

# Challenging the majority rule in matters of truth

BERND LAHNO

*Frankfurt School of Finance & Management*

**Abstract:** The majority rule has caught much attention in recent debate about the aggregation of judgments. But its role in finding the truth is limited. A majority of expert judgments is not necessarily authoritative, even if all experts are equally competent, if they make their judgments independently of each other, and if all the judgments are based on the same source of (good) evidence. In this paper I demonstrate this limitation by presenting a simple counterexample and a related general result. I pave the way for this argument by introducing a Bayesian model of evidence and expert judgment in order to give a precise account of the basic problem.

**Keywords:** competence, evidence, social epistemology, testimony, trust in experts, two-expert problem

**JEL Classification:** C11, D71, D82, D83

Surprises can be useful in epistemology. Epistemology is most helpful when it leads to normative recommendations that are surprising in that they are counterintuitive or in contradiction with established practice (Miriam Solomon 2006, 30).

Seeking and utilising the advice of experts is a very common and useful practice in a complex world; this is especially so given the ever-increasing stream of information, which no single individual can comprehend and process entirely on her own. We regularly ask experts for advice. And in many cases we ask different experts for their independent advice on one and the same issue. If we, for instance, fear a serious disease we may well ask several medical specialists for their diagnoses. But what if the diagnoses given are inconsistent?

---

**AUTHOR'S NOTE:** The work on this project was supported by a fellowship at the Käte Hamburger Kolleg/Centre for Global Cooperation, University of Duisburg-Essen, Germany.

How should and how do we actually cope with disagreement among experts? In an explorative paper Alvin Goldman (2001) investigates what good reasons a novice might have for trusting one putative expert more than another. He first presents a (non-exhaustive) list of such reasons. According to the second entry in his list, Goldman's advice to the layman confronted with conflicting expert judgments is to check for "agreement from additional putative experts on one side of the subject in question" (Goldman 2001, 93).

As a matter of fact, this piece of advice is the one Goldman discusses most extensively in his paper. He gives the following formal argument:

**Goldman's proposition.** *If expert judgments are sufficiently reliable and independent of each other, then additional experts confirming some proposition  $\varphi$  add positive credibility to  $\varphi$  (Goldman 2001, 99-101).*

The reader may be somewhat disappointed about this result. It does not seem to suit the needs of a layperson confronted with conflicting expert judgments very well. What we would rather have is something like this:

**Proposition (\*).** *If expert judgments are sufficiently reliable and independent of each other, then  $\varphi$  is more probably true than not, if the number of experts confirming  $\varphi$  exceeds the number of experts confirming  $\neg\varphi$ .*

The idea that the majority of expert judgments may play a decisive role as a basis of informed decisions appears very natural to us. This is somewhat reflected in the vast literature on the aggregation of opinions in psychology and management science (see Budescu 2006; Yaniv 2004 for overviews). There is much evidence to suggest that people predominantly use simple averaging rules to aggregate information from multiple sources. Averaging has recently also become salient in the debate about swarm intelligence and the wisdom of crowds.<sup>1</sup> In the case

---

<sup>1</sup> See, most prominently, Surowiecki 2004. Exploiting the wisdom of crowds presupposes some sort of aggregation of the information dispersed across the crowd. Most of the cases Surowiecki discusses are—like the classic Galton example—cases of simple averaging. The other aggregation procedure prominently mentioned is the market price (e.g., on the stock market) as a representation of a collective assessment of expected value or success (chapter 11). On one occasion Surowiecki refers to Bayesian updating (pp. xx-xxi), but the case discussed is a Bayesian search based on an aggregate expert opinion. How this aggregate opinion is produced is not specified (but it does not seem to be itself based on some form of Bayesian updating).

of a categorical binary judgment (e.g., whether or not a surgery is needed to cure a disease), assigning equal weight to the independent judgments of various experts and then averaging them amounts to following the advice of the majority (Yaniv 2004, 76).

In social epistemology, the rediscovery of Condorcet's jury theorem<sup>2</sup> eventually made the majority rule a standard reference point in the debate about the demands of epistemic rationality when merging the opinions of independent judges, and in other places such as the recent debate about the so-called discursive dilemma (Kornhauser and Sager 1986, see discussion below).

Thus, there seem to be good reasons to consider the majority rule not only a widely used, but also a well-founded guiding principle in forming an opinion on the basis of diverging expert judgments. However, a claim like (\*) is not true in general and Goldman is well advised not to make it! There are important cases in which the majority just cannot decide on matters of truth, even though all judgments are made independently by equally competent jurors with exactly the same opportunity to obtain information on the issue. The object of this paper is to demonstrate the existence of these cases, mainly by presenting a simple counterexample.

Within the theory of judgment aggregation (see List 2012 for an overview) it has been acknowledged that simply following the majority of judgments is not necessarily optimal if truth is the object. The main reason for this is that a rational decision maker would consider that judgments are given by individuals with different degrees of competence. Consequently, the judgments of individuals should be differently weighted according to their competence and/or other qualities of the individual and her informational resources, which might influence the reliability of her judgment. In what follows, my argument presupposes that all individuals are equally competent and that all are in essentially the same position to obtain relevant information on the matter at hand; thus all judgments should be assigned the same weight. Nevertheless, following the majority of judgments will turn out to be suboptimal.

