Deliberation and the Wisdom of Crowds

Franz Dietrich
Paris School of Economics & CNRS

Kai Spiekermann
London School of Economics

Revised and substantively extended version of December 2023[1]

Abstract

Does pre-voting group deliberation increase majority competence? To address this question, we develop a probabilistic model of opinion formation and deliberation. Two new jury theorems, one pre-deliberation and one post-deliberation, suggest that deliberation is beneficial. Successful deliberation mitigates three voting failures: (1) overcounting widespread evidence, (2) neglecting evidential inequality, and (3) neglecting evidential complementarity. Formal results and simulations confirm this. But we identify four systematic exceptions where deliberation reduces majority competence, always by increasing Failure 1. Our analysis recommends deliberation that is 'participatory', 'neutral', but not necessarily 'equal', i.e., that involves substantive sharing, privileges no evidences, but might privilege some persons.

1 Introduction: Deliberation and Voting

Does group deliberation improve group decisions? Many scholars of deliberation assert that it does, though others have warned that deliberation can fall into epistemic traps. Since the formal understanding of the epistemic merits of deliberation is at an early and disjointed stage, it is hard to assess who is right.

We present a formal analysis of *deliberation as sharing and absorbing*. The closest precursor is Ding and Pivato's (2021) model of deliberation as information disclosure, but our approach is social-choice-theoretic rather than game-theoretic. We construe deliberation as an *attempt* to exchange possibly complex 'evidences', such as arguments, intuitions, empirical facts or personal perspectives. Realistically, this exchange succeeds only partly, because many evidences are hard for someone to share (express, describe) and for others to absorb (understand, incorporate), given limitations in language, concepts, and awareness.

Our analysis provides a clearer understanding of when and how pre-voting deliberation benefits the voting outcome. Following the epistemic paradigm, we take everyone to vote for what they individually believe to be socially correct. Ideally, the voting outcome is informationally efficient, i.e., responds optimally to the total evidence dispersed across voters. Such efficiency can fail for at least three reasons:

[1] We thank anonymous referees for excellent suggestions. We are also grateful for helpful feedback received when presenting earlier versions of this work in various economic, political or philosophical seminars, and at the conferences Freedom and Reason (LSE, 2021), Aggregation across Disciplines (University Cergy-Pontoise, 2021), Social Choice under Risk and Uncertainty (Warwick Business School, 2022), Rationality and Democracy (Mainz, 2022), and Advances in Economic Design (Conservatoire National des Arts et Métiers, Paris, 2023). Franz Dietrich acknowledges support by the French National Research Agency through three grants (ANR-17-CE26-0003, ANR-16-FRAL-0010 and ANR-17-EURE-0001).

- *Failure 1: Overcounting widespread evidence.* Evidence held by more voters has exaggerated influence, by affecting more votes.
- *Failure 2: Neglecting evidential inequality.* Voters have the same weight, despite their unequally strong total evidence.
- *Failure 3: Neglecting evidential complementarity.* Information obtainable after combining different evidences dispersed across voters is undercounted, because few or no voters access all these evidences simultaneously.

All three failures stem from bad management of available but dispersed evidences. Figure 1 gives a stylised example with three voters and three evidences. Failure 1 arises

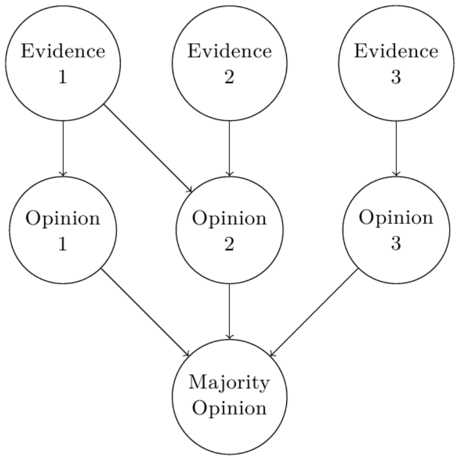


Figure 1: Example of growing access to evidence by deliberation

because evidence 1 is overcounted; it affects two votes while evidences 2 and 3 each affect only one vote. Failure 2 arises because voter 2 has stronger total evidence.[2] Failure 3 arises, for instance, if evidences 1 and 3 are complementary, because no voter has them both. For example, evidences 1 and 3 could be arguments that are uninformative in isolation but highly informative in combination.

The hope is that deliberation can improve the use of evidence. Specifically:

- Deliberation could reduce Failure 1 by increasing the spread of previously private or almost private evidences.
- Deliberation could reduce Failure 2 by letting voters with initially weak total evidence accumulate evidence.
- Deliberation could reduce Failure 3 by letting voters collect evidences from others and then recognize and use evidential complementarities.

But are these conjectures correct? We introduce a formal model of deliberation, and, for the first time, prove jury theorems that address the effect of deliberation on voting outcomes. We then analyse each failure, formally and using simulations. Our jury theorems and failure analysis partly confirm the optimistic take on deliberation, but surprisingly also identify some systematically harmful forms of deliberation. We present a typology of beneficial and harmful deliberation, allowing us to state more precisely which caveats apply to the thesis of the 'wisdom of deliberating crowds'. Deliberation can be harmful for four reasons: (i) persons have non-overlapping ('private') evidence; (ii)

[2] Assuming her two evidences are stronger in total than evidence 1 and than evidence 3.

persons share very badly but absorb very well; (iii) some evidences are far easier to share than others; (iv) some evidences are far easier to absorb than others. In all four cases, deliberation is harmful because it increases Failure 1. By contrast, deliberation robustly reduces Failures 2 or 3.

This paper is in 9 sections. After presenting existing approaches towards pre-voting deliberation in Section 2, Section 3 develops our new formal model of opinion formation and deliberation. Section 4 then presents two jury theorems that set an upper bound to collective competence, while suggesting that this upper bound is easier to reach post-deliberation. We also decompose the group's 'competence gap' into two gaps: the 'efficiency gap', which deliberation can potentially close, and the 'availability gap', which an increase in group size can potentially close. We then analyse Failures 1 and 2 formally in Section 5), before presenting exploratory simulations in Section 6, which suggest that deliberation is much better at reducing Failure 2 than Failure 1, and support a recommendation for 'participatory' and 'neutral' deliberation. After generalising the framework in Section 7, we finally address Failure 3 in Section 8, where we show that deliberation robustly reduces Failure 3. Section 9 offers concluding considerations.

2 Deliberation and Voting in Context

A large interdisciplinary literature addresses the interaction of deliberation and voting. Three perspectives dominate: the social-choice-theoretic, game-theoretic, and normative-democratic perspective. We now describe the first two perspectives and explain how our own approach relates to them. We set aside the extensive normative-democratic literature (see Min and Wong 2018 and Estlund and Landemore 2018 for reviews).

A first approach to deliberation comes from social choice theory, especially its epistemic branch. There is a long tradition of thinking about the epistemic effects of deliberation (Aristotle 1988 [350BC]; Condorcet 1785; Rousseau 1972 [1762]). Condorcet, a founder of social choice theory, was very much engaged with its epistemic aspect; not only did he interpret preferences epistemically as judgments of social betterness (McLean and Hewitt 1994, p. 38), he also proved the first of many jury theorems (e.g., Grofman et al. 1983, Ladha 1992, List and Goodin 2001, Dietrich and Spiekermann 2013; Pivato 2017; for a review see Dietrich and Spiekermann 2021). Unfortunately, deliberation has ambiguous effects on the assumptions of traditional jury theorems, potentially promoting voter competence while undermining voter independence. Although more recent jury theorems escape the concern that deliberation might undermine voter independence (e.g., Dietrich and Spiekermann 2013), no jury theorem addresses deliberation effects on (majority) outcomes. Our deliberation-specific jury theorems will aim to fill this gap. For other social-choice-theoretic takes on the epistemic virtues of deliberation, see Betz (2013), Perote-Pena and Piggins (2015), Goodin and Spiekermann (2018), Hartmann and Rafiee Rad (2018, 2020), and Hoek and Bradley (2022). Of course, pre-voting deliberation also serves non-epistemic purposes, such as enabling stable collective preferences (Dryzek and List 2003; Rafiee Rad and Roy 2021).

The game-theoretic literature interprets deliberation and voting as strategic interactions. Voters choose strategically what and when to communicate, and then how to vote. One insight of this literature is that, even if all share the same 'epistemic' preference for correct outcomes, incentives for strategic manipulation can arise in deliberation and voting, depending on the environment (e.g., Coughlan 2000; Austen-Smith and Feddersen 2006; Gerardi and Yariv 2007). We emphasise two semi-game-theoretic models of deliberation: Chung and Duggan's (2020) model of myopic discussion, constructive discussion and debate, and Ding and Pivato's (2021) model of deliberation as a process of information disclosure. These analyses share with us the focus on the dynamics of information exchange. In general, while a typical game-theoretic approach starts from assumptions about individual motivations and predicts how individuals deliberate, we will start from how they deliberate and predict how deliberation affects the information distribution and the correctness of individual and collective judgments. Our analysis therefore starts where a typical game-theoretic analysis ends. These two approaches are complementary. By setting aside the micro-foundations of behaviour, we sacrifice some explanatory power to gain in parsimony, focus, and generality, as explained in Section 3.5.

3 A Model of Opinions and Deliberation

This section presents our formal model, in a simple version later generalised in Section 7.

3.1 Opinions and their sources

A group of persons, labelled $1, ..., n$, faces two options, labelled 1 and -1. The group is denoted $N = \{1, ..., n\}$ and has any finite size $n \geq 1$. Following the epistemic 'Condorcetian' paradigm, exactly one option is objectively or intersubjectively *correct*; it is called the *state of the world*, for short the *state*. We represent it by a random variable $\mathbf{x}$ taking the value 1 or -1. In general, we denote random variables in bold letters, their particular values in non-bold letters, and the probability function by 'Pr', all of which refer to an underlying probability space (left implicit).

Each person forms an opinion about which opinion is correct. There are three possible opinions: the opinion that option 1 is correct (labelled 1), the opinion that option -1 is correct (labelled -1), and a neutral or undecided opinion (labelled 0). Opinions are based on 'evidences', in the broadest sense that includes empirical facts, arguments, normative aspects, and other inputs into opinion formation (but we set aside non-evidential 'noise' inputs, captured later in our generalised model). Formally, let S be a finite non-empty set of *sources*, and for each source $s \in S$ let $\mathbf{e}_s$ be a real-valued random variables, the *evidence* from source s. A positive, negative, or zero value of an evidence represents support for option 1, support for option -1, or evidential neutrality, respectively. The strength of this support is represented by the absolute value of the evidence. For instance, if the source s is an argument, then the evidence $\mathbf{e}_s$ measures which opinion it supports, and how strongly.

Each person i accesses some set of sources, her *source set*, represented by a random variable $\mathbf{S}_i$ whose values are subsets of S. In a court jury, a juror's source set might contain a witness report, a legal argument, and a legal text interpreting the law, while another juror's source set might contain the defendant's facial expression when interrogated, and other sources. In the introductory example of Figure 1, the source sets of persons 1, 2, and 3 contain one, two, and one source, respectively.

We can now define several derivative concepts. The *opinion of a person* i is the option supported by i's total evidence:

$$\mathbf{o}_i = \begin{cases} 1 & \text{if } \sum_{s\in\mathbf{S}_i} \mathbf{e}_s > 0 \\ -1 & \text{if } \sum_{s\in\mathbf{S}_i} \mathbf{e}_s < 0 \\ 0 & \text{if } \sum_{s\in\mathbf{S}_i} \mathbf{e}_s = 0. \end{cases}$$

The *majority opinion* is:

$$\mathbf{o}_{maj} = \begin{cases} 1 & \text{if } |\{i : \mathbf{o}_i = 1\}| > |\{i : \mathbf{o}_i = -1\}|, \text{ equivalently } \sum_i \mathbf{o}_i > 0 \\ -1 & \text{if } |\{i : \mathbf{o}_i = 1\}| < |\{i : \mathbf{o}_i = -1\}|, \text{ equivalently } \sum_i \mathbf{o}_i < 0 \\ 0 & \text{if } |\{i : \mathbf{o}_i = 1\}| = |\{i : \mathbf{o}_i = -1\}|, \text{ equivalently } \sum_i \mathbf{o}_i = 0. \end{cases}$$

The *competence* of a person i is the probability of a correct opinion $p_i = Pr(\mathbf{o}_i = \mathbf{x})$. The *majority competence* is the probability of a correct majority opinion $p_{maj} = P(\mathbf{o}_{maj} = \mathbf{x})$.

Diversity is key to successful deliberation.[3] It can be construed as heterogeneity in sources, i.e., dissimilarity between the source sets $\mathbf{S}_i$ (the arguments, the empirical knowledge, etc.) of different persons i. Under minimal diversity, people have identical source sets, so identical opinions. Under maximal diversity, they have pairwise disjoint source sets. A different concept is that of *intrapersonal diversity*. Someone has high intrapersonal diversity if they have a large source set, hence an opinion with a broad basis. As will emerge, deliberation tends to 'internalise' diversity: it lets sources be more widely held, which transforms interpersonal into intrapersonal diversity.

We make three simplifying assumptions (lifted later):

Equiprobable States: the state $\mathbf{x}$ takes both values 1 and -1 with probability $\frac{1}{2}$.

Simple Gaussian Evidence: Given any state $x \in \{\pm 1\}$, the evidences $\mathbf{e}_s$ ($s \in S$) have independent Gaussian distributions with mean x and some variance σ^2 that is the same across states x and sources s. So, each evidence correlates positively with the state: positive evidence objectively supports state 1, negative evidence objectively supports state -1. This positive correlation is the rationale behind our definition of opinions $\mathbf{o}_i$, according to which each evidence indeed pulls the opinion towards the state of same sign.

