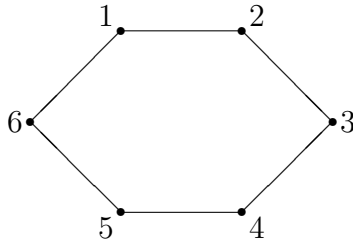
## 4 Social Networks

The social sciences study networks for describing information spreading and mutual social influence in groups. Usually, social networks are represented by means of graphs where agents are identified with vertices, and dependencies are modeled through edges connecting the vertices. Several formal approaches have been taken on the problem: DeMarzo et al. (2003) investigate the connection between persuasion capacities and connectedness in the group, i.e. how social influence is determined by the particular structure of the group network. Golub and Jackson (2007) propose a consensus model where information in the group is conveyed through the existent network (see also Zollman 2007). They give necessary and sufficient conditions for convergence of opinions in group deliberation. However, they do not allow agents to give different weights to one another if they are equidistant in the network.

To our mind, this condition is overly restrictive. Surely, if agents  $i$  and  $j$  are direct neighbors in a graph, the weight that they assign to each other will be higher than the weight they assign to agents that are further away. This seems to be part of the meaning of such a graphical representation. But why should this entail that the weights for all neighbors have to be equal?

Some links may be stronger, others weaker. As it turns out, relaxing that assumption does not preclude results of theoretical interest.

In the remainder of the section, we have a look at certain prominent network structures some of which are related to the results of the previous section.<sup>6</sup>



**Figure 1:** A ring-network of agents (dots), with edges representing connections.

Figure 1 represents a characteristic network structure, namely a *ring*. In a ring, agents are connected to themselves, and to two neighbors, but all connections to other agents are mediated via their neighbors. Since the ring structure seems to capture several aspects of information flow in reality, i.e. through geographic proximity or closeness of subject area, it is certainly worth to be studied in the Lehrer-Wagner model. Now, we give a definition of how the Lehrer-Wagner model captures the structure of social networks:

**Definition 4:** A group of agents  $G = \{1, \dots, N\}$ , together with weight matrices  $(W^{(m)})_{m \in \mathbb{N}}$ , is arranged according to a graph  $\langle V, E \rangle$  ( $V$  being a set of vertices,  $E \subset V^2$  a set of edges) if and only if for all  $m \geq 1$

- $w_{ij}^{(m)} \neq 0$  if  $(i, j) \in E$ ,
- $w_{ij}^{(m)} = 0$  in all other cases.

The definition captures the idea that the impact of a remote agent  $A$  on our judgment is only mediated via the direct/indirect neighbors, and their judgment on  $A$ 's competence. Thus, the weight matrices of a ring structure (which is understood as an undirected graph) have positive entries only on the main diagonal and the two neighboring diagonals. Theorem 1 ensures that if the agents treat each other as epistemic peers, a ring structure will converge to the arithmetic average of the individual opinions, as one might expect from the graphical representation. The notable thing about this application of theorem 1 is that weights between direct neighbors need not be equal to each other, as presupposed by Golub and Jackson (2007). (In such a case, the result would be trivial, due to symmetry considerations.)

Also, it is interesting to study the case of a *subgroup* of epistemic peers in a larger group:

**Definition 5:** In a group  $G = \{1, \dots, N\}$  of  $N$  agents with weight matrices  $(W^{(m)})_{m \in \mathbb{N}}$ , the agents  $S = \{i_1, i_2, \dots, i_L\} \subset G$  are a *subgroup of epistemic peers* if and only if:

- For each  $m \geq 1$ , the matrix  $W_S^{(m)} = (w_{i_j i_k}^{(m)})_{j, k \leq L}$  is such that

$$\left( \frac{1}{\sum_{j \leq L} w_{i_1 i_j}^{(m)}} \right) W_S$$

is a doubly stochastic matrix, and

- For each  $j \notin S$  and  $k, l \in S$ ,  $w_{jk}^{(m)} = w_{jl}^{(m)}$ .

While the first condition captures the overall respect balance inside the subgroup, the second condition ensures that this balance is not disturbed by diverging competence assignments that are made by outsiders. Indeed, such a subgroup of epistemic peers will end up assigning equal weights to each other:

**Theorem 2:** Let  $G = \{1, \dots, N\}$  be a group of  $N$  agents with weight matrices  $(W^{(m)})_{m \in \mathbb{N}}$  which contains a subgroup  $S = \{i_1, \dots, i_L\}$  of epistemic peers and which, as a whole, converges to consensus. Then, all agents in  $S$  obtain the same consensual weights.