The argument in this paper is related to more general findings from the literature on 'opinion pooling' (Dietrich and List 2014). This literature is concerned with aggregating opinions, which are represented

---

<sup>2</sup> Condorcet 1785 did not explicitly formulate the theorem in his famous essay. The basic proposition was first explicitly stated by Duncan Black (1958).

by an assignment of probability to some proposition. ‘Linear pooling’ determines the average of the—possibly weighted—individual opinions as the correct collective opinion and corresponds to the majority rule in judgment aggregation. Franz Dietrich and Christian List (2014) argue that this procedure is suboptimal as a truth-tracking procedure. Their reasoning is that collective opinions generated by linear pooling may not adequately incorporate the whole body of information upon which individual opinions rest.

I will use Bayesian updating here as the benchmark of truth tracking. This corresponds to an opinion aggregation procedure known as ‘supra-Bayesian opinion pooling’ (Morris 1974), which Dietrich and List refer to at the end of their survey, but do not discuss in detail. Their reason for not engaging thoroughly with this procedure is that it rests on unrealistic assumptions that pertain to the information available to individuals. Rather than to formulate a consistent theory of judgment aggregation, my aim in this paper is to challenge the majority rule as it is used in epistemic discourse on judgment aggregation. My argument proceeds as follows:

I will begin with a short presentation of the discursive dilemma as presently discussed. The goal is to identify hidden assumptions and unquestioned premises that underwrite the majority rule as a guiding principle in opinion aggregation. In subsequent sections, I present a counterexample and the formal results—these serve to show that the assumptions of majority rule are by no means indisputable. However, before presenting a counterexample the underlying problem must be articulated. I will, therefore, motivate the primary argument by introducing a simple Bayesian model of evidence and expert judgment.<sup>3</sup> The paper concludes with some general remarks.

## THE DISCURSIVE DILEMMA

In philosophical debates about epistemic rationality, the majority rule, as an aggregation procedure for judgment, has been a predominant topic in the context of the so-called discursive dilemma. Here is a quote and an example from a recent paper to illustrate the focus of this debate (Pigozzi 2006, 285fn.):<sup>4</sup>

---

<sup>3</sup> Goldman (1999, 103-109) introduces a similar model for the special case of testimonial evidence.

<sup>4</sup> The problem first became known as “the doctrinal paradox” in the discipline of jurisprudence. The example given here has the same logical structure as the original example by Lewis A. Kornhauser and Lawrence G. Sager (1986).

A department of a prestigious university offers one career development fellowship to a candidate (proposition R) if and only if the candidate proposed a good research project (P) and if she has an excellent track record of publication (Q) [...] Suppose that there are three members in the departmental committee. Each of them consistently casts her vote on R (the conclusion) depending on her judgments on P and Q (the premises). The three members vote as shown in the table below.

	<i>P: Good project?</i>	<i>Q: Excellent publication?</i>	<i>R: Fellowship?</i>
Member 1	Yes	Yes	Yes
Member 2	Yes	No	No
Member 3	No	Yes	No
Majority	Yes	Yes	No

The problem is that the majority rule produces inconsistent results. While a majority certifies that the candidate has a good project and that publications of the candidate are excellent, there is no majority to accept the candidate as a fellow. Based on the above propositions, majority voting implies that premises P and Q are accepted whereas the conclusion R is rejected.

Two ways to escape the paradox suggest themselves and have come to dominate the debate. One can either confine aggregation to premises and derive the conclusion deductively according to these premises (the *premise-based procedure*); or an individual can start by drawing the conclusion according to her individual assessment of the premises and the votes on the conclusion are then aggregated (the *conclusion-based procedure*). Note that both procedures adhere to the majority rule as an essential standard of judgment aggregation; albeit, they restrict its use to a proper subset of the judgments to prevent inconsistencies.

List and Pettit (2002) introduce a general result that extends the discursive dilemma to a wider class of aggregation procedures. The defining property of this class is “systematicity” (List and Pettit 2002, 99), which requires that the collective acceptance of a claim depends solely on the pattern of individual acceptance and not, for example, on the content of the judgment or the evidence it is based on. To determine the collective judgment of some claim it suffices to consider individual votes (whatever the claim is).<sup>5</sup> This is a characteristic of all voting

<sup>5</sup> See List and Pettit 2003, 99; compare also the related but probably more transparent characterization “systematic responsiveness” in List 2006, 376.

procedures that are based solely on counting votes. Not surprisingly, the debate that follows from this—and other attempts to generalise the observation of inconsistency in the discursive dilemma—focuses on both the variations and strengths of the majority rule as a truth-tracking aggregation procedure.<sup>6</sup>

The root of the whole debate is the fundamental finding that using the majority rule in aggregating individual judgment may result in inconsistencies if several, logically connected judgments are at issue. Yet, those participants leading the debate never seem to seriously consider that something might be fundamentally wrong with applying the majority rule to matters of truth. Occasionally scholars recognise tension between democratic procedures and the fundamental quest for truth.<sup>7</sup> But the general reaction to the discursive dilemma is not a closer inspection of the relation between truth and majority voting in principle. Instead, alternative procedures are considered that constrain majority voting to suitable selections of judgments or by quota rules (Grossi and Pigozzi 2012). It is not the general truth-tracking capabilities of the majority rule that are called into question, but rather the proper pattern of its application.