Independent Sources: The source-accessing events are independent across people and sources, and jointly independent of the state and the evidences. Formally, for each person $i \in N$ and source $s \in S$, we consider the event that person i accesses source s, '$s \in \mathbf{S}_i$', and we require these source-accessing events to be mutually independent, and jointly independent of the state-evidence combination $(\mathbf{x}, (\mathbf{e}_s)_{s\in S})$.[4]

The probability that a person i accesses a source s will be denoted $p_{s\to i} = Pr(s \in \mathbf{S}_i)$ and called an *access probability*. The access probabilities $(p_{s\to i})_{s\in S, i\in N}$ fully determine the distribution of the source profile $(\mathbf{S}_i)$. How? We use Independent Sources twice. First, the probability that a person i has a source set S_i is the product of the probabilities of

[3] For an influential approach to diversity see Hong and Page (2004, 2012) and Page (2007).

[4] For instance, when deputies of Congress form opinions about the effectiveness of a law (the state), whether deputy 1 has listened to (accesses) the verdict of some expert (a source) is independent of which other sources she and other deputies access, and also independent of the law's effectiveness (the state) and the evidence from each source (e.g., the verdict of the expert).

accessing any source in S_i and *not* accessing any other source:

$$Pr(S_i) = \left(\prod_{s\in S_i} p_{s\to i}\right)\left(\prod_{s\in S\setminus S_i} \overline{p_{s\to i}}\right) \tag{1}$$

where $\overline{p}$ stands for $1-p$. Second, the probability of an entire source profile (S_i) is the product $\prod_i Pr(S_i)$, with $Pr(S_i)$ given by (1).

To summarise, our formal primitive is a *simple opinion structure*, by which we mean a triple $(\mathbf{x}, (\mathbf{e}_s)_{s\in S}, (\mathbf{S}_i)_{i\in N})$, in short $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$, that contains:

(1) a random variable $\mathbf{x}$, the *state* or *correct option*, taking the value 1 or -1 with equal probability;

(2) a family $(\mathbf{e}_s)$, indexed by some set S of *sources* (non-empty and finite), consisting of real-valued random variables, the *evidences* from these sources, which have state-conditionally independent Gaussian distributions with mean $\mathbf{x}$ and with some fixed variance $\sigma^2 > 0$;

(3) a family $(\mathbf{S}_i)$, indexed by some set $N = \{1, ..., n\}$ of *persons* $(1 \leq n < \infty)$, consisting of random subsets of S, the *source sets* of these persons, such that the accessing events '$s \in \mathbf{S}_i$' are independent across sources s and persons i and of the state and the evidences.

Terminology: The *source profile* is the combination of source sets across persons $(\mathbf{S}_i)_{i\in N}$, in short $(\mathbf{S}_i)$. Person i's *evidence bundle* is the family of her evidences $(\mathbf{e}_s)_{s\in\mathbf{S}_i}$; it is doubly random, through her source set $\mathbf{S}_i$ and the evidences $\mathbf{e}_s$ from her sources s. The *evidence profile* is the combination of evidence bundles across people $((\mathbf{e}_s)_{s\in\mathbf{S}_i})_{i\in N}$, in short $((\mathbf{e}_s)_{s\in\mathbf{S}_i})$.

3.2 The rationality of opinions

Is this opinion model arbitrary from a rationality perspective? The worry is natural, as we presuppose a seemingly naive heuristic for forming opinions: adding up one's evidences and comparing the sum with zero. In fact, such opinion formation *is* rational in a perfectly classical sense. Why?

Classic rationality requires evaluating opinions (decisions) by expected utility. Given our epistemic setting, let us identify 'utility' with 'correctness level', defined as 1 if the opinion is correct, 0 if it is incorrect, and $\frac{1}{2}$ if it is neutral, i.e., zero. Technically, a person i or her opinion $\mathbf{o}_i$ is *classically rational* if the expected correctness level of $\mathbf{o}_i$ weakly exceeds that of all her other possible opinions $\mathbf{o}$. Here, a *possible opinion* of person i is any random variable $\mathbf{o}$ that generates 1, -1 or 0 as a function of i's information $(\mathbf{e}_i)_{i\in\mathbf{S}_i}$; its *correctness level* is 1 if $\mathbf{o} = \mathbf{x}$ (correct opinion), 0 if $\mathbf{o} = -\mathbf{x}$ (false opinion), and $\frac{1}{2}$ if $\mathbf{o} = 0$ (neutral opinion).

Theorem 1 *Given any simple opinion structure* $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$, *the opinion* $\mathbf{o}_i$ *of any person* i *is classically rational.*

Later, non-simple opinion structures can also model non-rational opinions.

3.3 Ideal and efficient opinion

The *ideal opinion* is the hypothetical opinion based on all sources:

$$\mathbf{o}_{ideal} = \begin{cases} 1 & \text{if } \sum_{s\in S} \mathbf{e}_s > 0 \\ -1 & \text{if } \sum_{s\in S} \mathbf{e}_s < 0 \\ 0 & \text{if } \sum_{s\in S} \mathbf{e}_s = 0. \end{cases}$$

Its correctness probability $p_{ideal} = Pr(\mathbf{o}_{ideal} = \mathbf{x})$ is the *ideal competence*. The ideal opinion is a little too 'ideal', as it neglects that some sources are not available to anyone. The 'efficient' opinion builds only on *available* evidence. Formally, the *available source set* is the union of the personal source sets $\cup_i \mathbf{S}_i$. The *efficient opinion* is the hypothetical opinion based on the available sources:

$$\mathbf{o}_{eff} = \begin{cases} 1 & \text{if } \sum_{s\in \cup_i \mathbf{S}_i} \mathbf{e}_s > 0 \\ -1 & \text{if } \sum_{s\in \cup_i \mathbf{S}_i} \mathbf{e}_s < 0 \\ 0 & \text{if } \sum_{s\in \cup_i \mathbf{S}_i} \mathbf{e}_s = 0. \end{cases}$$

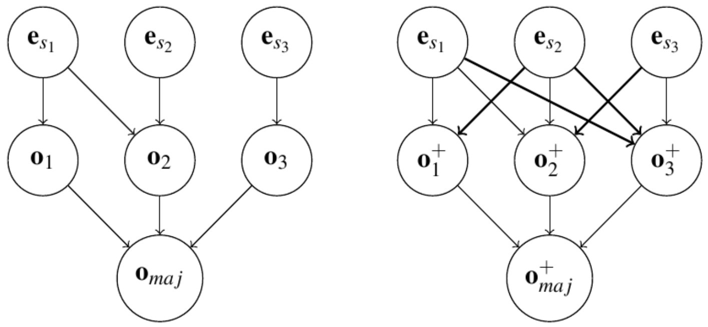
Its correctness probability $p_{eff} = Pr(\mathbf{o}_{eff} = \mathbf{x})$ is the *efficient competence*. Efficiency corresponds to what is sometimes called *full information equivalence*, with 'full information' understood as 'available information' (e.g., Barelli et al. 2022).

The ideal and efficient opinion are in fact the rational opinions based on total evidence $(\mathbf{e}_s)_{s\in S}$ resp. available evidence $(\mathbf{e}_s)_{s\in\cup_i \mathbf{S}_i}$, assuming a simple opinion structure. This can be shown in analogy to Theorem 1.

3.4 Deliberation as sharing and absorbing

We construe group deliberation as a process of evidence transmission. To capture this idea, we now define the notion of a *share-absorb process*. Such a process is given by parameters of two types, namely, for each source $s \in S$ and person $i \in N$, a 'sharing probability' $p_{s,i\rightarrow}$ and an 'absorbing probability' $p_{s,i\leftarrow}$, both in $[0, 1]$. The process transforms the initial source profile $(\mathbf{S}_i)$ into a post-deliberation source profile $(\mathbf{S}_i^+)$, in two steps. First, each person i shares each of her initial sources $s \in \mathbf{S}_i$ with an independent probability of $p_{s,i\rightarrow}$. Second, for each source s shared by at least someone, each person i with $s \notin \mathbf{S}_i$ absorbs s with an independent probability of $p_{s,i\leftarrow}$. The new source set of a person i contains i's initial sources *and* i's absorbed sources: $\mathbf{S}_i^+ = \mathbf{S}_i \cup \{s \in S : i \text{ absorbs } s\}$. The process is defined more formally in Appendix B. Figure 2 gives an illustration, which starts from the introductory example and adds a post-deliberation stage. Thick arrows indicate new sources absorbed during deliberation. The three persons' source sets grow from $S_1 = \{s_1\}$, $S_2 = \{s_1, s_2\}$ and $S_3 = \{s_3\}$ pre-deliberation to $S_1^+ = \{s_1, s_2\}$, $S_2^+ = \{s_1, s_2, s_3\}$ and $S_3^+ = \{s_1, s_2, s_3\}$ post-deliberation. To anticipate later sections, this mitigates Failures 1, 2, and 3, because – roughly speaking – sources and their complementarities have become more widely accessible.

Sharing or absorbing a source can be easy or hard, take seconds or hours, and involve verbal or non-verbal communication. For instance, statistical facts might be easier to share or absorb than complex arguments. So the probabilities $p_{s,i\rightarrow}$ and $p_{s,i\leftarrow}$ can be source-dependent. They can also be person-dependent, partly because some persons are more able or willing than others to share or absorb.



(a) Pre-Deliberation (b) Post-Deliberation

Figure 2: Example of growing access to evidence by deliberation

A share-absorb process generates a new opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i^+))$ with richer source sets $\mathbf{S}_i^+$. Like the initial opinion structure, the new one induces derivative constructs, namely opinions, competence levels, and (as will soon be seen) imbalance measures capturing failures. They are defined as usual, but based on the new opinion structure; we denote them by the usual symbol with an additional superscript '+'. Specifically, any person i has new opinion

$$\mathbf{o}_i^+ = \begin{cases} 1 & \text{if } \sum_{s\in\mathbf{S}_i^+} \mathbf{e}_s > 0 \\ -1 & \text{if } \sum_{s\in\mathbf{S}_i^+} \mathbf{e}_s < 0 \\ 0 & \text{if } \sum_{s\in\mathbf{S}_i^+} \mathbf{e}_s = 0 \end{cases}$$

and new competence p_i^+ $(= Pr(\mathbf{o}_i^+ = \mathbf{x}))$, resulting in a new group opinion $\mathbf{o}_{maj}^+$ and competence p_{maj}^+ $(= Pr(\mathbf{o}_{maj}^+ = \mathbf{x}))$. This machinery will allow us to operationalise our enquiry into the effects of deliberation. For instance, whether deliberation is beneficial overall depends on whether $p_{maj}^+ > p_{maj}$.

Clearly, deliberation creates cross-personal correlations of sources. So, the new opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i^+))$ violates Independent Sources, hence is no longer of the simple kind (we come to general opinion structures in Section 6). Nonetheless our rationality result (Theorem 1) continues to apply, so that post-deliberation opinions remain rational, as shown in Appendix A.

3.5 Comparison with the game-theoretic approach

The sharing and absorbing probabilities $p_{i,s\rightarrow}$ and $p_{i,s\leftarrow}$ of our share-absorb process could be taken to emerge from (equilibrium) behaviour in an underlying 'deliberation game'. The nature of this game varies considerably with the intended interpretation. For instance, one might be tempted to interpret sharing and absorbing as *actions* (choices). Then the sharing and absorbing probabilities could represent mixed strategies in a suitable dynamic game with imperfect information, or represent probabilities of player types that share/absorb, relative to a different game.[5] Yet an 'action interpretation' is often inappropriate because individuals cannot control whether they succeed in sharing or

[5] This dynamic game might have the following structure. *Stage 1*: nature randomly draws a state x, evidences $(e_s)_{s\in S}$, personal source sets $(S_i)_{i\in N}$, and some personal character traits (types), where each person (player) i is informed only of her evidence bundle $(e_s)_{s\in S_i}$ and her type. *Stage 2*: simultaneously,

absorbing arguments or intuitions. A game-theoretic underpinning then requires a less parsimonious game. Actions are now *attempts* to share or absorb, the success of which depends on chance moves and possibly player types; $p_{i,s\rightarrow}$ and $p_{i,s\leftarrow}$ then represent probabilities of a successful attempt to share or absorb, i.e., of the combination of an action and a chance move.

For some important forms of deliberation, sharing and absorbing cannot be underpinned game-theoretically for principled reasons. Deliberation often involves growing awareness since individuals learn 'surprising' information that they did not even consider (and so could not anticipate strategically). Standard game-theory notoriously rules out changing awareness and genuine 'surprises'.[6] It reduces deliberation to mere 'information transmission', without awareness change. An emerging unorthodox branch of game theory tackles awareness growth, fundamentally revising the notions of game and equilibrium (e.g., Feinberg 2021). We do not model the awareness dynamics of deliberation, to focus on our main goals. However, we permit low awareness – no one needs to know the set S of possible sources of evidence, let alone the opinion structure.

In sum, we provide no game-theoretic underpinning of share-absorb processes so as to gain parsimony and, more importantly, allow for various psychological interpretations and phenomena, including ones ruled out by standard game theory. For instance, failures to share or absorb could stem from a conscious choice, or unsuccessful attempt, or inability to even attempt. Moreover, deliberators could be instrumentally or intrinsically motivated, have stable or variable preferences, reason strategically or not, be fully rational or use simple heuristics, and acquire only new information or even new awareness and concepts. Game-theoretic models need to commit on all these issues, often in simplified ways that seem in tension with thinking in democratic theory about the cognitive and motivational structure of deliberation (e.g., Cohen 1996; Gutmann and Thompson 1996). The game-theoretic approach, of course, has its own advantages. For example, Landa and Meirowitz (2009) show how game theory reveals which institutional proposals are strategically stable, though this is not a question investigated in this paper.

Overall, the game-theoretic approach emphasises *decisions* to share, while we emphasise *abilities* to share (express, describe) and to absorb (understand, incorporate), and effects of deliberation on the information distribution and on individual and collective competence.