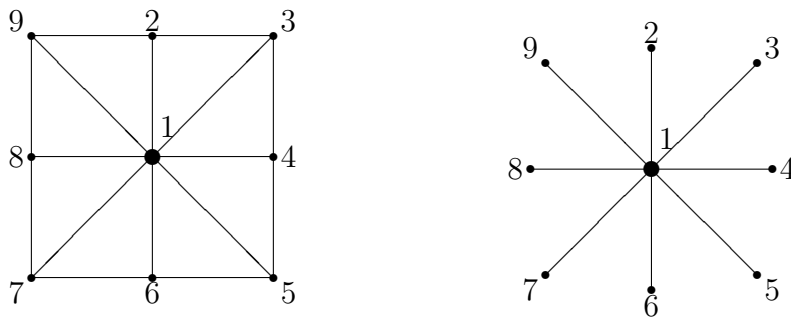
**Proof:** See appendix.

Now, we turn to the problem of modeling (lack of) expertise. We distinguish two cases: first, a group of epistemic peers recognizes an expert or opinion leader, and second, the group of agents is star-like connected. The latter case amounts to the presence of an opinion leader, too, but here, all respect for other group members is mediated by the leader. Interestingly, the results are the same for either case. First, the case of a subgroup of epistemic peers:

**Corollary 1:** Let  $G = \{1, \dots, N\}$  be a group of  $N$  agents with a subgroup  $S = \{2, \dots, N\}$  of epistemic peers, weight matrices  $(W^{(m)})_{m \in \mathbb{N}}$  and individual estimates  $v_i$ ,  $i \leq N$ . Let  $l := w_{j1}^{(m)} = w_{k1}^{(m)}$  for all  $j, k \leq N$  and  $m \geq 1$ . Then, the consensus value is equal to  $l v_1 + (1 - l) \frac{1}{N-1} \sum_{k=2}^N v_k$ .

**Proof:** It is easy to check that the entries in the first column remain unchanged by matrix multiplication. The rest follows from Theorem 2.

Thus, the consensus value is the weighted average of the consensus value in  $S$  and the leader's submission, as one might expect. An important application of this result consists in the “wheel” structure (see figure 2), which is basically a ring structure where an opinion leader sits in the middle of the network, being connected to any member of the group. Whenever the opinion leader does not distort the respect balance between the outer nodes, the other agents form a subgroup of epistemic peers. Hence, the consensus value will be the weighted average of the leader opinion and the arithmetic average of the other members' opinion.



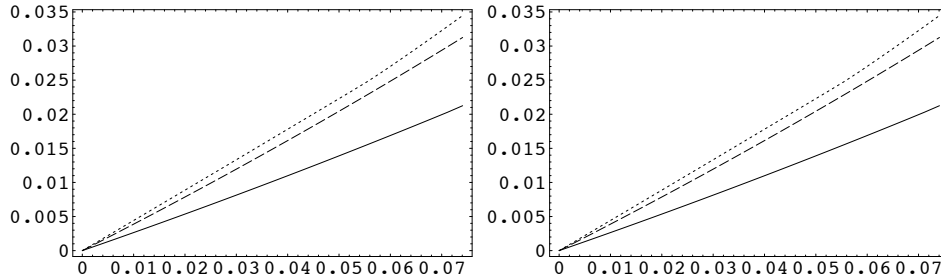
**Figure 2:** A wheel- and star-network of agents (dots), with edges representing connections.

Now, we turn to the other structure represented in figure 2, namely a star-like network where all connections between the agents go through the opinion leader.

**Theorem 3:** Let  $G = \{1, \dots, N\}$  be a group of  $N$  agents with weight matrix  $W := W^{(m)} = W^{(n)}$  for all  $m, n \geq 1$  that is arranged in a star-like network structure. Agent 1 denotes the central vertex, and moreover  $w_{1j} = w_{1k}$  and  $w_{j1} = l$  for all  $j, k \leq N$ . Then, a consensus is reached, and the consensus value is equal to  $lv_1 + (1 - l) \frac{1}{N-1} \sum_{k=2}^N v_k$ .