To the outside observer this may appear somewhat strange. Two related observations from the history of this body of thought might help to better understand the approach taken in this debate.

First, one aspect of the debate (and a familiar domain for many of its participants) stems from social choice theory. From the very beginning the standard view of the problem has included elements of social choice theory—such as Arrow's impossibility theorem or the Condorcet paradox—which are formally similar and were perceived as substantially related. So, as an intuitive point of departure, the focus on voting with the majority rule was transferred from social choice theory to social epistemology.

Second, Condorcet's jury theorem has always been understood as a guiding background insight. Roughly, the theorem says that the majority judgment on some yes/no issue by a group of  $N$  sufficiently competent individuals is more probably true than a single judgment by one of its members, and the probability of its truth approaches 1 as  $N$  increases.<sup>8</sup>

---

<sup>6</sup> See, e.g., List 2005; Bovens and Rabinowicz 2006; Hartmann and Sprenger 2012.

<sup>7</sup> See the distinction between procedural and epistemic democracy in List and Goodin 2001.

<sup>8</sup> See, e.g., Bovens and Rabinowicz 2006 for an overview and generalizations of the theorem.

Looking at the theorem in this rough formulation, it may well seem that the supremacy of the majority rule as a truth-tracking procedure is not a matter of debate but a formally proven and indisputable result.

Notice, however, that Condorcet's jury theorem states that the majority rule outperforms individual judgment under certain conditions, it does not say anything about how it relates to other aggregation rules. Moreover, it is worthwhile to have a closer look at the conditions, particularly at the assumption of sufficient competence. This assumption could be specified as follows: There is a  $p > 0.5$ , such that the probability that an individual judgment is true is at least  $p$  for every individual in the group. The hidden assumption behind this condition is that for every individual  $i$  there is a constant probability  $p_i$  that one of his judgments is correct, independently of the specific content of the judgment. If his judgment is that there will be a thunderstorm tomorrow, this will come true with probability  $p_i$ ; but if his judgment would have been 'no thunderstorm tomorrow', this would have indicated 'no thunderstorm' with probability  $p_i$  as well. Whatever an expert says about whatever issue has the same authority in terms of truth. As we will see, this is a very strong condition, which cannot be assumed to be satisfied in general—not even among individuals whom we unhesitatingly and rightly classify as competent experts.

One aim of this paper is to challenge the social choice perspective and to sow some uneasiness regarding the majority rule when truth is at stake. This will be done in a much simpler context than that of the discursive dilemma by way of illustrating a counterexample that focuses on a single statement. But before this subversive task is undertaken, I will introduce some formal specifications of evidence and expert knowledge to get a firm grip on the problem.

## EVIDENCE

As is common in the debate about the discursive dilemma I will confine my analysis to simple yes/no problems. An individual has evidence for some proposition  $\varphi$  being true if he has information on some matter of fact that is suitably related to the truth of  $\varphi$ . A convenient example is evidence for prostate cancer. The PSA (Prostate-Specific Antigen) blood level is the most common and—although contested to some extent—the best prostate cancer marker available. Prostate cancer causes, as a rule, a high PSA blood level, which can easily be detected. An increased PSA blood level, however, can have other causes as well, e.g., prostatitis

(prostate inflammation) or benign prostatic hyperplasia (a swelling of the prostate). A high PSA blood level is evidence for prostate cancer because the probability of a high PSA blood level is significantly greater in the event of prostate cancer than not.

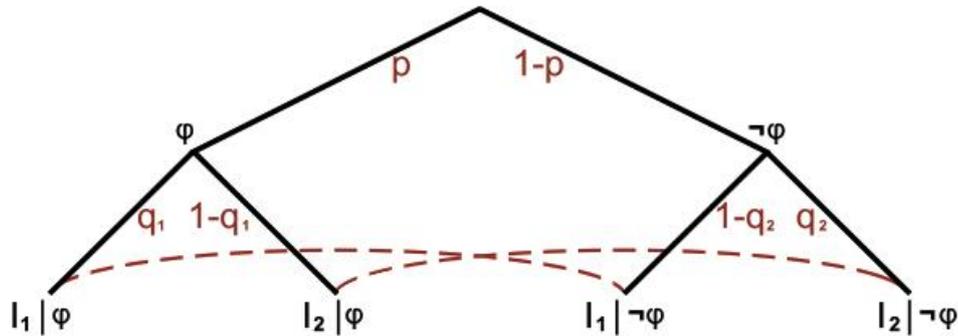


Figure 1: An indicator model of evidence with two indicator states

Figure 1 shows an abstract model of the basic interrelations. Call this an indicator model of evidence in a yes/no problem with two indicator states. An individual  $A$  is interested in knowing whether  $\varphi$  ( $A$  has prostate cancer) or  $\neg\varphi$  ( $A$  does not have prostate cancer) is true. His prior estimation of  $\varphi$  is  $p > 0$ .  $A$  is observing an indicator  $I$  (the PSA blood level) with (as is assumed to simplify the problem) two indicator states  $I_1$  (high PSA blood level) and  $I_2$  (low PSA blood level).  $A$  can correctly judge the state of the indicator, but he does not know whether  $\varphi$  is the case, i.e., he cannot discriminate between  $I_1$  on condition that  $\varphi$  and  $I_1$  on condition that  $\neg\varphi$  (indicated by the curved line connecting the nodes  $I_1|\varphi$  and  $I_1|\neg\varphi$ ). However, he does know the conditional probabilities  $q_1 = \text{prob}(I_1|\varphi)$  and  $q_2 = \text{prob}(I_2|\neg\varphi)$ . It should be clear now how the graph in Figure 1 represents the situation.<sup>9</sup>