---

each person i chooses which sources in S_i she shares. *Stage 3:* simultaneously, each person i chooses which sources she absorbs among the sources that she did not acquire in Stage 1 and that someone shared in Stage 2. One could include a final *Voting Stage*: simultaneously, everyone casts a vote in $\{1, -1, 0\}$. A player's utility function depends on her type, and might reflect that sharing and absorbing are costly and (importantly, to capture an epistemic motivation) that collective outcomes should be successful. If there is a Voting Stage, 'successful' could mean that the voting outcome matches the state. Without the Voting Stage, 'successful' could mean that post-deliberation knowledge in the group is high, as measured for instance by the number of persons whose final opinion matches the state.

[6] All players of a game are supposed to know the game form, hence know what others *can* do. So, a deliberation *game* cannot model 'surprises', in which a player learns things that she did not foresee as possible (and so cannot anticipate strategically). Surprises lead to awareness growth, not just information growth, and are at the heart of many real-life deliberation processes. A failure to absorb evidence can stem from insufficient awareness (which prevents one from 'seeing' or 'understanding' evidence), whereas a successful absorption can stem from sufficient awareness, or from spontaneously growing awareness. An emerging non-classical branch of game theory tackles awareness growth. Instead of taking this (fascinating but complex) route, we proceed non-game-theoretically.

4 The Wisdom of Crowds Pre- and Post-Deliberation: Two Jury Theorems

The wisdom of crowds is often defended by appealing to jury theorems, but the connection to deliberation has so far remained informal. We now present two jury theorems – one pre-deliberation, one post-deliberation. Compared to classical jury theorems, the message will be revisionary at two levels.

For one, the new jury theorems will draw a less optimistic picture, by setting an objective bound to the wisdom of crowds instead of postulating asymptotically infallible groups. However large, the group cannot beat the ideal opinion – i.e., the hypothetical opinion based on total evidence. Yet even the ideal opinion is fallible, because total evidence can lie. Worse, the group can fail to reach the ideal opinion and thus perform 'sub-ideally', because firstly some evidences are accessed by nobody and secondly the accessed evidences are scattered across members and therefore hard to exploit.

Here deliberation steps in, by improving the spread of evidences and thereby helping the group make better use of its evidence and approach the ideal opinion, as our jury theorems suggest. By contrast, classical jury theorems make deliberation appear inessential (as large groups find the truth anyway) or even harmful (by undermining voter independence). This rehabilitation of deliberation is the second revisionary message of our jury theorems.

Our jury theorems operate in the framework of a simple opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ and a share-absorb process, although some generalisations would be possible.

4.1 Pre-Deliberation

Since jury theorems vary the group size n, we straightforwardly extend the simple opinion structure $(\mathbf{x}, (\mathbf{e}_s)_{s\in S}, (\mathbf{S}_i)_{i\in N})$ by letting the set of persons N be the infinite set $\{1, 2, ...\}$, called the 'population'. We then talk of a 'simple opinion structure *for an infinite population*'. In such a structure, we can consider groups $\{1, ..., n\} \subseteq N$ of any finite size $n \geq 1$, with a corresponding majority opinion denoted $\mathbf{o}_{maj,n}$ or simply $\mathbf{o}_{maj}$, and majority competence $Pr(\mathbf{o}_{maj,n} = \mathbf{x})$ denoted $p_{maj,n}$ or simply p_{maj}.

Our first jury theorem says that a finite group performs sub-ideally as long as people are not utterly perfect at accessing sources ('Imperfect Access'), but the group reaches the ideal asymptotically if people are good enough at accessing sources ('Access Competence'). Formally:

Imperfect Access: At least one source $s \in S$ is not surely accessed, i.e., has access probability $p_{i\to s} < 1$ for each person i.

Access Competence: The probability $p_{s\to i}$ that a person $i \in N$ accesses a source $s \in S$ is at least $2^{-1/|S|} + \epsilon$, for some $\epsilon > 0$ independent of i and s.

Pre-Deliberation Jury Theorem: *Given a simple opinion structure* $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ *for an infinite population, the majority competence* $p_{maj,n}$

(a) *is at most the ideal competence* p_{ideal}*, and less than it under Imperfect Access,*

(b) *converges to the ideal competence* p_{ideal} *as* $n \to \infty$ *under Access Competence.*

The ideal competence p_{ideal} in this (and the next) jury theorem takes a simple form. As shown in Appendix C, it is the probability that a standard-normal variable takes a value below $\frac{\sqrt{|S|}}{\sigma}$:

$$p_{ideal} = Pr(\mathbf{o}_{ideal} = \mathbf{x}) = F_{N(0,1)}\left(\frac{\sqrt{|S|}}{\sigma}\right), \tag{2}$$

where $F_{N(0,1)}$ is the standard-normal distribution function. This competence is always below 1, reflecting the objective limits of evidence. It is increasing in the number of sources $|S|$ and decreasing in the noise parameter σ. For instance, it is $p_{ideal} \approx 0.868$ if $|S| = 5$ and $\sigma = 2$.

Access Competence is very demanding. For example, with $|S| = 5$ sources the access probability $p_{s\to i}$ must exceed $2^{-1/5} \approx 0.87$ for all persons i and sources s. However, a deliberating group only needs a much weaker competence assumption. Why?

4.2 Post-Deliberation

Now suppose the group deliberates before voting. So, consider a share-absorb process. To make the process apply to arbitrarily large group sizes n, we assume that its sharing and absorbing probabilities $(p_{s,i\to}, p_{s,i\leftarrow})_{s\in S, i\in N}$ run over the infinite population $N = \{1, 2, ...\}$. We call the so-extended process a share-absorb process *for an infinite population*. For any finite group $\{1, ..., n\} \subseteq N$ (where $n \geq 1$), the extended process induces a standard share-absorb process for this group, defined by the (sub)family of parameters restricted to persons from $\{1, ..., n\}$, i.e., $(p_{s,i\to}, p_{s,i\leftarrow})_{s\in S, i\in\{1,...,n\}}$. This (sub)process generates a post-deliberation source set $\mathbf{S}^+_{i,n}$ for each group member $i \in \{1, ..., n\}$, and hence a post-deliberation opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}^+_{i,n}))$, with personal opinions $\mathbf{o}^+_{i,n}$, personal competences $p^+_{i,n}$, a group opinion $\mathbf{o}^+_{maj,n}$, and a group competence $p^+_{maj,n}$. All these concepts are defined as usual. The extra index 'n' signals the dependence on the current group size n. Crucially, the same person i can (and will) develop different post-deliberation opinions $\mathbf{o}^+_{i,n}$ depending on the size of the deliberating group: the larger the group, the more sources are shared, hence absorbed. Note that personal post-deliberation competence grows with group size: $p^+_{i,n} \leq p^+_{i,n+1} \leq p^+_{i,n+3} \leq \ldots$ This, however, does not automatically translate into growth of *majority* competence $p^+_{maj,n}$, because the new members may be less competent. Still, our post-deliberation jury theorem brings positive news: majority opinions are asymptotically ideal under a far weaker competence condition than the pre-deliberation competence condition of Access Competence. This weaker competence condition pertains not just to people's ability to access sources initially, but also to their ability to absorb sources during deliberation. We use the label 'acquisition' to refer to both phenomena, initial access and later absorption:

Acquisition Competence: Informally, for all persons i and sources s, the person has a high access probability $p_{s\to i}$ *or* a high absorbing probability $p_{s,i\leftarrow}$ (or both). Formally, for all persons $i \in N$ and sources $s \in S$, the product $(1 - p_{s\to i})(1 - p_{s,i\leftarrow})$ is at most $1 - 2^{-1/|S|} - \epsilon$, for some $\epsilon > 0$ independent of i and s.

Crucially, if people violate Access Competence because of too low access probabilities, they can still satisfy Acquisition Competence because their absorbing probabilities can make up for their low access probabilities. Deliberation gives them a second chance to acquire sources. Formally:

Proposition 1 *Acquisition Competence is strictly weaker than Access Competence.*

Proof. Given Access Competence, Acquisition Competence holds because, for any $i \in N$ and $s \in S$, $(1 - p_{s\to i})(1 - p_{s,i\leftarrow}) \leq 1 - p_{s\to i} \leq 1 - 2^{-1/|S|} - \epsilon$, where the second '$\leq$' uses Access Competence. Acquisition Competence is *strictly* weaker because under many parameter constellations only Acquisition Competence holds (example: $p_{s\to i} = 0$ and $p_{s,i\leftarrow} = 1$ for all s and i). ■

Our result also uses a minimal condition on participation: new group members do not stop sharing in the limit. This ensures that larger groups have a richer, more diverse deliberation. Technically, recall that $p_{s,i\to}$ represents the *conditional* sharing probability given that the source was accessed in the first place. Our condition pertains instead to the *unconditional* sharing probability, i.e., the probability of accessing and sharing the source, $p_{s\to i} \times p_{s,i\leftarrow}$. We now state our condition, followed by the jury theorem.

Non-Vanishing Participation: For each source $s \in S$, the probability that a person i accesses and shares s, $p_{s\to i} \times p_{s,i\to}$, does not tend to 0 as $i \to \infty$.[7]

Post-Deliberation Jury Theorem: *Given a simple opinion structure* $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ *and a share-absorb process, both for an infinite population, the post-deliberation majority competence* $p^+_{maj,n}$

(a) *is at most the ideal competence* p_{ideal}*, and less than it under Imperfect Access,*

(b) *converges to the ideal competence* p_{ideal} *as* $n \to \infty$ *under Acquisition Competence and Non-Vanishing Participation.*

By this theorem, the interplay of deliberation and group increase makes the group opinion asymptotically ideal under interesting conditions. Going beyond ideal group opinions remains impossible, no matter how much the group deliberates or is increased, because of objectively limited evidence.

An upshot is that deliberation can lead to asymptotically ideal majority opinions even when people are arbitrarily bad at accessing sources (so that Access Competence fails), provided that during deliberation they absorb sources well enough and participate at least minimally, i.e., Acquisition Competence and Non-Vanishing Participation hold.

4.3 Closing the competence gap: by deliberation or group increase?

Group competence usually falls short of ideal competence. The difference $p_{ideal} - p_{maj}$ defines the *competence gap*. To reduce it, two instruments are available: deliberation and group increase. How do they complement one another? Recall that the efficient opinion is the opinion based on the available source set $\cup_{i=1}^n \mathbf{S}_i$, denoted $\mathbf{o}_{eff,n}$ or just $\mathbf{o}_{eff}$; its correctness probability is the efficient competence, denoted $p_{eff,n}$ or just p_{eff}. Now the competence gap $p_{ideal} - p_{maj}$ decomposes into the sum of two gaps:

- The *efficiency gap* is the gap from the actual to the efficient competence, $p_{eff} - p_{maj}$, which stems from imperfect use of available evidences.
- The *availability gap* is the gap from the efficient to the ideal competence, $p_{ideal} - p_{eff}$, which stems from the unavailability of some evidences.

Deliberation is an attempt to reduce the efficiency gap. It cannot reduce the availability gap because it does not 'discover' new sources (formally, because the new available set

[7] This holds for instance if all $p_{s\to i}$ and $p_{s,i\to}$ exceed some fixed level $\epsilon > 0$.

$\cup_{i=1}^{n}\mathbf{S}_i^+$ is no larger than the old one $\cup_{i=1}^{n}\mathbf{S}_i$).[8] The availability gap can instead be reduced by increasing group size. Indeed, $p_{eff,n}$ converges to p_{ideal} as $n \to \infty$, under the minimal assumption that the access probability $p_{s\to i}$ does not converge to 0 as $i \to \infty$. The reason is that, under this assumption of 'non-vanishing access competence', each source is ultimately accessed by *someone* when adding persons.[9] Increasing the group can also reduce the efficiency gap; it even closes this (and the other) gap asymptotically under the fortunate conditions of Access Competence, by the Pre-Deliberation Jury Theorem. But normally Access Competence fails, and a mere group size increase cannot close the efficiency gap – which allows for deliberation. Figure 3 shows how deliberation and

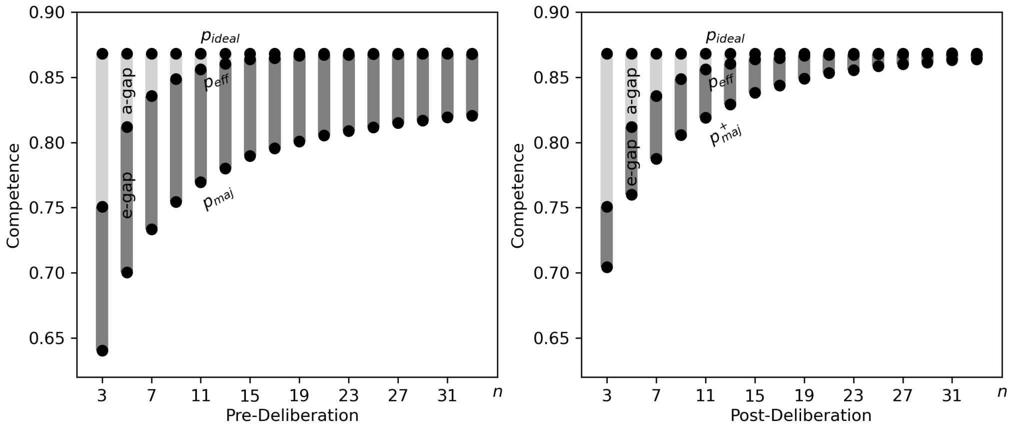


Figure 3: How both competence gaps depend on deliberation and on group size ('e-gap' is short for 'efficiency gap', 'a-gap' for 'availability gap')

group size affect both competence gaps, for some typical parameter values chosen such that Access Competence fails while Acquisition Competence holds.[10] The efficiency gap shrinks considerably though deliberation, as is seen by comparing the pre- and post-deliberation plots. The availability gap is deliberation-invariant but shrinks when adding persons, due to increasing available evidence. The ideal competence p_{ideal} (top line) represents a hard upper bound; it is well below 1, underscoring the objective limitation of evidence. The efficiency gap persists at all group sizes without deliberating (left) but disappears asymptotically with deliberating (right). Exactly this was expected from our jury theorems, as Access Competence fails but Acquisition Competence holds.