**Proof:** See appendix.

Thus, the results are independent of whether a subgroup of epistemic peers deals with an opinion leader (wheel network), or take each other only seriously to the extent that the opinion leader does (star network): in both cases, the final result exactly mirrors the leader's initial impact on the group whereas all other submissions are considered to be equally important. Notice that we did not address the question of whether the leader bias is beneficial or not – this depends on the leader's factual competence and honesty. The Lehrer-Wagner model is silent on these questions.



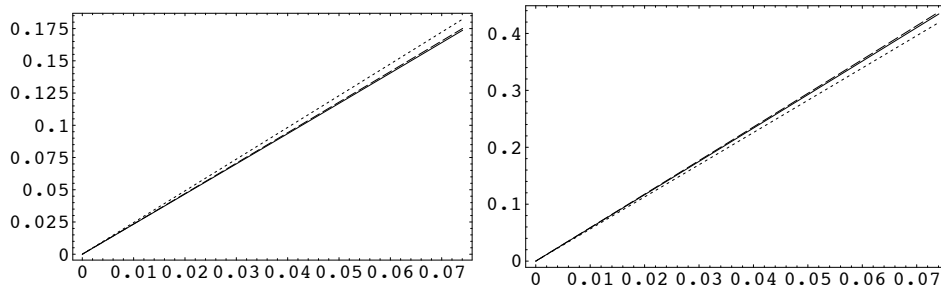
**Figure 3:** Relative mean difference between the consensus values of a group in overall balance and the consensus values of a slightly disturbed group (ego bias), as a function of the strength of disturbance  $\varepsilon$ , for several parameter values (details in the appendix). Left figure: relative mean difference of all weights. Right figure: relative mean difference of assignments of self-respect.

## 5 Simulations on the Stability of the Results

The results in the previous two sections provide valuable generalizations of the equal weight view for the case of agents that do not have perfectly symmetrical doxastic attitudes towards each other. In particular, we have pointed out that epistemic peers, understood as a set of agents in overall respect balance, end up assigning equal weights to each other. We believe that these results are of high theoretical interest, but the central condition of double stochasticity of the weight matrices is clearly idealized: in real situations, agents will only *roughly*, but probably not precisely, be in overall respect balance. Thus, we ask how stable our theorems are when this crucial premise is violated.

We model two particular deviations from double stochasticity – namely a situation where agents are slightly biased towards themselves (ego bias) and situations where agents are slightly biased towards a particular person in the group (leader bias). By means of numerical simulations, we investigate the effect of slight disturbances on the results. More precisely, we add a matrix of the order  $\varepsilon$ , with  $0 < \varepsilon \ll 1$ , where the row sums vanish.

Figure 3 plots the relative mean difference between the consensus values of a group of epistemic peers and a group whose members “approximately” treat each other as epistemic peers, except for a *ego bias*, implying that everybody reserves some extra weight for herself. Arguably, this is the most natural and most frequently found violation of the overall balance condition. We see that in both cases, the ego bias does not lead to a relative mean difference of more than 3.5%, even for the most unfavorable parameter values. Moreover, it does not make a difference whether the relative mean difference is taken



**Figure 4:** Relative mean difference between the consensus values of a group in overall balance and the consensus values of a slightly disturbed group (leader bias), as a function of the strength of disturbance  $\varepsilon$ , for several parameter values (details in the appendix). Left figure: relative mean difference of all weights. Right figure: relative mean differences of the relative weight of the opinion leader.

over assignments of self-respect only, or over all mutual respect assignments. Thus, our analytical results for groups of epistemic peers are highly stable if agents are sticking to their guns a little bit – they will nevertheless end up with an outcome close to equal weights. We suspect that the various ego biases tend to cancel each other out on a whole because in principle, each agent is ready to settle on a consensus.

To complement the picture, we also considered the case of a *leader bias* where all group members reserve some extra weight for a particular person  $l$  in the group (figure 4). Here, the disturbance effects in favor of  $l$  add each other up and cause a higher relative mean difference than for the case of the ego bias ( $\approx 17\%$  instead of 2-3.5%). This is even more pronounced if we calculate the relative mean difference merely over the weights  $l$  is assigned (right graph in figure 4): we obtain a relative mean difference of up to 40%, implying that the leader’s weight increases sharply as a function of the leader bias. This numerical result confirms a conjecture a special case of which was proven in Theorem 3: a leader bias is neither augmented nor decreased in the deliberation process. Indeed, the observed relative mean difference in the consensus matrix almost exactly corresponds to the relative mean difference which was initially imposed.