Observing the indicator may provide information about  $\varphi$  to  $A$ . Bayes' rule describes how the information about  $\varphi$  is conveyed by  $A$ 's observations. Given his basic information about the probabilistic interrelations,  $A$  can update his initial probability estimate for  $\varphi$  being true ( $\text{prob}(\varphi) = p$ ) after observing  $I_1$  or  $I_2$ :

<sup>9</sup> The representation follows the conventions of the theory of dynamic games. In fact, Figure 1 displays the game form of a very simple interaction in which nature is the only player choosing over alternatives.

$$\begin{aligned} \text{prob}(\varphi|I_1) &= \frac{\text{prob}(I_1|\varphi)\text{prob}(\varphi)}{\text{prob}(I_1|\varphi)\text{prob}(\varphi) + \text{prob}(I_1|\neg\varphi)\text{prob}(\neg\varphi)} \\ &= \frac{q_1p}{q_1p + (1-p)(1-q_2)} \\ \text{prob}(\varphi|I_2) &= \frac{\text{prob}(I_2|\varphi)\text{prob}(\varphi)}{\text{prob}(I_2|\varphi)\text{prob}(\varphi) + \text{prob}(I_2|\neg\varphi)\text{prob}(\neg\varphi)} \\ &= \frac{(1-q_1)p}{(1-q_1)p + (1-p)q_2} \end{aligned}$$

If the probability that  $\varphi$  is true rises after observing indicator state  $I_i$  one may say that  $I_i$  *indicates*  $\varphi$ . Or, to give a more formal, general definition:

**Definition 1.** *An event  $\omega$  is said to indicate that  $\psi$  or to be evidence for  $\psi$  iff  $\text{prob}(\psi|\omega) > \text{prob}(\psi)$ .*

There is a convenient characterisation of  $I_1$  (i.e., the event of observing  $I_1$ ) being evidence for  $\varphi$  in our indicator model of evidence:

**Remark 1.** *Consider an indicator model of evidence in a yes/no problem as represented in Figure 1.*

$$I_1 \text{ is evidence for } \varphi \Leftrightarrow I_2 \text{ is evidence for } \neg\varphi \Leftrightarrow q_1 + q_2 > 1.$$

The proof is by simple algebra (see Appendix).

We are especially interested in those indicators whose indicator states signal that the state of affairs they indicate is the most probable state:

**Definition 2.** *An event  $\omega$  is said to be **decisive** for  $\psi$  iff  $\text{prob}(\psi|\omega) > 0.5$ .*

**Definition 3.** *In an indicator model of evidence in a yes/no problem as represented in Figure 1, the indicator  $I$  is said to be **decisive** iff  $I_1$  is decisive for  $\varphi$  and  $I_2$  is decisive for  $\neg\varphi$ .*

Here is a convenient characterisation of an indicator being decisive:

**Remark 2.** *Consider an indicator model of evidence as represented in Figure 1. With  $q^* := q_1p + q_2(1-p)$ , the following equivalences hold:*

$$I_1 \text{ is decisive for } \varphi \Leftrightarrow \frac{1-q_2}{q_1} < \frac{p}{1-p} \Leftrightarrow q^* > 1-p$$

$$I_2 \text{ is decisive for } \neg\varphi \Leftrightarrow \frac{p}{1-p} < \frac{q_2}{1-q_1} \Leftrightarrow q^* > p.$$

$$I \text{ is decisive} \Leftrightarrow \frac{1-q_2}{q_1} < \frac{p}{1-p} < \frac{q_2}{1-q_1} \Leftrightarrow q^* > \max\{p; 1-p\}.$$

Again, the proof is by simple algebra (see Appendix).

Notice that  $q^*$  represents the overall probability that the state of the indicator  $I$  indicates the true state of affairs.

## EXPERT JUDGMENTS

Expert judgments are fundamentally related to evidence in at least two ways.

First, an expert has privileged access to evidence. She is a person who is in a particularly favourable situation to obtain information about some subject and/or who is particularly competent to process such information. Call the circumstances that favour or disfavour the access of some person to information on some subject her *opportunity* and her capacity to process this information correctly her *competence* (see Goldman 1999, 109). The evidence available to an expert is a function of her opportunity and her competence. Whatever the particular opportunity and competence of an expert may be, the evidence available to her will ultimately be of the same basic form as any other evidence. Thus, if an expert has evidence pertinent a yes/no problem it may be provided by an indicator as shown in the model above (e.g., given her opportunity and competence, the expert may be in a position to observe the PSA blood-level).