5 Failures 1 and 2 analysed formally

Jury theorems cannot reveal the concrete mechanisms by which deliberation helps or harms. We argue that the reduction or increase of Failures 1–3 are part of this mechanism.

[8] Under a broader concept of deliberation (formalised in Section 7), deliberation can also discover sources that nobody held initially, and thereby help close also the availability gap. This would strengthen the case for deliberation further.

[9] With probability one, the available source set $\cup_{i=1}^{n}\mathbf{S}_i$ converges to the full set S as $n \to \infty$.

[10] Specifically, $|S| = 5$, $\sigma = 2$, $p_{s\to i} = 0.2$, $p_{s,i\to} = 0.5$ and $p_{s,i\leftarrow} = 0.85$. Access Competence fails as $p_{s\to i} < 2^{-1/|S|} \approx 0.871$. Acquisition Competence holds as $(1 - p_{s\to i}) \times (1 - p_{s,i\leftarrow}) = 0.12 < 1 - 2^{-1/|S|} \approx 0.129$. The values in Figure 3 were computed using Monte Carlo simulations.

Failures 1–3 are *processing failures*: anomalies in how information is processed. Processing failures should be contrasted with *outcome failures*: deviations of majority outcomes from efficient decisions. Intuitively, processing failures promote outcome failures. The results of this section substantiate this conjecture for Failures 1 and 2, by establishing formal links between these processing failures and 'bad' majority decisions. We first define numerical proxies of Failures 1 and 2 (Section 5.1), and then study the existence and frequency of inefficient majority decisions in the presence of these failures (Section 5.2). Failure 3 is set aside for now.

5.1 Imbalance measures as proxies of Failures 1 and 2

Failures 1 and 2 reflect two forms of imbalance – either between sources (with different spread) or between persons (with unequal evidence). How can both forms of imbalance be measured?

Spread imbalance. Each source s is accessed by some set of persons, to be denoted $\mathbf{N}_s = \{i : s \in \mathbf{S}_i\}$. The *spread* of a source s is the number of source owners $\#\mathbf{N}_s$. The absolute variation of spread between two distinct sources s and t is $|\#\mathbf{N}_s - \#\mathbf{N}_t|$. The relative variation is more relevant. It is calculated by dividing the absolute variation $|\#\mathbf{N}_s - \#\mathbf{N}_t|$ by the average spread $\frac{1}{2}(\#\mathbf{N}_s + \#\mathbf{N}_t)$. Here and elsewhere, divisions of 0 by 0 are handled by setting $\frac{0}{0} = 0$. Now the *spread imbalance* is the average relative variation of spread across all pairs of distinct sources:

$$\begin{aligned}\mathbf{SI} &= \frac{1}{|S|\,(|S|-1)} \sum_{(s,t)\in S^2 : s\neq t} \text{'imbalance in spread between } s \text{ and } t\text{'} \\ &= \frac{1}{|S|\,(|S|-1)} \sum_{(s,t)\in S^2 : s\neq t} \frac{|\#\mathbf{N}_s - \#\mathbf{N}_t|}{\frac{1}{2}(\#\mathbf{N}_t + \#\mathbf{N}_t)}.\end{aligned}$$

Here $|S|\,(|S|-1)$ is the number of pairs (s,t) of distinct sources.

Interpersonal imbalance. Each person i has some total evidence, to be denoted $\mathbf{E}_i = \sum_{s\in\mathbf{S}_i} \mathbf{e}_s$. Her *evidence strength* is her absolute total evidence $|\mathbf{E}_i|$. The absolute variation of evidence strength between two distinct persons i and j is $||\mathbf{E}_i| - |\mathbf{E}_j||$. What matters is, however, the relative variation of evidence strength, obtained by dividing the absolute variation $||\mathbf{E}_i| - |\mathbf{E}_j||$ by the average strength $\frac{1}{2}(|\mathbf{E}_i| + |\mathbf{E}_j|)$. The *interpersonal imbalance* is the average relative variation of evidence strength across all pairs of distinct persons:

$$\begin{aligned}\mathbf{II} &= \frac{1}{n(n-1)} \sum_{(i,j)\in N^2 : i\neq j} \text{'imbalance in evidence strength between } i \text{ and } j\text{'} \\ &= \frac{1}{n(n-1)} \sum_{(i,j)\in N^2 : i\neq j} \frac{||\mathbf{E}_i| - |\mathbf{E}_j||}{\frac{1}{2}(|\mathbf{E}_i| + |\mathbf{E}_j|)}.\end{aligned}$$

Here, $n(n-1)$ is the number of pairs (i,j) of distinct persons.

The imbalance indices as proxies of processing failures. We will use the two imbalance indices – spread imbalance and interpersonal imbalance – as proxies for the extent of Failure 1 and 2, respectively. The rationale is simple: Failure 1 ('overcounting widespread evidence') occurs to the extent that evidences have differently strong spread, which is measured by spread imbalance, and Failure 2 ('neglecting evidential inequality') occurs to the extent that there is evidential inequality, which is measured by interpersonal imbalance.

5.2 How Failures 1 and 2 can harm voting outcomes

In the presence of Failures 1 or 2, majority outcomes can be inefficient: they can be suboptimal in light of the evidence scattered across voters. Let us see why. Consider a simple opinion structure. Majority rule is *inefficient* at an evidence profile $((e_s)_{s\in S_i})$ (a value of $((\mathbf{e}_s)_{s\in\mathbf{S}_i})$) if the majority opinion $\mathbf{o}_{maj}$ differs there from the efficient opinion $\mathbf{o}_{eff}$, which is based on the group's available evidence $(\mathbf{e}_s)_{s\in\cup_i\mathbf{S}_i}$ (see Section 3.3). Our first result shows that Failures 1 and 2 can lead to inefficiency, and that each failure can do so without the other failure:

Proposition 2 *Given a simple opinion structure with non-zero access probabilities and at least two sources and persons,*
(a) the majority opinion is inefficient at some evidence profile at which $\mathbf{SI} > 0$ *and* $\mathbf{II} = 0$*, i.e., at which Failure 1 occurs without Failure 2,*
(b) the majority opinion is inefficient at some evidence profile at which $\mathbf{II} > 0$ *and* $\mathbf{SI} = 0$*, i.e., at which Failure 2 occurs without Failure 1.*

Intuitively, a majority selects the inefficient decision for two different reasons:
- In (a), strong evidence in one direction is spread less widely than weak opposite evidence, without any evidential inequality between voters.
- In (b), a minority of voters holds strong total evidence in one direction while a majority holds weak total evidence in the opposite direction, without any difference in spread between single evidences.

Less technically oriented readers might skip the rest of this subsection. Proposition 2 is a mere existence result: it leaves open the frequency of inefficiencies under Failure 1 or 2. In fact, inefficiencies happen in abundance under Failure 1 or 2. We will show this by focusing on a perhaps more basic property than efficiency. We call majority rule *epistemically monotonic* if an increase in evidential support for an option never harms this option in the outcome: for all evidence profiles $((e_s)_{s\in S_i})$ and $((e'_s)_{s\in S'_i})$,

$$P(\mathbf{x}=1|(e_s)_{s\in\cup_i S_i}) < Pr(\mathbf{x}=1|(e'_s)_{s\in\cup_i S'_i}) \Rightarrow o_{maj} \leq o'_{maj}, \tag{3}$$

where o_{maj} and o'_{maj} denote the majority outcomes in $\{-1,0,1\}$, respectively.[11] If (3) is violated, we talk of a violation of epistemic monotonicity '*at* $((e_s)_{s\in S_i})$' and '*at* $((e'_s)_{s\in S'_i})$'. Epistemic monotonicity is intimately linked to efficiency:

Remark 1 *A violation of epistemic monotonicity at an evidence profile implies inefficiency at this profile or at the other profile of the violation (or at both profiles).*

[11] *Technical detail*: A probability conditional on a random variable (such as $Pr(\mathbf{x}=1|(\mathbf{e}_s)_{s\in\cup_i\mathbf{S}_i})$, where the variable is the vector $(\mathbf{e}_s)_{s\in\cup_i\mathbf{S}_i}$ with random components and random dimensionality) is a function of that variable and is defined only 'almost uniquely', where two versions coincide at all values of that variable except for a set of values of probability zero. The condition (3) should be read as applying to the uniquely existing 'natural' version of $Pr(\mathbf{x}=1|(\mathbf{e}_s)_{s\in\cup_i\mathbf{S}_i})$ that depends continuously on the $\mathbf{e}_s$'s, assuming a simple opinion structure with non-zero access probabilities. So we need not bother with non-uniqueness problems.

Proof. Let (3) be violated, i.e., $Pr(\mathbf{x} = 1|(e_s)_{s\in\cup_i S_i}) < Pr(\mathbf{x} = 1|(e'_s)_{s\in\cup_i S'_i})$ but $o_{maj} > o'_{maj}$. Let o_{maj} be efficient; we show that o'_{maj} is inefficient. As $o_{maj} > o'_{maj}$, we have $o_{maj} \geq 0$ and $o'_{maj} \leq 0$. As o_{maj} is efficient and $o_{maj} \geq 0$, we have $Pr(\mathbf{x} = 1|(e_s)_{s\in\cup_i S_i}) \geq 0$. So $Pr(\mathbf{x} = 1|(e'_s)_{s\in\cup_i S'_i}) > 0$, i.e., outcome 1 is efficient at $(e'_s)_{s\in\cup_i S'_i}$. Hence, o'_{maj} (≤ 0) is inefficient. ■

By Remark 1, the following analysis of monotonicity violations has direct implications for inefficiency. In principle, the support of an evidence profile $((e_s)_{s\in S_i})$ for an option can change in two ways: either the evidence itself changes or the access to evidence changes, i.e., either some evidences e_s change or some source sets S_i change. This leads to two subtypes of monotonicity:

- *Monotonicity in evidence* holds if (3) holds for those evidence profiles $((e_s)_{s\in S_i})$ and $((e'_s)_{s\in S'_i})$ such that $S_i = S'_i$ for all persons $i \in N$.
- *Monotonicity in sources* holds if (3) holds for those evidence profiles $((e_s)_{s\in S_i})$ and $((e'_s)_{s\in S'_i})$ such that $e_s = e'_s$ for all sources s contained in some S_i and some S'_j $(i, j \in N)$.

Failures 1 and 2 can lead to violations of both types of monotonicity, causing inefficiency. Each type can be violated '*at an evidence profile*', in a sense defined like for general epistemic monotonicity.

We first address violations of monotonicity in evidence. At an evidence profile $((e_s)_{s\in S_i})$, let us call an available source $s \in \cup_i S_i$ a *minority source* if fewer persons access s than do not not access s among the persons accessing at least one source, i.e., $\#\{i : s \in S_i\} < \#\{i : s \notin S_i \neq \varnothing\}$.

Proposition 3 *Given a simple opinion structure with non-zero access probabilities and at least three sources, there exist thresholds $\Delta, \Delta' > 0$ such that*

(a) monotonicity in evidence is violated at each evidence profile at which Failure 1 occurs to a degree $\mathbf{SI} \geq \Delta$ *(and Δ is low enough that such profiles exist),*

(b) monotonicity in evidence is violated at each evidence profile with at least one minority source, in such a way that at the other evidence profile of the violation Failure 2 occurs to a degree $\mathbf{II} \geq \Delta'$.

To paraphrase this result informally, monotonicity in evidence is frequently violated in the presence of Failures 1 or 2: it is violated at all evidence profiles with large enough Failure 1, and at almost all evidence profiles when paired with some other evidence profile with large enough Failure 2. In fact, a stronger result holds with precise thresholds given by $\Delta = \frac{6|S|-1}{5|S|-5}$ and $\Delta' = \frac{4}{n+2}$, as shown in Appendix D.

We finally turn to violations of monotonicity in sources in the presence of Failures 1 or 2.

Proposition 4 *Given a simple opinion structure with non-zero access probabilities and at least three persons, there exist thresholds $\Delta, \Delta' > 0$ such that*

(a) monotonicity in sources is violated at each evidence profile with at least two evidences supporting each option[12]*, in such a way that at the other evidence profile of the violation Failure 1 occurs to a degree* $\mathbf{SI} \geq \Delta$,

[12] i.e., with $|\{s \in \cup S_i : e_s > 0\}|, |\{s \in \cup S_i : e_s < 0\}| \geq 2$, where $((e_s)_{s\in \mathrm{Si}})$ denotes the evidence profile

(b) *monotonicity in sources is violated at each evidence profile with at least two evidences supporting each option, in such a way that at the other evidence profile of the violation Failure 2 occurs to a degree* $\mathbf{II} \geq \Delta'$.

So, in short, monotonicity in sources is violated frequently in the presence of Failure 1 or 2: it is violated at almost all evidence profiles when paired with some other evidence profile with large enough Failure 1 or 2. In Appendix D we prove this result in a stronger version with precise thresholds given by $\Delta = 4\frac{n-2}{n|S|}$ and $\Delta' = (3-2\sqrt{2})\frac{n-1}{n}$.

6 A Typology of Beneficial and Harmful Deliberation

This section uses Monte Carlo simulations to help us understand how deliberation can mitigate (or worsen) Failures 1 and 2, and ultimately raise (or lower) group competence. After brief preliminaries (Section 6.1), we will see that share-absorb processes perform well in baseline scenarios (Section 6.2). We then identify four harmful scenarios (Section 6.3), an analysis of which suggests that deliberation should be 'participatory' and 'neutral', but not necessarily 'equal' (Section 6.4).

6.1 Preliminaries

Resulting vs. systemic failures. Our simulation-based analysis will take an ex-ante rather than ex-post perspective on Failures 1 and 2. In what sense? Our indices **SI** (spread imbalance) and **II** (interpersonal imbalance) measure the *resulting or ex-post imbalance*, i.e., the imbalance created as an outcome of the particular values taken by the source sets $\mathbf{S}_i$ and evidences $\mathbf{e}_s$. By contrast, the *systemic or ex-ante imbalance* is the tendency towards resulting imbalance, to be measured by the *expected* resulting imbalance, denoted $\mathcal{SI} = \mathbb{E}(\mathbf{SI})$ resp. $\mathcal{II} = \mathbb{E}(\mathbf{II})$.