We can summarize the results as follows: For the important case of the ego bias, Theorem 1 remains valid because the relative mean differences are negligible. On the other hand, a leader bias has visible consequences. In other words, only if the bias is assigned in a systematic way, favoring a particular person, violation of the premises of Theorem 1 yields significant effects. Still, we would like to point out that the magnitude of the effect does not exceed the magnitude of the cause: Both figure 3 and figure 4

exhibit linear dependencies between the dimension of the disturbance and the dimension of the effect on the final estimate, vindicating the robustness of our results.

## 6 Conclusions

This paper has motivated and investigated the Lehrer-Wagner model as a tool for studying the dynamics of consensual decision-making, and applied it to an analysis of how to resolve disagreement under various circumstances. In particular, we have explicated the notion of an epistemic peer for a group of  $N$  agents and argued that overall respect balance captures the core intuition underlying the concept. Then, we have proved that as long as the overall balance of mutual respect is maintained in that group, agents end up ascribing equal weight to each other. This result transfers to subgroups of epistemic agents as long as other agents do not disturb the balance.

Moreover, we have applied these results to various social networks, such as a ring, a wheel and a star. The latter cases are tailor-made for studying the case of generally recognized expertise and opinion leadership. We observed that the initial relative weight of the opinion leader exactly corresponds to her consensual relative weight.

All our results are derived for special cases of the Lehrer-Wagner model. Due to the linearity of the model, we conjectured that slight violations of those assumptions, as they might occur in practice, will only lead to slight violations of the results. Numerical simulations have confirmed our conjecture: the analytical results are robust, and the observed deviations depend linearly on the scale of the disturbance. Moreover, a leader bias brings about greater changes than an ego bias.

We believe that our results facilitate the understanding of disagreement resolution among epistemic peers, and we plan to combine analytical and simulation-based results in further work in social epistemology.

## Acknowledgements

We would like to thank Jason McKenzie Alexander, Vincent Buskens, Jake Chandler, Igor Douven, Adam Elga, Joost Mulder, Pieter Ruys, Jonah Schupbach, and Carl Wagner for their encouraging and helpful feedback.



## A Proofs of the Theorems

**Proof of Theorem 1:** First, we argue for the convergence of  $(\prod_{m=1}^n W^{(m)})_{n \in \mathbb{N}}$ , second, we show that the consensus matrix  $W^\infty$  must be doubly stochastic, and third and last, we conclude that the consensual weights have to be equal to  $1/N$ .

The first claim follows from Lehrer and Wagner's (1981, 152) Theorem 8.3 which shows the convergence of  $(\prod_{m=1}^n W^{(m)})_{n \in \mathbb{N}}$  towards a consensus matrix  $W^\infty$  if the group satisfies equation (1) ("minimal respect", see page 8).

For the second claim, we observe that since  $G$  is a group of epistemic peers, any  $W^{(m)}$  is doubly stochastic. Next, we show that  $\prod_{m=1}^n W^{(m)}$  is also doubly stochastic, for any  $n \in \mathbb{N}$ . Then we are done, because the set of doubly stochastic matrices is a closed subset of  $\mathbb{R}^{N^2}$ , which entails that the consensus matrix  $W^\infty$  has to lie in that set as well.

Hence it suffices to show that the product  $C$  of two doubly stochastic matrices  $A$  and  $B$  is again doubly stochastic. For the columns, we calculate for each  $i \leq N$

$$\begin{aligned} \sum_{j \leq N} c_{ji} &= \sum_{j \leq N} \left( \sum_{k \leq N} a_{jk} b_{ki} \right) \\ &= \sum_{k \leq N} b_{ki} \left( \sum_{j \leq N} a_{jk} \right) = \sum_{k \leq N} b_{ki} \\ &= 1, \end{aligned}$$

and analogously for the rows.