Second, an expert judgment about some subject is itself evidence of or for the subject. If an expert is sufficiently competent and has the opportunity to obtain relevant information about some subject (i.e., if she has sufficient evidence), then her judgment may be understood as an indicator on this subject.

Consider an expert judgment on a yes/no problem with the expert choosing between two options only:  $I_1 = \text{'yes, } \varphi \text{ is true'}$  or  $I_2 = \text{'no, } \varphi \text{ is not true'}$ . In presuming that competence and opportunity are sufficient, this judgment is evidence for  $\varphi$  in exactly the same way as described in our model in Figure 1, with  $q_1$  and  $q_2$  jointly representing the competence and opportunity of the expert. Moreover, if the evidence for  $\varphi$  used by the expert is based on a *decisive* indicator  $I^*$  with indicator states  $I_1^*$  and  $I_2^*$ , then the evidence given by the expert judgment is formally identical and equivalent to the evidence the judgment is based

on, i.e., it is described by a formally identical indicator with formally identical indicator states and the same conditional probabilities.

Notice that the judgment of an expert may differ from his testimony. An expert may judge that such and such is the case, but testify something else for various reasons. Testimony may, of course, still be evidence of or for some subject. To give an account of such evidence, not only are competence and opportunity to be included in the analysis but also the properties of 'honesty' and 'integrity'.<sup>10</sup> I abstract away the related strategic problems here by assuming that the judgment of the expert is correctly communicated and can be correctly assessed. Thus, my analysis concentrates on the evidential character of pure judgment only.

If two or more experts have the same competence (with respect to some issue  $\varphi$ ) and are in the same general position to obtain information on the issue (i.e., have the same 'opportunity'), then the evidence provided by a judgment of one expert will be characterized by the same conditional probabilities  $q_1$  and  $q_2$  as the evidence provided by a judgment of another expert. The experts *judge under equivalent epistemic conditions* relative to  $\varphi$  iff their judgment on  $\varphi$  can be described by the same indicator model with identical parameters  $q_1$ ,  $q_2$ , and  $p$ .

If experts judge under equivalent epistemic conditions, their judgments that  $\varphi$  should be assigned the same weight. A judgment ' $\varphi$ ' of one expert represents exactly the same evidence for  $\varphi$  as a judgment ' $\varphi$ ' of another expert under equivalent epistemic conditions. But this does not mean that a judgment in favour of  $\varphi$  should also have the same weight as a judgment for  $\neg\varphi$ . In fact, as we will see in the next section, an expert judgment ' $\varphi$ ' may produce stronger support for  $\varphi$  than the opposite judgment for  $\neg\varphi$  under equivalent epistemic conditions. This suggests that, in such cases, positive judgments should have more weight than negative ones. But this is exactly what is implicitly neglected by the assumptions behind Condorcet's theorem and explicitly ruled out by the systematicity condition.

---

<sup>10</sup> See Lahno 2012. Goldman discriminates 'opportunity', 'competence', and 'sincerity' (or 'honesty') as basic determinants of testimonial information. He does not make a distinction between integrity and honesty. See Goldman 1999, 109.

## ARE THERE SHARKS IN THE ISLAND WATERS?

We are now in a position to give a precise formulation of the majority rule problem of expert judgments in simple yes/no problems:

Suppose an uneven number of  $m$  experts<sup>11</sup> judge (independently) under equivalent epistemic conditions on a decisive indicator as given in Figure 1, with  $l$  experts judging ' $\varphi$  is true' and  $m-l$  experts judging ' $\neg\varphi$  is true': *Is the indicator  $I^*$  with indicator states  $I_1^* = 'l > m-l'$  and  $I_2^* = 'm-l > l'$  decisive?* The answer, as indicated above, is 'not necessarily!' A simple counterexample demonstrates this.

Imagine an island about which fishermen are interested to know whether there are sharks in the waters ( $\varphi$ ) or not ( $\neg\varphi$ ) before they go out to work. Assume that the probability of sharks in the water is 10% (this knowledge is shared by all fisherman):

$$\text{prob}(\varphi) = p = 0.1.$$

Also assume that there are three individuals, each of which are located on top of one of the three hills on the island with excellent panoptic visibility. If sharks are in the waters, sometimes a shark fin cuts across the surface of the sea ( $I_1$ ) and, therefore, may become visible to an observer on one of the hills. Assume that the three 'experts' on the hills share the same basic competence and opportunity in watching for shark fins. If there are sharks in the waters ( $\varphi$ ), then for each of them, independently of the other two, there is a 5% chance to spot a shark fin:

$$\text{prob}(I_1|\varphi) = q_1 = 0.05.$$

If there are no sharks ( $\neg\varphi$ ), no fin will be observed ( $I_2$ ):

$$\text{prob}(I_2|\neg\varphi) = q_2 = 1.$$

For each observer spotting a fin is evidence for sharks being in the water and spotting no fin is evidence for 'no sharks' ( $q_1+q_2 = 1.05 > 1$ ). Moreover, the evidence given to each observer is decisive ( $q^* > 0.9$ ).