Failures 1 and 2 (and even 3) can indeed be understood either as 'resulting' (ex-post) failures or as 'systemic' (ex-ante) failures, i.e., tendencies towards resulting failures. Which understanding matters is context-dependent. While Section 5 focused on resulting failures, using the proxies **SI** and **II**, the current simulation-based section will focus on systemic failures, using the proxies $\mathcal{SI}$ and $\mathcal{II}$.

Our simulation setting. All simulations apply share-absorb processes to simple opinion structures $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$. Previous notation applies. We shall estimate the old and new group competence p_{maj} and p^+_{maj}, and the old and new failure proxies $\mathcal{SI}$, $\mathcal{SI}^+$, $\mathcal{II}$ and $\mathcal{II}^+$ – under various parameter constellations. In principle, one could vary all model parameters: the group size n, source number $|S|$, noise parameter σ, access probabilities $p_{s\to i}$, sharing probabilities $p_{s,i\to}$, and absorbing probabilities $p_{s,i\leftarrow}$. While we have explored several parameter constellations privately, we only report results that vary the parameters $p_{s\to i}$, $p_{s,i\to}$ and $p_{s,i\leftarrow}$, while assuming that $n = 9$, $|S| = 5$ and $\sigma = 2$ (one exception will be highlighted). These choices of n, $|S|$ and σ are rich enough for making meaningful comparisons, and limited enough for inspecting results visually and keeping computational costs low. Our private robustness checks for other values of n, $|S|$ and σ suggest that not much is lost by focusing on our particular values of n, $|S|$ and σ.[13] Our estimates are obtained by taking averages over 1,000,000 rounds of Monte Carlo

[13]Other values of n, $|S|$ and σ affect the extent and frequency of harmful outcomes, but seem not to

simulation (our Python code is available as Supplementary Material).

6.2 Beneficial deliberation in baseline cases

We now present simulation results. They will show that share-absorb processes can generate diverse but not erratic aggregate phenomena, which can be systematised, explained, and exploited for recommendations. The current subsection treats cases of beneficial deliberation; the next two subsections turn to harmful deliberation and recommendations.

We call a share-absorb process:

- *neutral* if its parameters are source-independent. Intuitively, no evidence is privileged.
- *equal* (or *anonymous*) if its parameters are person-independent. Intuitively, everyone takes part equally in deliberation.

The labels 'neutral' and 'equal' can also be applied to sharing alone, or to absorbing alone, or to access, meaning that the corresponding parameters are source- resp. person-independent. Finally, deliberation is:

- *participatory* if every person i shares substantially, in the sense that her average sharing probability $\frac{1}{|S|}\sum_{s\in S} p_{s,i\rightarrow}$ exceeds some threshold δ (of for instance 0.5). For neutral deliberation this condition simplifies: every person i has a (source-independent) sharing probability of $p_{s,i\rightarrow} > \delta$. There are stronger and weaker notions of 'participatory', depending on the choice of δ.

Figure 4 gives examples of how deliberation performs in the baseline case of neutral, equal and participatory scenarios. Here $p_{s\rightarrow i}$, $p_{s,i\rightarrow}$ and $p_{s,i\leftarrow}$ are all independent of s and i, leaving us with just three parameters to vary. In all these scenarios, deliberation raises

	Parameters			Pre-Deliberation			Post-Deliberation			change from...		
#	$p_{s\rightarrow i}$	$p_{s,i\rightarrow}$	$p_{s,i\leftarrow}$	p_{maj}	$\mathcal{SI}$	$\mathcal{II}$	p^+_{maj}	$\mathcal{SI}^+$	$\mathcal{II}^+$	p_{maj} to p^+_{maj}	$\mathcal{SI}$ to $\mathcal{SI}^+$ in %	$\mathcal{II}$ to $\mathcal{II}^+$ in %
1.1	0.5	0.5	0.5	.839	.394	.842	.859	.300	.602	.020	-24.0	-28.5
1.2	0.8	0.8	0.8	.865	.189	.528	.868	.064	.166	.003	-66.3	-68.5
1.3	0.5	0.8	0.8	.839	.395	.842	.867	.146	.344	.028	-63.0	-59.1
1.4	0.2	0.5	0.5	.755	.847	1.239	.800	.953	.825	.045	12.5	-33.4

Figure 4: Results for neutral, equal and participatory deliberation.

majority competence and reduces Failure 2. Deliberation occasionally raises Failure 1, as $\mathcal{SI}$ grows in Scenario 1.4, but the effect does not dominate since group competence still grows.

6.3 Four types of harmful deliberation

Outside the neutral, equal and participatory baseline case, deliberation can become harmful, i.e., lower majority competence. We have identified four elementary types of harmful deliberation; these types and some 'hybrid' types that combine them seem to exhaust the space of harmful share-absorb processes (except for degenerate cases discussed in Section 6.4). Figure 5 gives an example of each elementary harmful type. Remarkably, in all

add entirely new types of harmful deliberation.

#	$p_{s\to i}$	$p_{s,i\to}$	$p_{s,i\leftarrow}$	p_{maj}	$\mathcal{SI}$	$\mathcal{II}$	p^+_{maj}	$\mathcal{SI}^+$	$\mathcal{II}^+$	p_{maj} to p^+_{maj}	$\mathcal{SI}$ to $\mathcal{SI}^+$ in %	$\mathcal{II}$ to $\mathcal{II}^+$ in %
	Parameters			**Pre-Deliberation**			**Post-Deliberation**			**change from...**		
2.1	Fully private evidence											
	1|0	0.5	0.5	.890	0	.879	.874	.726	.780	-.016	∞	-11.3
2.2	Non-participatory deliberation											
	0.2	0.1	1	.754	.847	1.239	.749	1.013	.822	-.005	19.5	-33.6
2.3	Non-neutral sharing 1:4											
	0.2	1|0	1	.755	.847	1.239	.740	1.077	.703	-.015	27.1	-43.3
2.4	Non-neutral absorbing 1:4											
	0.2	1	1|0	.755	.847	1.239	.741	1.077	0.702	-.015	27.2	-43.3

Figure 5: Examples of the four harmful types of deliberation

four examples deliberation raises Failure 1, not 2: deliberation harms by raising source imbalance, not evidential inequality. We now discuss the four elementary harmful types. By a 'scenario' we mean a parameter constellation, i.e., a family of access, sharing and absorbing probabilities $(p_{s\to i}, p_{s,i\to}, p_{s,i\leftarrow})$

Type 1: some private-evidence scenarios. Deliberation harms in some of the scenarios in which most or all members have few or no evidences in common: their source sets have little or no overlap. This happens in Scenario 2.1, in which each person i accesses only one source s_i, i.e., $\mathbf{S}_i = \{s_i\}$ for sure, with $s_i \neq s_j$ if $i \neq j$. (To model this scenario, we let $|S|$ equal $n = 9$ rather than 5.) Part of why deliberation can harm in such scenarios is that the group is already highly competent pre-deliberation. Why? For one, the lack of source overlap creates (state-conditional) independence between votes, so that the law of large numbers kicks in already for relatively small n, lifting majority competence close to 1. For another, Failure 1 is fully absent in scenario 2.1: each source is accessed by exactly one voter. Deliberation unsettles this fine balance, creating dependence between voters as well as Failure 1. Whether this then reduces majority competence, as it does in Scenario 2.1, depends on the precise parameters.[14]

Scenario 2.1 is of special interest because it yields the classic Condorcet jury setting.[15] It is essentially Austen-Smith and Banks' (1996) standard jury model, which follows Condorcet but adds the previously implicit informational basis of opinions.[16] Two insights follow. First, since Scenario 2.1's access structure is artificial, classic jury theorems implicitly rely on an implausible opinion structure. Second, this implicit assumption has skewed the debate about the relevance of deliberation for voting: the cards have been stacked against deliberation. Deliberation is far more useful in reality than is being suggested by classic jury settings. Authors who find that deliberation *is* useful indeed often work in less classical settings (e.g., Barelli et al. 2022 consider complementary evidence,

[14] If for instance all sharing and absorbing probabilities are close to one, deliberation is beneficial, as it tends to give everyone all available information.

[15] It implies Condorcet's controversial assumptions of voter independence and homogeneous competence: voters have independent and identical probabilities above $\frac{1}{2}$ of holding a correct opinion.

[16] Scenario 2.1 and Austen-Smith and Banks' model both let voters access (state-conditionally) independent evidences of homogeneous quality. Scenario 2.1 differs from Austen-Smith and Banks' model in that evidences are Gaussian rather than binary. But this difference is small if each voter has a single evidence and hence need not aggregate different evidences.

as discussed in Section 8.1).

Type 2: some non-participatory scenarios. Deliberation harms in some of the scenarios in which many voters have low average sharing probability. An example is Scenario 2.2. Here, everyone shares any source with probability of only 0.1. Although Scenario 2.2 is equal and neutral in both access and deliberation, deliberation surprisingly harms majority competence, driven by rising spread imbalance (Failure 1). By the combination of low sharing and high absorbing, deliberation puts very few evidences on the table; these are then widely absorbed and become overinfluential, letting Failure 1 rise. Figure 6 gives an example of what can happen: only one source s_3 – with misleading evidence – comes on the table, and spreads fully. Thin arrows indicate initial access, thick arrows indicate new post-deliberation access. The evidence values and post-deliberation opinions are as displayed. The correct option being 1, source s_3 supports the incorrect option, namely -1. By spreading misleading evidence, deliberation turns a correct majority opinion $o_{maj} = 1$ into an incorrect one $o^+_{maj} = -1$. But total available evidence was non-misleading: $e_{s_1} + \cdots + e_{s_5} > 0$. So the problem lies in the evidence management, not the evidence availability.

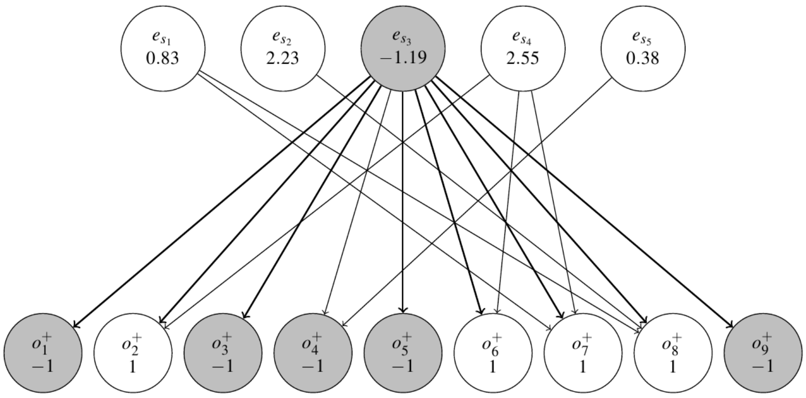


Figure 6: An epistemically harmful outcome of even deliberation

Types 3 & 4: some non-neutral-sharing or non-neutral-absorbing scenarios: Deliberation harms in some of the scenarios in which sources are shared non-neutrally (see Scenario 2.3) or absorbed non-neutrally (see Scenario 2.4). The origin of the problem is in plain sight: some evidences spread overly compared to others, being over-shared or over-absorbed. Deliberation creates a bottleneck where few evidences become dominant, which feeds into Failure 1 and possibly lowers majority competence. The effect is at its worst if very few or just one evidence is put on the table (Scenario 2.3) or picked up (Scenario 2.4).

Although the four harmful scenarios differ structurally, they share two features. First, members have low access probability on average; otherwise enough evidence is available to prevent the negative deliberation effects identified. Second, as mentioned, deliberation always harms through raising Failure 1, not 2. We did privately identify some rare scenarios where deliberation raised Failure 2, but this never translated into falling majority competence *unless* also Failure 1 rose. So the drawbacks of a deliberative rise in Failure 2 seem to be compensated by a deliberative reduction in Failure 1 whenever existent.

6.4 Recommendations and discussion

Our analysis yields a clear-cut recommendation: the group should engage in *participatory and neutral* deliberation, as a means to improve group decisions. Such deliberation is characterised by source-independent sharing and absorbing probabilities $p_{s,i\rightarrow} \equiv p_{i\rightarrow}$ and $p_{s,i\leftarrow} \equiv p_{i\leftarrow}$ ('neutral') and a sufficiently high $p_{i\rightarrow}$ ('participatory'). More precisely, our analysis warrants the following general conjecture:

Conjecture: *Deliberation in the form of a participatory and neutral share-absorb process improves collective competence, given any plausible simple opinion structure as the starting point.*

The Conjecture is warranted because participatory and neutral deliberation blocks the four harmful types of scenario: Types 1 and 2 are only little participatory, and Types 3 and 4 are non-neutral. The Conjecture excludes 'implausible' initial opinion structures, as participatory and neutral deliberation can be non-beneficial for certain highly artificial access parameters. Two such settings stand out. First, deliberation has no effect at all if all sources are certainly accessed by everyone, or more generally if some sources are certainly accessed by everyone and the other sources are never accessed by anyone; in such cases everyone has the same source set, hence learns nothing in deliberation. Second, if some minority of persons certainly accesses the same sources (at least one source) while everybody else never accesses any source, then deliberation harms, because the pre-deliberation majority opinion is the minority's opinion (as the other persons abstain), and this opinion is efficient by being based on all available evidence, whereas the post-deliberation majority opinion can become inefficient.

This second scenario also helps us understand whether deliberation is more beneficial if there is evidential *ex-ante* inequality between voters. Ex-ante inequality is inequality prior to observing evidence. Technically, the group is ex-ante equal if the access-probabilities $p_{s\rightarrow i}$ are person-independent.[17] One might conjecture that deliberation is particularly important under ex-ante inequality, since more 'evidential equalization' work is then to be done. Investigating this conjecture goes beyond this paper's remit. We can however exclude that the conjecture holds *in full generality*: in the second scenario, ex-ante inequality is high and yet deliberation can be harmful (and would become beneficial if people's access probabilities were suitably equalized).