Since  $W^\infty$  is a consensus matrix, every row has the same entries. Together with double stochasticity, this implies

$$1 = \sum_{j \leq N} w_{ji}^\infty = N w_{ki}^\infty$$

for any  $i, k \leq N$ . We conclude  $w_{ki}^\infty = 1/N$  for all  $i, k \leq N$ , completing the proof.  $\square$

**Proof of Theorem 2:** Without loss of generality, let  $S = \{1, 2, \dots, L\}$  denote the subgroup of epistemic peers in  $G = \{1, \dots, N\}$ .

By assumption,  $(\prod_{m=1}^n W^{(m)})_{n \in \mathbb{N}}$  converges to a consensus matrix  $W^\infty$ . It remains to be shown that the subgroup of epistemic peers arrives at the

equal weight outcome. First we show that for any  $C \in \cup_{n \in \mathbb{N}} \{\prod_{m=1}^n W^{(m)}\}$ , and any  $i \notin S$  and any  $j, k \in S$ ,

$$c_{ij} = c_{ik}. \quad (2)$$

Then it follows that  $w_{ij}^\infty = w_{ik}^\infty$ , because the relevant set of matrices is a closed subset of  $\mathbb{R}^{N^2}$  (cf. the argument in the proof of Theorem 1). Then, because all rows of  $W^\infty$  are identical, we obtain  $w_{ij} = w_{kl}$  for all  $i, j, k, l \in S$ , completing the proof.

To show (2), we consider two matrices  $A$  and  $B$  that contain a subgroup of epistemic peers  $S = \{1, \dots, L\}$ . We demonstrate that the column sums over the elements of  $S$  in  $C := A \cdot B$  are equal to each other, and that  $C$  satisfies (2). Then we infer that each product of the  $W^{(m)}$  satisfies these conditions, too, and we are done.

First, we show that the column sums in  $S$  are equal to each other. To that purpose, let  $i, j \leq L$ :

$$\begin{aligned} \sum_{l=1}^L c_{li} - c_{lj} &= \sum_{l=1}^L \sum_{k=1}^N a_{lk} b_{ki} - \sum_{l=1}^L \sum_{k=1}^N a_{lk} b_{kj} \\ &= \sum_{k=1}^L (b_{ki} - b_{kj}) \sum_{l=1}^L a_{lk} + \sum_{k=L+1}^N (b_{ki} - b_{kj}) \sum_{l=1}^L a_{lk} \\ &= \left( \sum_{l=1}^L a_{l1} \right) \left( \sum_{k=1}^L b_{ki} - \sum_{k=1}^L b_{kj} \right) + 0 \\ &= 0. \end{aligned}$$

(In the third line, we have used that  $\sum_{l \leq L} a_{lk}$  is constant in  $k$ .) Second, we show that  $c_{ij} = c_{ik}$  for all  $i \notin S$  and any  $j, k \in S$ :

$$\begin{aligned} c_{ij} - c_{ik} &= \sum_{l=1}^N a_{il} b_{lj} - \sum_{l=1}^N a_{il} b_{lk} \\ &= \sum_{l=1}^L a_{il} b_{lj} - \sum_{l=1}^L a_{il} b_{lk} + \sum_{l=L+1}^N a_{il} (b_{lj} - b_{lk}) \\ &= a_{i1} \left( \sum_{l=1}^L b_{lj} - \sum_{l=1}^L b_{lk} \right) + 0 \\ &= 0. \end{aligned}$$

In the last argument, we have only used the equality of column sums of  $B$  and the fact that  $A$  and  $B$  satisfy (2).  $\square$

**Proof of Theorem 3:** The weight matrix of the star network is

$$W = \begin{pmatrix} l & \frac{1-l}{N-1} & \frac{1-l}{N-1} & \cdots & \frac{1-l}{N-1} \\ l & 1-l & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ l & 0 & 0 & \cdots & 1-l \end{pmatrix}.$$