So we have three experts independently (in the relevant sense) judging under equivalent epistemic conditions on whether  $\varphi$  or  $\neg\varphi$  is true. Each of the three judgments is evidence for  $\varphi$  described by an indicator model as in Figure 1 with identical parameters for all experts.

Suppose expert 1 spots a fin and the other two do not. Then the evidence suggests the presence of sharks:

---

<sup>11</sup> We assume an uneven number of experts to avoid problems with ties.

$\text{prob}(\varphi | 1 \text{ observes } I_1 \text{ AND } 2 \text{ observes } I_2 \text{ AND } 3 \text{ observes } I_2) = 1.$

But the majority of the experts judge  $\neg\varphi$ : No sharks!

## A GENERAL RESULT

The example illustrates that the aggregated judgment of a minority of experts may have more weight than the majority judgment even if all experts judge independently and under equivalent epistemic conditions. One may suspect that the result is an exception, the consequence of an extreme and exceptional parameter arrangement. A particularly striking feature of our example is the assumption that  $q_2 = 1$ : if there is no shark, no fins will be seen. This assumption, in fact, simplifies the problem dramatically. However, the force of the counterexample is not necessitated by this peculiar assumption. The example maintains its rebutting force if  $q_2$  is changed to a value below but sufficiently near to 1.

There is a general mechanism behind the example: one indicator state may have more evidential force than another; if this is the case,  $n$  observations of the first indicator may outweigh  $n+1$  observations of the second (for some integer  $n$ ).

To be more precise, let us first specify what can plausibly be meant by saying that one indicator state has 'more weight' than another.

**Definition 4.** Consider an indicator model of evidence in a yes/no problem as represented in Figure 1. Let  $I_1 I_2$  denote the event that two observations are independently made:  $I_1$  is observed first and  $I_2$  is observed second. The indicator state  $I_1$  is said to **have more weight than** the indicator state  $I_2$  iff the event  $I_1 I_2$  indicates  $\varphi$ , i.e., iff  $\text{prob}(\varphi | I_1 I_2) > p$ .

Here is a convenient characterisation:

**Remark 3.** Consider an indicator model of evidence as given in Figure 1 with  $I_1$  being evidence for  $\varphi$ . Then:

$I_1$  has more weight than  $I_2 \Leftrightarrow q_1(1-q_1) > q_2(1-q_2) \Leftrightarrow q_1 < q_2.$

We can now state our general result (see the Appendix for proofs):

**Main proposition.** Consider an indicator model of evidence in a yes/no problem as represented in Figure 1. Let  $I$  be decisive, let  $I_1$  be evidence for  $\varphi$  and  $I_2$  evidence for  $\neg\varphi$ . For any integer  $n$ , let  $I_1^n I_2^{n+1}$  denote the event that  $2n+1$  observations are independently made,

$I_1$  is observed  $n$  times and  $I_2$  is observed  $n+1$  times. Then the following holds:

If  $I_1$  has more weight than  $I_2$  (i.e.,  $q_2 > q_1$ ) and  $q_1 \neq 0$ , there is an integer  $n_0$  such that for all  $n \geq n_0$ ,  $I_1^n I_2^{n+1}$  is decisive for  $\varphi$ .

If the indicator states have different weights, then there is always a number  $n$  such that  $n$  expert judgments made with regard to independent observations of the weightier indicator state will outweigh a majority of  $n+1$  judgments to the contrary by other experts.

## CONCLUSION

The general question behind the problem discussed here is: How should we aggregate or accumulate the evidence obtained by different expert judgments? A seemingly natural suggestion seems to be: weigh the evidence of any single expert judgment in some suitable form representing her competence and her access to the relevant information; the total evidence then is determined by the weighted sum of the evidence given by the total multitude of individual judgments. If expert judgments come in the form of yes-or-no statements, and if all experts possess the same competence and have equivalent access to relevant information, then this amounts to accepting that the aggregated evidence favours the statement that is endorsed by a majority. The simple counterexample and the main proposition in the last section show that this is not a good way to aggregate the expert evidence in general.

The indicator model of evidence suggests a different way to aggregate expert evidence, namely by updating probability judgments according to Bayes' rule. In simple situations, similar to those analysed in this paper, this aggregation method will also amount to counting votes and forming judgment based on a simple number rule:

*Let all experts independently judge under equivalent epistemic conditions and let expert judgments be aggregated according to Bayes' rule. If  $I$  denotes the number of experts that testify  $\varphi$ , then there is a number  $k$  such that ' $I > k$ ' is decisive for  $\varphi$ .*

In the example above  $k = 0$ . Notice that  $k$  not only depends on  $q_1$  and  $q_2$ ; it is also dependent on the prior probability  $p = \text{prob}(\varphi)$ . So it will change in the process of learning about  $\varphi$  as more and more evidence is accumulated. The amount of confirmative expertise needed to believe a certain proposition  $\varphi$  is dependent upon the sort of claim that  $\varphi$  states

and on its prior credibility. There is no general ratio of 'yes' to 'no' testimonies that can be used as a reliable decision rule when it comes to believing or disbelieving a statement.

How does all this relate to the fundamental role of the majority rule in theories of judgment aggregation? I will conclude with two short remarks to this question.