The asymmetry between deliberative effects on $\mathcal{SI}$ and $\mathcal{II}$ is remarkable: according to the simulations, $\mathcal{SI}$ may well increase, which then often harms the majority competence, whereas $\mathcal{II}$ rarely increases, and when it does then the majority competence is not harmed (except if also $\mathcal{SI}$ increases). What explains this asymmetry? Under a share-absorb processes, $\mathcal{SI}$ has an intrinsic tendency for self-reinforcement, because a source that is spread more than others is usually more likely to be shared by someone, and hence often more likely to be newly absorbed by someone (although this tendency is frequently counterbalanced by the fact that highly spread sources have less potential for additional spread). And such an increase in $\mathcal{SI}$ may be collectively harmful, as few evidences come to dominate the decision. By contrast, a self-reinforcement mechanism is harder to identify for $\mathcal{II}$: there is usually no systematic reason for someone with stronger evidence to gain more evidence than others through deliberation – on the contrary, deliberation may allow

[17] A different (weaker) notion of ex-ante inequality is that $\mathcal{II} \neq 0$.

others to catch up in evidence strength, reducing $\mathcal{II}$. And if $\mathcal{II}$ nonetheless increases, then the majority competence need by no means fall, since no evidence was necessarily overcounted. Rather some *voter* was *undercounted* (by neglecting her evidential superiority). A raise in $\mathcal{II}$ does not tend to make majority outcomes worse, but rather tends to make majority *rule* a worse voting rule in comparison, since voters become more heterogeneous in competence and thus rival voting rules with competence-sensitive voting power suddenly perform much better than majority rule.

7 A Generalised Framework

Important real phenomena go beyond the framework used above, with respect to both opinion formation and deliberation. To name just a few limitations, simple opinion structures preclude irrational opinions (cf. Theorem 1) and a treatment of Failure 3 (as will be seen). In addition, share-absorb processes preclude deliberative phenomena such as discovery of sources outside everyone's initial access, communication within subgroups or networks, and sharing or absorbing with a bias towards some option. To capture such phenomena and prepare our analysis of Failure 3, we now generalise our model of opinion formation (Sections 7.1 and 7.2) and deliberation (Section 7.3).

7.1 General opinion structures

No major departure from simple opinion structures $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ is needed to model opinion formation very generally. It suffices to lift assumptions that were made 'for simplicity'. For instance, by no longer assuming that all $\mathbf{e}_s$ correlate with the state, we can model irrational opinions that are affected by 'noises', i.e., state-independent variables $\mathbf{e}_s$ without objective evidential value. This, for instance, allows modelling verdicts of jurors influenced by the defendant's skin colour, the room temperature, or other noises. In general, $(\mathbf{e}_s)$ will then consist of 'influences', be they evidences or noises. Further, by no longer assuming that all $\mathbf{e}_s$ follow Gaussian distributions, we can model opinion formation as a discrete rather than continuous process, in the simplest case driven by binary influences $\mathbf{e}_s$ taking only the values 1 ('support for 1') and -1 ('support for -1'). In fact, we will not even require evidences to be real-valued.

Specifically, we will lift the three distributional assumptions (Equiprobable States, Simple Gaussian Evidence, and Independent Sources), and we no longer assume that a person i's influences $\mathbf{e}_s$ ($s \in \mathbf{S}_i$) are real numbers that are aggregated additively. So, we replace the additive expression '$\sum_{s\in\mathbf{S}_i} \mathbf{e}_s$' with a general expression '$g((\mathbf{e}_s)_{s\in\mathbf{S}_i})$'. Here, g is called the 'influence aggregator' and is some function transforming any influence bundle into a real number, the 'total influence'. Simple opinion structures implicitly assume real-valued evidences and an additive influence aggregator g given by

$$g((e_s)_{s\in S'}) = \sum_{s\in S'} e_s$$

for any influence bundle $(e_s)_{s\in S'}$ over any source set $S' \subseteq S$.

Formally, a (general) *opinion structure* is thus a quadruple $(\mathbf{x}, (\mathbf{e}_s)_{s\in S}, (\mathbf{S}_i)_{i\in N}, g)$, in short $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i), g)$, that contains:

(1) a random variable $\mathbf{x}$, the *state* or *correct option*, taking the value 1 or -1 with arbitrary non-zero probabilities;

(2) a family $(\mathbf{e}_s)$, indexed by some set S of *sources* (non-empty and finite), consisting of random variables, the *influences* from these sources, where each $\mathbf{e}_s$ ranges over some (discrete or continuous) space $\mathcal{E}_s$;

(3) a family $(\mathbf{S}_i)$, indexed by some set $N = \{1, ..., n\}$ of *persons* $(1 \leq n < \infty)$, consisting of random subsets of S, the *source sets* of these persons, again with arbitrary distributions;

(4) a function g, the *influence aggregator*, mapping any influence bundle $(e_s)_{s\in S'}$ $(S' \subseteq S)$ to its 'total influence' $g((e_s)_{s\in S'})$ (technically, a function from $\cup_{S'\subseteq S}\Pi_{s\in S'}\mathcal{E}_s$ to $\mathbb{R}$ that is measurable[18]).

Allowing influences $\mathbf{e}_s$ to take values other than real numbers offers additional modelling flexibility, since influences such as arguments, sensory perceptions or even 'moods' naturally take non-numerical values.[19]

In the default case of an additive influence aggregator g with real-valued influences, we abbreviate the opinion structure by '$(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$', leaving g implicit. Examples of this are *simple* opinion structures, which moreover satisfy the three distributional conditions. An influence $\mathbf{e}_s$ is called a *noise* if it is independent of the state $\mathbf{x}$ (even conditional on the other influences), and an *evidence* otherwise.

As different influences can now be (state-conditionally) dependent, aggregating (real-valued) influences additively can in fact be irrational. For instance, *positively* dependent influences, say from similar sources, are best aggregated subadditively, to avoid double-counting. It is now clear why our generalised notion of opinion structure allows g to be non-additive: otherwise we would require irrational responses to correlated evidences. Still, g *could* be additive, even for correlating evidences. In sum, our model is very flexible and can capture irrational or rational opinions, formed using simple heuristics or sophisticated evidence aggregation.

Our entire earlier machinery carries over to a general opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i), g)$. The *opinion* of a person i is determined by her (now possibly non-additive) aggregate influence:

$$\mathbf{o}_i = \begin{cases} 1 & \text{if } g((\mathbf{e}_s)_{s\in\mathbf{S}_i}) > 0 \\ -1 & \text{if } g((\mathbf{e}_s)_{s\in\mathbf{S}_i}) < 0 \\ 0 & \text{if } g((\mathbf{e}_s)_{s\in\mathbf{S}_i}) = 0. \end{cases}$$

All other derivative concepts – notably the majority opinion $\mathbf{o}_{maj}$, personal competence p_i, majority opinion $\mathbf{o}_{maj}$ and competence p_{maj}, ideal opinion $\mathbf{o}_{ideal}$ and competence p_{ideal}, efficient opinion $\mathbf{o}_{eff}$ and competence p_{eff}, and spread imbalance **SI** or $\mathcal{SI}$ – keep their original definitions, except that the definition of interpersonal imbalance **II** or $\mathcal{II}$ should be generalised by using g instead of summation to aggregate personal influences. One should now more properly call $(\mathbf{e}_i)_{i\in\mathbf{S}_i}$ person i's *influence bundle* and call $((\mathbf{e}_i)_{i\in\mathbf{S}_i})$ the *influence profile*, since the earlier labels 'evidence bundle' and 'evidence profile' neglect the possibility of non-evidential influences.

[18] The co-domain $\cup_{S'\subseteq S}\Pi_{s\in S'}\mathcal{E}_s$ carries the natural σ-algebra, i.e., the union σ-algebra of the spaces $\Pi_{s\in S'}\mathcal{E}_s$ $(S' \subseteq S)$, where each $\Pi_{s\in S'}\mathcal{E}_s$ carries the product σ-algebra of the spaces $\mathcal{E}_s$ $(s \in S')$.

[19] For instance, arguments might take sentences as values. Simple opinion structures effectively capture and represent an evidence by its degree of support for option 1 over option -1, which is indeed a real number. General opinion structures allow modelling evidences and other influences as the 'raw' data or phenomena that they initially are.

7.2 Rationality in general opinion structures

Less technically oriented readers might skip this subsection. Part of the point of general opinion structures was to allow for irrational influences on opinions, in the form of noise rather than evidence. But let us return here to the fully rational paradigm of opinion formation. In *simple* opinion structures, aggregating one's evidences *additively* was rational, by Theorem 1. What sort of non-additive aggregator g is rational in general opinion structures with possibly correlated or even non-real-valued $\mathbf{e}_s$'s? A general Bayesian answer is that the total evidence contained in a bundle $((e_s)_{s\in S'})$ is the log-likelihood-ratio

$$g((e_s)_{s\in S'}) = \log \frac{f_{S'}((e_s)_{s\in S'}|1)}{f_{S'}((e_s)_{s\in S'}|-1)}, \tag{4}$$

where, for each $S' \subseteq S$, $f_{S'}(\cdot|x)$ is a probability mass or density function of the vector $(\mathbf{e}_s)_{s\in S'}$ given state x.[20] A so-defined aggregator g creates classically rational opinions (as defined in Section 3.2), assuming that $Pr(\mathbf{x} = 1) = \frac{1}{2}$ ('Equiprobably States') and every person i's source set $\mathbf{S}_i$ is independent of the state and influences, i.e., of $(\mathbf{x}, (\mathbf{e}_s)_{s\in\mathbf{S}})$ (a condition weaker than 'Independent Sources'). The reason is, in short, that the log-likelihood ratio of someone's information exceeds 0 just in case the likelihood-ratio exceeds 1, which happens just in case state 1 is more likely than state -1, given the information.[21]

Let us give a concrete formula for g in the special case of a 'generalised simple opinion structure', in which the 'Simple Gaussian Evidence' property is generalised such that different evidences can be correlated and of different (expected) strength. That is, the distribution of the evidences is still Gaussian, but with arbitrary means and correlations, given the state. More precisely:

Generalised Gaussian Evidence: Given any state $x \in \{\pm 1\}$, the evidences $\mathbf{e}_s$ $(s \in S)$ are real-valued with a multivariate Gaussian distribution on $\mathbb{R}^S$ with mean vector $x\mu$ and covariance matrix Σ, for some vector $\mu = (\mu_x)_{x\in S} \in \mathbb{R}^S$ and some positive-definite matrix $\Sigma = (\Sigma_{s,t})_{s,t\in S} \in \mathbb{R}^{S\times S}$ (with μ and Σ independent of x).

Thus, given any state $x \in \{\pm 1\}$, the evidences $\mathbf{e}_s$ are still Gaussian variables over $\mathcal{E}_i = \mathbb{R}$, but now with any means $\mathbb{E}(\mathbf{e}_s|x) = x\mu_s$ (rather than the same mean of x) and with any covariances given by the off-diagonal entries of Σ (rather than zero covariance). The simple condition emerges if $\mu_s = 1$ for all $s \in S$ and $\Sigma_{s,t} = 0$ for $s \neq t$ while $\Sigma_{s,t} = \sigma^2$ for $s = t$. The generalised condition will rationalise a non-additive opinion formation in which each evidence bundle $(e_s)_{s\in S'} \in \mathbb{R}^{S'}$ $(S' \subseteq S)$ is aggregated into the total evidence of

$$g((e_s)_{s\in S'}) = (\mu|_{S'})^T(\Sigma|_{S'})^{-1}(e_s)_{s\in S'}, \tag{5}$$

where $(e_s)_{s\in S'}$ is the evidence bundle seen as a column vector, $\mu|_{S'} = (\mu_s)_{s\in S'}$ is the subvector of μ with dimensions in S', again seen as a column vector, $(\mu|_{S'})^T$ is its transpose (a row vector), and $\Sigma|_{S'} = (\Sigma_{s,t})_{s,t\in S'}$ is the submatrix of Σ with row and column

[20] Density functions are of course defined w.r.t. a measure on the co-domain $\Pi_{s\in S'}\mathcal{E}_s$ of $(\mathbf{e}_s)_{s\in S'}$, e.g., the Lebesgue measure on $\mathbb{R}^{S'}$ if all evidences are real-valued.

[21] The second step uses Equiprobably States. The first step uses our independence assumption on each $\mathbf{S}_i$, which guarantees that a person i learns nothing about the state from $\mathbf{S}_i$, i.e., from the fact of *which* sources she could access, so that only the $\mathbf{e}_s$'s carry information about the state. Without assuming Equiprobable States, rational opinions are obtained through amending (4) by adding the prior log-odds $\log \frac{Pr(\mathbf{x}=1)}{Pr(\mathbf{x}=-1)}$.

dimensions in S' (an $S' \times S'$ matrix), and $(\Sigma|_{S'})^{-1}$ is its inverse. One should interpret $(\mu|_{S'})^T(\Sigma|_{S'})^{-1}(e_s)_{s\in S'}$ as 0 if $S' = \varnothing$, i.e., if $\mu|_{S'}$, $(e_s)_{s\in S'}$, and $\Sigma|_{S'}$ are degenerate 0-dimensional objects.

Formally, we call an opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i), g)$ *generalised simple* if it satisfies Equiprobable States, Independent Sources, Generalised Gaussian Evidence, and g is given by (5) or a positive multiple of (5). Of these four properties, the first two are shared with simple opinion structures and the last two are more general.

Theorem 2 *Under a generalised simple opinion structure* $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i), g)$*, the opinion* $\mathbf{o}_i$ *of any person* i *is classically rational.*

The rational evidence aggregator g given by (5) has some notable properties. First, under a simple opinion structure, g reduces to the additive evidence aggregator up to a multiplicative constant, i.e., $g((e_s)_{s\in S'}) = \frac{1}{\sigma^2}\sum_{s\in S'} e_s$ for all $(e_s)_{s\in S'}$.[22] Second, although g is usually non-additive, it is *linear*: $g((e_s)_{s\in S'})$ is a linear combination of the e_s's ($s \in S'$). The coefficients reflect the means and correlations of evidences, to account for evidence strength and avoid double-counting highly correlated evidences.