We cannot apply Theorem 2 because  $\{2, \dots, N\}$  is no subgroup of epistemic peers. But  $W$  satisfies the conditions for convergence of the series  $(W^n)_{n \in \mathbb{N}}$  (Theorem 7.2 in Lehrer and Wagner 1981, 131). It remains to be shown that the consensual weights are  $(l, (1-l)/(N-1), \dots, (1-l)/(N-1))$ . To this end, we first prove by means of induction that  $(W^n)_{11} = l$ . Obviously, the condition is satisfied for  $n = 1$ . Then, the inductive step

$$\begin{aligned} (W^n)_{11} &= \sum_{j=1}^N w_{1j}^{(n-1)} w_{j1} = l^2 + l \sum_{j=2}^N w_{1j}^{(n-1)} \\ &= l^2 + l(1 - w_{11}^{(n-1)}) = l^2 + l(1 - l) \\ &= l \end{aligned}$$

completes the proof of our claim. Now, the theorem follows easily: the symmetry of  $W$  in the second to  $N$ -th column yields  $(W^n)_{1i} = (W^n)_{1j}$  for all  $2 \leq i, j \leq N$ . Therefore,  $W^\infty$  will have  $(l, (1-l)/(N-1), \dots, (1-l)/(N-1))$  as the first row entry, and thus also as the consensual weight assignment.  $\square$

## B Simulation Details

For both simulations (ego bias and leader bias), we consider the case of five agents with one single weight matrix  $W = W^{(m)}$ , with

$$W = \frac{1}{15} \begin{pmatrix} 5 & 4 & 3 & 2 & 1 \\ 1 & 5 & 4 & 3 & 2 \\ 2 & 1 & 5 & 4 & 3 \\ 3 & 2 & 1 & 5 & 4 \\ 4 & 3 & 2 & 1 & 5 \end{pmatrix}.$$

Thus, the relative weights correspond to a standardized ordinal ranking of all group members. Obviously,  $W$  is doubly stochastic. We disturb  $W$  by

adding the matrix

$$D_E = \varepsilon \begin{pmatrix} 1 & a & b & 0 & 0 \\ a & 1 & 0 & 0 & b \\ a & 0 & 1 & b & 0 \\ b & a & 0 & 1 & 0 \\ 0 & 0 & b & a & 1 \end{pmatrix}$$

(ego bias) respectively

$$D_L = \varepsilon \begin{pmatrix} 1 & a & b & 0 & 0 \\ 1 & a & 0 & 0 & b \\ 1 & 0 & 0 & b & a \\ 1 & a & 0 & b & 0 \\ 1 & 0 & b & a & 0 \end{pmatrix}$$

(leader bias).  $a$  and  $b$  sum up to  $-1$ , thus the row sums vanish. We calculate the limiting matrix  $W^\infty$  of the sequence  $((W + D_{E/L})^n)_{n \in \mathbb{N}}$  and plot the overall relative mean differences to equal weights:

$$\text{RMD} = \frac{1}{1/N} \left( \frac{1}{N^2} \sum_{i,j=1}^N \left| w_{ij}^\infty - \frac{1}{N} \right| \right) = N \left( \frac{1}{N} \sum_{j=1}^N \left| w_{1j}^\infty - \frac{1}{N} \right| \right)$$

as a function of  $\varepsilon$ . For the right subgraphs of figures 3 and 4, we also plot the relative mean differences for self-respect, respectively leader-respect ( $l$  denoting the leader):

$$\text{RMD}(\text{SR}) = \frac{1}{1/N} \left( \frac{1}{N} \sum_{i=1}^N \left| w_{ii}^\infty - \frac{1}{N} \right| \right) = \text{RMD}$$

$$\text{RMD}(\text{LR}) = \frac{1}{1/N} \left( \frac{1}{N} \sum_{i=1}^N \left| w_{il}^\infty - \frac{1}{N} \right| \right) = N \left| w_{1l}^\infty - \frac{1}{N} \right|.$$

In the graphs of figure 3 and 4, we set  $a$  and  $b$  to  $-0.1/-0.9$  (dotted lines),  $-0.4/-0.6$  (dashed lines), and  $-0.5/-0.5$  (full lines).

## Notes

<sup>1</sup>This idea is similar to the common knowledge rationale used by Aumann (1976), Milgrom (1981) and others.

<sup>2</sup>Notably, these weights need *not* depend on how close the agents are to each other with respect to their estimates, contrary to Regan et al. 2006 or Hegselmann and Krause 2002 where the weight assignment also depend on the distance between the agents' opinions on

the subject under discussion. Still, the model permits us to assign relative weights as a function of an opinion difference.