First, the alleged foundation rests on the assumption that *the amount of affirmation that is needed to justify a belief in a certain claim should not depend on the nature of the claim at stake*. This is the essence of the condition of systematicity in Pettit and List's (2002) impossibility result concerning the discursive dilemma. When spelled out and made explicit in this way it loses much of its initial plausibility. If I see a shark this is better evidence for the proposition 'there are sharks' than not seeing sharks is for 'no sharks'. And this is not a peculiarity of the example given here. The idea that the weight given to an affirmation should be the same for all claims independently of their specific content is also embodied by the common assumption (of most versions) of Condorcet's jury theorem that the competence of a juror is defined by one single and constant probability  $p$  of judging correctly.<sup>12</sup> Again, the example shows that this may not be appropriate. The probability that somebody who judges that there are no sharks on the basis that he did not observe one is wrong may well be larger than the probability of being wrong when judging the presence of sharks on the basis of observation. As statisticians would claim: the probability of an error of type 1 may not be identical to the probability of a type 2 error. And we know from many contexts such as medical diagnostics that type 1 and type 2 errors are, in fact, not equally probable. In our model the probability of type 1 and type 2 errors are represented by  $q_1$  and  $q_2$ . The general result, thus, shows that whenever the type 1 and type 2 errors are not equally likely, we have a good reason to mistrust the majority rule in aggregating expert judgments.

Second, democracy is a normative standard in defining collective measures on the basis of individual interests. Individual opinions cannot have the same fundamental role in epistemology. Whereas interests are the ultimate, independent reference points in politics, opinions and judgments ought to be assessed in terms of a single more fundamental criterion: truth. Democracy (or the majority rule) and truth are not necessarily related. To be sure: the majority may (and often does)

---

<sup>12</sup> Kirchstein and Wangenheim 2010 is an exception to this rule.

indicate truth. But this is a contingent relationship that calls for critical investigation and cautious assessment on the basis of all relevant circumstances.

## APPENDIX

*Proof of Remark 1.*

$$\begin{aligned} I_1 \text{ is evidence for } \varphi &\Leftrightarrow \frac{q_1 p}{q_1 p + (1-p)(1-q_2)} > p \\ &\Leftrightarrow q_1 > q_1 p + (1-p)(1-q_2) \\ &\Leftrightarrow q_1 > 1 - q_2 \Leftrightarrow q_1 + q_2 > 1. \end{aligned}$$

$$\begin{aligned} I_2 \text{ is evidence for } \neg\varphi &\Leftrightarrow \frac{(1-p)q_2}{(1-q_1)p + (1-p)q_2} > 1 - p \\ &\Leftrightarrow q_2 > (1-q_1)p + (1-p)q_2 \\ &\Leftrightarrow q_2 > 1 - q_1 \Leftrightarrow q_1 + q_2 > 1. \end{aligned}$$

*Proof of Remark 2.*

$$\begin{aligned} \text{prob}(\varphi|I_1) > 0.5 &\Leftrightarrow \frac{q_1 p}{q_1 p + (1-p)(1-q_2)} > \frac{1}{2} \\ &\Leftrightarrow 2q_1 p > q_1 p + (1-p)(1-q_2) \\ &\Leftrightarrow q_1 p + q_2(1-p) > 1-p \Leftrightarrow \frac{1-q_2}{q_1} < \frac{p}{1-p}. \end{aligned}$$

$$\begin{aligned} \text{prob}(\neg\varphi|I_2) > 0.5 &\Leftrightarrow \frac{(1-p)q_2}{(1-q_1)p + (1-p)q_2} > \frac{1}{2} \\ &\Leftrightarrow 2(1-p)q_2 > (1-q_1)p + (1-p)q_2 \\ &\Leftrightarrow q_1 p + q_2(1-p) > p \Leftrightarrow \frac{p}{1-p} < \frac{q_2}{1-q_1}. \end{aligned}$$

The third equivalence is a combination of the first two equivalences.

*Proof of Remark 3.*

From Bayes' rule we get:

$$\text{prob}(\varphi|I_1 I_2) = \frac{q_1(1-q_1)p}{q_1(1-q_1)p + q_2(1-q_2)(1-p)}.$$

Then:

$$\begin{aligned} \text{prob}(\varphi|I_1I_2) > p &\Leftrightarrow \frac{q_1(1-q_1)p}{q_1(1-q_1)p + q_2(1-q_2)(1-p)} > p \\ &\Leftrightarrow q_1(1-q_1) > q_1(1-q_1)p + q_2(1-q_2)(1-p) \\ &\Leftrightarrow q_1(1-q_1)(1-p) > q_2(1-q_2)(1-p) \\ &\Leftrightarrow q_1(1-q_1) > q_2(1-q_2). \end{aligned}$$

Remember that  $I_1$  is evidence for  $\varphi$ , and thus  $q_1+q_2 > 1$ . Therefore:

$$\begin{aligned} \text{prob}(\varphi|I_1I_2) > p &\Leftrightarrow q_1(1-q_1) > q_2(1-q_2) \\ &\Leftrightarrow q_1 - q_2 > q_1^2 - q_2^2 \\ &\Leftrightarrow q_1 - q_2 > (q_1 - q_2)(q_1 + q_2) \\ &\Leftrightarrow (q_1 - q_2 > 0 \wedge 1 > q_1 + q_2) \\ &\quad \vee (q_1 - q_2 < 0 \wedge 1 < q_1 + q_2) \\ &\Leftrightarrow q_1 < q_2. \end{aligned}$$