7.3 General deliberation processes

Deliberation may involve phenomena that go beyond a share-absorb process, such as: (1) the discovery of entirely new arguments, aspects or other sources outside anyone's initial awareness (Goodin 2017; more generally: Müller 2018), (2) non-public deliberation, in subgroups or networks, or (3) evidence-sensitive (possibly biased) sharing, where sources are shared only if they provide evidence of certain strength or direction. While these three phenomena can be captured by suitably generalised share-absorb processes, they call for a unified notion of 'deliberation process' that can accommodate these phenomena and others. We now spell out the three processes capturing (1)–(3), before presenting our unified notion of 'deliberation process'. Throughout we presuppose an arbitrary initial opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i), g)$. The three generalised share-absorb processes (and later the unified process) each generate a new source profile $(\mathbf{S}_i^+)$. How?

- *To model (1)*, we add a discovery stage to the share-absorb process, located between sharing and absorbing. Here is an example of the three-stage process: two arguments are shared, then this sparks the discovery of a third argument, and finally the three arguments on the table are selectively absorbed by members. Formally, besides the usual sharing and absorbing probabilities $p_{s,i\rightarrow}$ and $p_{s,i\leftarrow}$, we introduce probabilities $p_{S\rightarrow s}$ of discovering a sources $s \in S$ after a set of sources $T \subseteq S\backslash\{s\}$ was shared. These sharing, absorbing, and discovering probabilities jointly induce a *share-discover-absorb* process.[23] This can, for instance, model the likely discovery of an argument s after such-and-such arguments T are placed on the table: just set the discovery probability $p_{T\rightarrow s}$ high. Standard share-absorb processes emerge if all discovery probabilities are zero.

[22] Because in (5) $\mu|_{S'}$ consists of 1's and $\Sigma|_{S'}$ has diagonal entries σ^2 and off-diagonal entries 0, so that $(\mu|_{S'})^T(\Sigma|_{S'})^{-1} = \frac{1}{\sigma^2}(1, \ldots, 1)$.

[23] Based on an initial source profile (S_i), *first*, persons i share their sources $s \in S_i$ with independent probabilities of $p_{s,i\rightarrow}$; *second*, letting T be the set of sources shared, the non-shared sources $s \in S\backslash T$ are discovered with independent probabilities of $p_{T\rightarrow s}$; *third*, persons i absorb sources s that they did not own and were shared or discovered with independent probabilities of $p_{s,i\leftarrow}$.

- *To model (2)*, we must capture sharing to, or absorbing from, a subgroup. For each source s, person i, and subgroup $J \subseteq N \backslash \{i\}$, consider probabilities $p_{s,i \to J}$ and $p_{s,i \leftarrow J}$ that i shares s to J or absorbs s from J, respectively. These parameters again induce a generalised share-absorb process. This can model not only deliberation in subgroups or networks, but also biased absorbing, where someone absorbs more easily from certain members than from others, perhaps out of prejudice.
- *To model (3)*, the tendency to share or absorb must depend on the influence from the source. For each source s, influence value $e \in \mathcal{E}_s$, and person i, consider probabilities $p_{s,e,i\to}$ and $p_{s,e,i\leftarrow}$ that person i shares (resp. absorbs) source s emitting influence e. These parameters induce a generalised share-absorb process. It can, for instance, model deliberation where only sufficiently strong influences are transmitted: just set $p_{s,e,i\to}$ and $p_{s,e,i\leftarrow}$ to zero if e is not 'strong' in some sense. It can also model biased sharing, where some members i only share sources whose evidence supports option 1 (so that $p_{s,e,i\to} = 0$ if e supports -1) while other members i do the opposite (so that $p_{s,e,i\to} = 0$ if e supports 1). Biased absorbing can be modelled analogously.

What, then, is our unified notion of 'deliberation process' that encompasses all these specific processes and many others? A *deliberation process* is any transformation that stochastically generates a new source profile (S_i^+) based on individual inputs. The input of a person i is anything she has access to, i.e., maximally her influence bundle $(e_s)_{s\in S_i}$. Formally, the process is any (measurable) function D that maps each initial influence profile $((e_s)_{s\in S_i})$, i.e., each value of $((\mathbf{e}_s)_{s\in \mathbf{S}_i})$, to a lottery over source profiles (S_i^+), i.e., profiles of subsets of S. The probability of an (S_i^+) represents how likely (S_i^+) emerges from deliberation, starting from the initial influence profile $((e_s)_{s\in S_i})$. The process generates a new random source profile $(\mathbf{S}_i^+)$, defined as the random source profile whose conditional distribution is $D(((\mathbf{e}_s)_{s\in \mathbf{S}_i}))$ given $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ (hence also given $((\mathbf{e}_s)_{s\in \mathbf{S}_i})$). This generated source profile $(\mathbf{S}_i^+)$ is essentially unique, i.e., its distribution is unique. More generally, the new opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i^+), g)$ is essentially unique, i.e., its (joint) distribution is unique.[24]

8 Evidential complementarity and Failure 3

Failure 3 arises when some sources emit mutually complementary evidences and this complementarity is underappreciated because few or no voters access all these sources simultaneously. The generalised framework permits modelling complementarity and addressing Failure 3. We begin with illustrations (Section 8.1) and then turn to a more systematic analysis (Section 8.2).

8.1 Some illustrations

Let us start with two brief illustrations of complementarities leading to Failure 3. First, consider the generalised simple opinion structure of Section 7.2. Here, evidences combine linearly: the total support for option 1 contained in an evidence bundle $(e_s)_{s\in S'} \in \mathbb{R}^{S'}$

[24] That is, if $(\mathbf{S}_i^+)$ and $(\widehat{\mathbf{S}}_i^+)$ are each generated by D, then $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i^+))$ and $(\mathbf{x}, (\mathbf{e}_s), (\widehat{\mathbf{S}}_i^+))$ have the same (joint) distribution.

($S' \subseteq S$) is a linear combination

$$g((e_s)_{s\in S'}) = \sum_{s\in S'} a_{s,S'} e_s, \tag{6}$$

for certain coefficients $a_{s,S'} \in \mathbb{R}$ ($s \in S' \subseteq S$).[25] Complementaries arise once the weight $a_{s,S'}$ attributed to an evidence e_s depends on S', hence on the kind of other evidences available. If for instance $a_{s,\{s\}}, a_{s',\{s'\}} > 0$ but $a_{s,\{s,s'\}}, a_{s',\{s,s'\}} < 0$, then the evidences e_s and $e_{s'}$ will (if positive) support option 1 *when taken in isolation*, but support option -1 *when combined.* Failure 3 lurks: voters accessing just one of these sources (say, pre-deliberation) fail to use this complementarity.

Second, hidden common causes can often create correlations and ultimately complementarities between evidences, leading to Failure 3 if voters fail to access the complementarities. An insightful illustration of this can be found in Barelli et al.'s (2022) general analysis of voting (in)efficiency under private information. Adapting one of their examples, a group votes on whether to select (option 1) or dismiss (option -1) a political candidate. The candidate stands somewhere on a political left-right axis. We represent her position by a random variable $\mathbf{z}$ on $\mathbb{R}$, with some symmetric distribution around 0 ($\mathbf{z}$ will act as a common cause). As is agreed, acceptance (1) is correct just in case the candidate is non-extreme, i.e., her position $\mathbf{z}$ falls into the interval $[-10, 10]$ of centrist positions. So, $\mathbf{x} = 1$ if $|\mathbf{z}| \leq 10$ and $\mathbf{x} = -1$ if $|\mathbf{z}| > 10$. Each voter i observes a single noisy signal about the candidate's position, represented by an evidence $\mathbf{e}_{s_i} = \mathbf{z} + \epsilon_i$, where ϵ_i is an 'error' variable, which is distributed symmetrically around 0 (and is independent of $\mathbf{z}$). Each voter i thus accesses a single source: $\mathbf{S}_i = \{s_i\}$ with probability 1. She votes for 1 just in case her evidence is 'non-extreme', i.e., $|\mathbf{e}_{s_i}|$ is small enough. While the details depend on the exact distributional assumptions, it is clear that this sort of opinion structure normally leads to significant complementarities: two persons i and j can each own evidence for an extreme candidate (-1) while the combined evidence suddenly supports the opposite (1): just imagine $\mathbf{e}_{s_i} \gg 0$, $\mathbf{e}_{s_j} \ll 0$ and $\mathbf{e}_{s_i} + \mathbf{e}_{s_j} \approx 0$. Such unused complementarities can translate into inefficient voting outcomes – Failure 3 in action. Barelli et al. suggest that such settings call for deliberation. We agree: deliberation tends to reduce Failure 3 by uncovering complementarities. Barelli et al.'s analysis goes much further, by identifying general conditions on the information structure and/or voters' shared utilities such that voting outcomes are asymptotically efficient.

8.2 A more systematic analysis

What exactly is evidential complementarity, and why does deliberation robustly mitigate Failure 3? Informally, different evidences are 'complementary' if their combination contains different information from the aggregate information of these evidences in isolation. In short: the information in aggregate evidence differs from the aggregate information in evidence. For instance, two complementary arguments might be inconclusive in isolation but conclusive in combination, leading to Failure 3 if each voter owns just one of the arguments.[26]

[25] In the notation of Section 7.2, given any $S' \subseteq S$, the numbers $a_{s,S'}$ ($s \in S'$) are the components of the row vector $(\mu|_{S'})^T(\Sigma|_{S'})^{-1}$; in short, $((a_{s,S'})_{s\in S'})^T = (\mu|_{S'})^T(\Sigma|_{S'})^{-1}$.

[26] As discussed in Section 3.5, our framework allows not only for lack of information, but also more radically for unawareness, though without explicitly modelling it. Unawareness of unowned evidence

Instead of defining a measure of Failure 3 (as done for Failures 1 and 2), let us define what an instance of Failure 3 is. The definition will first be stated loosely:

Informal Definition: Failure 3 occurs over a set of sources S' if (the evidence from) S' is complementary and dispersed.[27]

What means 'complementarity' and 'dispersed'? We leave 'complementary' unspecified for a moment. We call S' *dispersed* if all sources in S' are available to the group but not to all members, i.e., $S' \subseteq \cup_i \mathbf{S}_i$ but $S' \nsubseteq \mathbf{S}_i$ for at least one person i. So, someone or even everyone does not fully access S', hence fails use the complementarity in $(\mathbf{e}_s)_{s \in S'}$ (if any). Worse, fewer persons use this complementarity than use any given evidence in $(\mathbf{e}_s)_{s \in S'}$, because anyone who fully accesses S' must access any given $s \in S'$. So, the complementarity is *underused* compared to isolated evidences. Exactly this constitutes the failure. The parallel to Failure 1 is clear: While in Failure 1 some evidence is less accessed and thus underappreciated, in Failure 3 some complementarity is less accessed and thus underappreciated.

Now assume deliberation creates a new source profile $(\mathbf{S}_i^+)$, following any monotonic deliberation process, e.g., a share-absorb process or one of its three extensions in Section 7.3. 'Monotonic' means that nobody loses sources: $\mathbf{S}_i \subseteq \mathbf{S}_i^+$ for all persons i. How is Failure 3 affected? The good news is: Failure 3 occurs over fewer source sets S'. Why? While deliberation does not affect the complementarity of S', it may eliminate the dispersion of S' by transforming people's partial access to S' into full access. The opposite effect is impossible, since monotonic deliberation never reduces anyone's access to sources.

There is a caveat to the claim that monotonic deliberation reduces Failure 3: if deliberation can discover new sources – so that $\cup_i \mathbf{S}_i^+$ grows beyond $\cup_i \mathbf{S}_i$, as is possible for certain deliberation processes such as share-*discover*-absorb processes – then deliberation can create a Failure 3 over particular sets S', namely sets S' that contain newly discovered source from $\cup_i \mathbf{S}_i^+ \setminus \cup_i \mathbf{S}_i$. But this fact can hardly count against deliberation, since the group's initial avoidance of Failure 3 over such sets S' came merely from not yet having discovered part of S'.

The complementarity of an evidence bundle $(e_s)_{s \in S'}$ can be given a precise statistical meaning. How? The central tool is that of *information.* For any evidence bundle $(e_s)_{s \in S'}$, consider a real number $info((e_s)_{s \in S'})$ representing the information in $(e_s)_{s \in S'}$ about the state, i.e., the evidential support for state 1 against state -1. This number could be $g((e_s)_{s \in S'})$, the aggregate evidence according to the opinion structure, or it could be 'statistical information' according to one of the powerful approaches developed in statistics – it will be both if the opinion structure models 'statistically rational' opinion formation. According to the most canonical statistical approach, the information in $(e_s)_{s \in S'}$ is defined by the likelihood-ratio, or equivalently (after changing to a logarithmic scale) by the logarithm of its likelihood-ratio. Formally, adopting this approach, let an *information function* for the opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i), g)$ (or simply for $(\mathbf{x}, (\mathbf{e}_s))$) be a function

usually comes with unawareness of complementarities between owned and unowned evidence. In creating awareness of new evidence, deliberation also brings the complementarity between old and new evidence to awareness.

[27] In saying 'evidence' rather than more generally 'influence', we anticipate that a (proper) complementarity can only exist between evidences, not noises, as will become clear.