<sup>3</sup>This feature of the Lehrer-Wagner model is the counterpart to Aumann's (1976) famous result that rational agents with common priors, commonly known posteriors, and commonly known expertise have to agree with each other.

<sup>4</sup>See Bradley 2006 for a recent critique.

<sup>5</sup>Moreover, it could be said that the truth orientation of epistemic agents is neglected in the Lehrer-Wagner model which is therefore an inadequate tool for modeling epistemic peers. This objection does not trouble us either. Nothing prevents us from stipulating that the agents are truth-directed while maintaining that their respect assignments can be captured through the Lehrer-Wagner model.

<sup>6</sup>See also Zollman's (2007) analysis of the epistemic benefits of certain kinds of network structures among scientists who are working on one and the same issue.

## References

- [1] Aumann, Robert. 1976. "Agreeing to Disagree". *Annals of Statistics* 4: 1236-1239.
- [2] Bradley, Richard. 2006. "Taking Advantage of Difference of Opinion". *Episteme: A Journal of Social Epistemology* 3: 141-155.
- [3] Christensen, David. 2007. "Epistemology of Disagreement: the Good News". *Philosophical Review* 116: 187-217.
- [4] DeGroot, Morris. 1974. "Reaching a Consensus". *Journal of the American Statistical Association* 69: 118-121.
- [5] DeMarzo, Peter, Dimistri Vayanos and Jeffrey Zwiebel. 2003. "Persuasion Bias, Social Influence and Unidimensional Opinions". *The Quarterly Journal of Economics* August 2003: 909-968.
- [6] Elga, Adam. 2007. "Reflection and Disagreement". *Noûs* 41: 478-502.
- [7] Elga, Adam. 2009. "How to Disagree About How to Disagree". In R. Feldman and T. Warfield (eds.), *Disagreement*. Oxford: Oxford University Press, in press.
- [8] Feldman, Richard. 2007. "Reasonable Religious Disagreements", in: L. Antony (eds.): *Philosophers Without Gods: Meditations on Atheism and the Secular Life*, 194-214. Oxford: Oxford University Press.
- [9] Golub, Benjamin and Matthew O. Jackson. 2007. "Naïve Learning in Social Networks: Convergence, Influence, and the Wisdom of Crowds". *Working Papers Series FEEM Working Paper No. 64.2007*.

- [10] Gutting, Gary, 1982. *Religious Belief and Religious Skepticism*. Notre Dame: University of Notre Dame Press.
- [11] Hegselmann, Rainer and Ulrich Krause. 2002. "Opinion Dynamics and Bounded Confidence: Models, Analysis and Simulation". *Journal of Artificial Societies and Social Simulation* 5(3).
- [12] Jackson, Matthew O. 2008. *Social and Economic Networks*. Princeton: Princeton University Press.
- [13] Kelly, Thomas. 2005. "The Epistemic Significance of Disagreement". In J. Hawthorne and T. Szabo (eds.), *Oxford Studies in Epistemology* 1, 167-196. Oxford: Oxford University Press.
- [14] Kelly, Thomas. 2009. "Peer Disagreement and Higher Order Evidence". In R. Feldman and T. Warfield (eds.) *Disagreement*. Oxford: Oxford University Press (in press).
- [15] Lehrer, Keith. 1976. "When Rational Disagreement is Impossible". *Noûs* 10: 327-332.
- [16] Lehrer, Keith and Carl Wagner. 1981. *Rational Consensus in Science and Society*. Dordrecht: Reidel.
- [17] Milgrom, Paul. 1981. "An Axiomatic Characterization of Common Knowledge". *Econometrica* 49: 219-222.
- [18] Regan, Helen, Mark Colyvan and Lisa Markovchick-Nicholls. 2006. "A Formal Model for Consensus and Negotiation in Environmental Management". *Journal of Environmental Management* 80: 167-176.
- [19] Steele, Katie, Helen Regan, Mark Colyvan and Mark A. Burgman. 2007. "Right Decisions or Happy Decision-Makers?". *Social Epistemology: A Journal of Knowledge, Culture and Policy* 21: 349-368.
- [20] Wagner, Carl. 1978. "Consensus through respect: a model of rational group decision-making". *Philosophical Studies* 34: 335-349.
- [21] Zollman, Kevin. 2007. "The Communication Structures of Epistemic Communities". *Philosophy of Science* 74: 574-587.