*Proof of Main proposition.*

From Bayes' rule we get:

$$\begin{aligned} \text{prob}(\varphi|I_1^n I_2^{n+1}) > 0.5 &\Leftrightarrow \frac{q_1^n(1-q_1)^{n+1}p}{q_1^n(1-q_1)^{n+1}p + q_2^{n+1}(1-q_2)^n(1-p)} > .5 \\ &\Leftrightarrow q_1^n(1-q_1)^{n+1}p > q_2^{n+1}(1-q_2)^n(1-p) \\ &\Leftrightarrow \frac{p}{1-p} > \left(\frac{q_2(1-q_2)}{q_1(1-q_1)}\right)^n \cdot \frac{q_2}{1-q_1} \end{aligned}$$

From Remark 3, we know that  $\frac{q_2(1-q_2)}{q_1(1-q_1)} < 1$ , so there is an  $n_0$  such that for all  $n > n_0$  the inequality holds.

## REFERENCES

- Berend, Daniel, and Jacob Proush. 1998. When is Condorcet's jury theorem valid? *Social Choice and Welfare*, 15 (4): 481-488.
- Black, Duncan. 1958. *The theory of committees and elections*. Cambridge: Cambridge University Press.
- Bovens, Luc, and Wlodek Rabinowicz. 2006. Democratic answers to complex questions: an epistemic perspective. *Synthese*, 150 (1): 131-153.

- Budescu, David V. 2006. Confidence in aggregation of opinions from multiple sources. In *Information sampling and adaptive cognition*, eds. Klaus Fiedler, and Perter Juslin. Cambridge: Cambridge University Press, 327-352.
- Condorcet, Jean-Antoine-Nicolas de Caritat Marquis de. 1785. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris: Imprimerie royale.
- Dietrich, Franz, and Christian List. 2014. Probabilistic opinion pooling. Forthcoming in *Oxford handbook in probability and philosophy*, eds. Christopher Hitchcock, and Alan Hajek. Oxford: Oxford University Press.
- Goldman, Alvin. 1999. *Knowledge in a social world*. Oxford: Clarendon Press.
- Goldman, Alvin. 2001. Experts: which ones should you trust? *Philosophy and Phenomenological Research*, 63 (1): 85-110.
- Grossi, Davide, and Gabriella Pigozzi. 2012. Introduction to judgment aggregation. In *ESSLLI 2010/2011 Lectures*, eds. V. Goranko, and N. Bezhanishvili. Springer-FoLLI Lecture Notes in Computer Science (LNCS 7388), 160-209.
- Hartmann, Stephan, and Jan Sprenger. 2012. Judgment aggregation and the problem of tracking the truth. *Synthese*, 187 (1): 209-221.
- Kornhauser, Lewis A., and Lawrence G. Sager. 1986. Unpacking the court. *Yale Law Journal*, 96 (1): 82-117.
- Kirchstein, Roland, and Georg von Wangenheim. 2010. A generalized Condorcet jury theorem with two independent probabilities of error. *Discussion paper, MAGKS no. 11*.
- Lahno, Bernd. 2012. Simple games of information transmission. *Analyse & Kritik*, 34 (2): 315-338.
- List, Christian 2012. The theory of judgment aggregation: an introductory review. *Synthese*, 187 (1): 179-207.
- List, Christian. 2006. The discursive dilemma and public reason. *Ethics*, 116 (2): 362-402.
- List, Christian. 2005. Group knowledge and group rationality: a judgment aggregation perspective. *Episteme*, 2 (1): 22-38.
- List, Christian, and Robert Goodin. 2001. Epistemic democracy: generalizing the Condorcet jury theorem. *The Journal of Political Philosophy*, 9 (3): 277-306.
- List, Christian, and Philip Pettit. 2002. Aggregating sets of judgments: an impossibility result. *Economics and Philosophy*, 18 (1): 89-110.
- Morris, Peter. A. 1974. Decision analysis expert use. *Management Science*, 20 (9): 1233-1241.
- Pigozzi, Gabriella 2006. Belief merging and the discursive dilemma: an argument-based account of paradoxes of judgment aggregation. *Synthese*, 152 (2): 285-298.
- Solomon, Miriam. 2006. Groupthink versus the wisdom of crowds: the social epistemology of deliberation and dissent. *The Southern Journal of Philosophy*, 44 (S1): 28-42.
- Surowiecki, James. 2004. *The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. New York: Doubleday.
- Yaniv, Ilan. 2004. The benefit of additional opinions. *Current Directions in Psychological Science*, 13 (2): 75-78.

**Bernd Lahno** is professor of philosophy at Frankfurt School of Finance & Management and the academic director of Frankfurt School's bachelor program 'Management, Philosophy & Economics'. His research interests include theories of trust and cooperation, foundational issues in decision theory, and the philosophy of economics.

Contact e-mails: <b.lahno@fs.de> <b.lahno@me.com>