$info$ that maps each bundle $(e_s)_{s\in S'}$ in $\cup_{S'\subseteq S}\Pi_{s\in S'}\mathcal{E}_s$ to a real number given by

$$info((e_s)_{s\in S'}) = \log \frac{f_{S'}((e_s)_{s\in S'}|1)}{f_{S'}((e_s)_{s\in S'}|-1)} \tag{7}$$

if $f_{S'}((e_s)_{s\in S'}|1), f_{S'}((e_s)_{s\in S'}|-1) \neq 0$, where, for each source set $S' \subseteq S$, $f_{S'}(\cdot|\mathbf{x})$ is a joint probability density or mass function of the bundle $(\mathbf{e}_s)_{s\in S'}$ conditional on the state $\mathbf{x}$ (with respect to some measure over $\Pi_{s\in S'}\mathcal{E}_s$). For instance, $f_{S'}(\cdot|\mathbf{x})$ could be a multivariate Gaussian density function over $\mathbb{R}^{S'}$, if evidences are Gaussian. The likelihood-ratio tells how much more likely state 1 makes the bundle than state -1. If this ratio is above (resp. below, equal to) 1, then $info((e_s)_{s\in S'})$ is positive (resp. negative, zero), indicating support for 1 (resp. for -1, for neither). When S' is a singleton set $\{s\}$, then we simply write '$info(e_s)$' and talk of the 'information in e_s'.

We call the source set S' or the evidence bundle $(e_s)_{s\in S'}$ *complementarity* if the combined information $info((e_s)_{s\in S'})$ differs from the total isolated information $\sum_{s\in S'} info(e_s)$. Such complementarity is ruled out if evidences are state-conditionally independent.[28] This is why simple opinion structures rule out Failure 3.

The complementarity of a set S' can be improper, being merely inherited from that of some subset $S'' \subsetneq S'$. If for instance three evidences are complementary merely because the first two are complementary while the third is independent, then the triple is improperly complementary. A formal definition of Failure 3 must rule out improper complementarity of S', since there is obviously no *real* neglect of complementarity if an *improperly* complementary set S' is not fully accessed whilst its properly complementary subsets are fully accessed. Formally, a complementary set S' is *properly* complementary if each subset S'' of S' is complementary with the rest of S', i.e., $info((\mathbf{e}_s)_{s\in S'}) \neq info((\mathbf{e}_s)_{s\in S''}) + info((\mathbf{e}_s)_{s\in S'\setminus S''})$.

We are ready to state the above notions and result formally:[29]

Definition 1 *Given an opinion structure* $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i), g)$ *(with an information function* $info$*) and a source set* $S' \subseteq S$*,*

- Dispersion of S' *is the event that* $S' \subseteq \cup_i \mathbf{S}_i$ *and* $S' \nsubseteq \mathbf{S}_i$ *for some* i*, i.e.,* S' *is available to the group but not to all members,*
- Complementarity of S' *is the event that* $info((\mathbf{e}_s)_{s\in S'}) \neq \sum_{s\in S'} info(\mathbf{e}_s)$*, i.e., information in combined evidence differs from total isolated information,*
- Proper complementarity of S' *is the event that* S' *is complementary and moreover each subset* $S'' \subseteq S'$ *is complementary with* $S'\setminus S''$*, i.e.,* $info((\mathbf{e}_s)_{s\in S'}) \neq info((\mathbf{e}_s)_{s\in S''}) + info((\mathbf{e}_s)_{s\in S'\setminus S''})$*.*
- Failure 3 over S' *is the event that* S' *is properly complementary and dispersed.*

Proposition 5 *Given any opinion structure* $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i), g)$ *(with an information function* $info$*), every monotonic deliberation process reduces Failure 3 in the sense that, for all* $S' \subseteq S$*, if Failure 3 occurs over* S' *post-deliberation then Failure 3 already occurred over*

[28] Under state-conditional independence, the joint likelihood function factorises: $f_{S'}((e_s)_{s\in S'}|x) = \prod_{s\in S'} f_{\{s\}}(e_s|x)$ for each state $x \in \{\pm 1\}$. Taking the ratio across states and then the logarithm on both sides yields $info((e_s)_{s\in S'}) = \sum_{s\in S'} info(e_s)$.

[29] Strictly speaking, density function need neither exist nor be unique. But in practical applications, there usually exists a single natural version of $f(\cdot|\mathbf{x})$, and hence of the information function $info$. Hence, the events to be defined based on $info$, i.e., events of complementarity and Failure 3, will exist and be essentially unique in practical applications.

S' pre-deliberation or S' was unavailable pre-deliberation, formally $F3^+_{S'} \subseteq F3_{S'} \cup \{S' \nsubseteq \cup_i \mathbf{S}_i\}$, where $F3_{S'}$ and $F3^+_{S'}$ are the events of Failure 3 over S' pre- resp. post-deliberation and $\{S' \subseteq \cup_i \mathbf{S}_i\}$ is the event that $S' \nsubseteq \cup_i \mathbf{S}_i$.

Proof sketch. Given the assumptions, we must show that $F3^+_{S'} \subseteq F3_{S'} \cup \{S' \nsubseteq \cup_i \mathbf{S}_i\}$, or equivalently that $F3^+_{S'} \cap \{S' \subseteq \cup_i \mathbf{S}_i\} \subseteq F3_{S'}$. Let $C_{S'}$ and $D_{S'}$ denote the events that S' is complementary resp. dispersed pre-deliberation, and $C^+_{S'}$ and $D^+_{S'}$ the corresponding events post-deliberation. By definition, $F3_{S'} = C_{S'} \cap D_{S'}$ and $F3^+_{S'} = C^+_{S'} \cap D^+_{S'}$. Now $C_{S'} = C^+_{S'}$, as complementarity events are unaffected by individual source sets, hence by deliberation. So it suffices to show that $C_{S'} \cap D^+_{S'} \cap \{S' \subseteq \cup_i \mathbf{S}_i\} \subseteq C_{S'} \cap D_{S'}$, or simply that $D^+_{S'} \cap \{S' \subseteq \cup_i \mathbf{S}_i\} \subseteq D_{S'}$. In words: if S' is dispersed post-deliberation and initially already available, then S' is already dispersed pre-deliberation. This holds as $\mathbf{S}_i \subseteq \mathbf{S}^+_i$ for all i by monotonicity. ■

The result simplifies if attention is restricted to *non-discovering* deliberation processes, in which no new sources are discovered, i.e., for which $\cup_i \mathbf{S}^+_i \subseteq \cup_i \mathbf{S}_i$:

Corollary 1 *Given any opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i), g)$ (with an information function $info$), every monotonic and non-disovering deliberation process reduces Failure 3 in the sense that, for all $S' \subseteq S$, if Failure 3 occurs over S' post-deliberation then Failure 3 already occurred over S' pre-deliberation, formally $F3^+_{S'} \subseteq F3_{S'}$.*

Proof. Note $F3^+_{S'} \subseteq \{S' \subseteq \cup_i \mathbf{S}^+_i\} = \{S' \subseteq \cup_i \mathbf{S}_i\}$; the second equality holds because $\cup_i \mathbf{S}^+_i = \cup_i \mathbf{S}_i$ by 'non-discovering' and 'monotonic'. So, in Proposition 5, '$F3^+_{S'} \subseteq F3_{S'} \cup \{S' \nsubseteq \cup_i \mathbf{S}_i\}$' is equivalent to '$F3^+_{S'} \subseteq F3_{S'}$'. ■

The above analysis suggests measuring the degree of complementarity of an evidence bundle $(e_s)_{s \in S'}$ by the gap between information in combined evidence and total isolated information:

$$comp((e_s)_{s \in S'}) = info((e_s)_{s \in S'}) - \sum_{s \in S'} info(e_s).$$

This measure captures the 'relational' rather than 'isolated' information in a bundle, i.e., the information contained in the relationship between evidences rather than in each evidences intrinsically. The evidence is complementary simpliciter whenever the measure takes a non-zero value.

9 Concluding Remarks

Knowledge that groups hold in a dispersed fashion is often used poorly because voting is bad at aggregating the knowledge underlying votes. The fundamental tension between respecting voter equality and achieving factually correct decisions is therefore hard to resolve. Giving up the principle of one-person-one-vote is unpalatable, but holding on to it is – short of rare symmetries – epistemically suboptimal.

We show that deliberation can mitigate the tension and enable electoral democracies and other groups to make better use of evidence. The effect of deliberation on collective decisions can be studied at two levels: the general level of overall correctness probability of outcomes, or the level of specific failures that threaten collective correctness. At the general level, we have presented the (to our best knowledge) first jury theorems about the effect of deliberation on majority decisions. They suggest that deliberation can increase group competence, though not overcoming the objective limits of available evidence. This message is different from that of orthodox jury theorems, which might suggest that deliberation is unnecessary (because large group perform very well anyway) or even harmful (because voter independence is undermined).

We have studied three information processing failures: overcounting widespread evidence, neglecting evidential inequality, and neglecting evidential complementarity. A mixture of theoretic results and simulations supports an overall positive picture: deliberation tends to mitigate these failures. But there are systematic exceptions. A typology of harmful deliberation has been presented. This analysis led to a recommendation for deliberation processes that are *participatory and neutral*: everyone contributes substantially and no evidences are privileged over others.

Interestingly, while our epistemic approach to democracy suggests participatory and neutral deliberation, a procedural-fairness approach to democracy might instead suggest participatory *and equal* deliberation, because participation adds legitimacy to outcomes and equality is a central fairness requirement (e.g., Manin 1987, Christiano 1996, Habermas 1996, Knight and Johnson 1997, Cohen 1997, Young 2002). A hybrid epistemic-procedural approach might therefore recommend participatory, neutral and equal deliberation.

Taken together, our results provide a robust argument in favor of pre-ballot deliberation on epistemic grounds. Deliberation is not only valuable because democratic citizens owe one another reasons, or because the practice of deliberation is intrinsically valuable, or because deliberation helps structuring voter preferences such as to escape voting paradoxes. It is also valuable because it helps groups make better use of their evidence when voting.

References

Aristotle, 1988. *The Politics*. Cambridge Texts in the History of Political Thought. Cambridge: Cambridge University Press.

Austen-Smith, David, and Jeffrey S. Banks. 1996. "Information Aggregation, Rationality, and the Condorcet Jury Theorem," *American Political Science Review* 90 (1): 34–45.

Austen-Smith, David, and Timothy J. Feddersen. 2006. "Deliberation, Preference Uncertainty, and Voting Rules." *American Political Science Review* 100 (2): 209–17. https://doi.org/10.1017/S0003055406062113.

Barelli, Paulo, Sourav Bhattacharya, and Lucas Siga. 2022. "Full Information Equivalence in Large Elections". *Econometrica 9*0(5): 2161–85. https://doi.org/10.3982/ECTA16376 .

Betz, Gregor. 2013. *Debate Dynamics: How Controversy Improves Our Beliefs*. Dordrecht: Springer.

Christiano, Thomas. 1996. "Deliberative Equality and Democratic Order". *Nomos* 38: 251–87.

Chung, Hun, and John Duggan. 2020. "A Formal Theory of Democratic Deliberation." *American Political Science Review* 114 (1): 14–35.

https://doi.org/10.1017/S0003055419000674.
Cohen, Joshua. 1997. "Deliberation and Democratic Legitimacy". In *Deliberative Democracy: Essays on Reason and Politics*, edited by James Bohman and William Rehg, 67–92. Cambridge MA: MIT Press.
Condorcet, Marquis De, Marie Jean Antoine Nicolas de Caritat. 1785. *Essai Sur l'application de l'analyse à La Probabilité Des Décisions Rendues à La Pluralité Des Voix.*
Coughlan, Peter J. 2000. "In Defense of Unanimous Jury Verdicts: Mistrials, Communication, and Strategic Voting." *American Political Science Review* 94 (2): 375–93. https://doi.org/10.2307/2586018.
Dietrich, Franz, and Kai Spiekermann. 2013. "Epistemic Democracy with Defensible Premises." *Economics and Philosophy* 29 (1): 87–120.
Dietrich, Franz, and Kai Spiekermann. 2021. "Jury Theorems." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2021. Metaphysics Research Lab, Stanford University.
https://plato.stanford.edu/archives/win2021/entries/jury-theorems/.
Ding, Huihui, and Marcus Pivato. 2021. "Deliberation and Epistemic Democracy." *Journal of Economic Behavior & Organization* 185 (May): 138–67.
https://doi.org/10.1016/j.jebo.2021.02.020.
Dryzek, John, and Christian List. 2003. "Social Choice Theory and Deliberative Democracy: A Reconciliation." *British Journal of Political Science* 33 (1): 1–28.
Estlund, David M. 2008. *Democratic Authority: A Philosophical Framework*. Princeton: Princeton University Press.
Estlund, David, and Hélène Landemore. 2018. "The Epistemic Value of Democratic Deliberation." *The Oxford Handbook of Deliberative Democracy.* edited by André Bächtiger, John S. Dryzek, Jane Mansbridge, and Mark E. Warren, Oxford: Oxford University Press.
Feinberg, Yossi. 2021. "Games with Unawareness." *The B.E. Journal of Theoretical Economics* 21 (2): 433–88. https://doi.org/10.1515/bejte-2018-0186.
Gerardi, Dino, and Leeat Yariv. 2007. "Deliberative Voting." *Journal of Economic Theory* 134 (1): 317–38. https://doi.org/10.1016/j.jet.2006.05.002.
Goodin, Robert E. 2017. "The Epistemic Benefits of Deliberative Democracy." *Policy Sciences* 50 (3): 351–66. https://doi.org/10.1007/s11077-017-9286-0.
Goodin, Robert E, and Kai Spiekermann. 2018. *An Epistemic Theory of Democracy.* Oxford: Oxford University Press.
Grofman, Bernard, Guillermo Owen, and Scott L Feld. 1983. "Thirteen Theorems in Search of the Truth." *Theory and Decision* 15: 261–78.
Habermas, Jürgen. 1996. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy.* Translated by William Rehg. Oxford: Polity.
Hartmann, Stephan, and Soroush Rafiee Rad. 2018. "Voting, Deliberation and Truth." *Synthese* 195 (3): 1273–93. https://doi.org/10.1007/s11229-016-1268-9.
Hartmann, Stephan, and Soroush Rafiee Rad. 2020. "Anchoring in Deliberations." *Erkenntnis* 85 (5): 1041–69. https://doi.org/10.1007/s10670-018-0064-y.
Hoek, Daniel, and Richard Bradley. 2022. "Million dollar questions: why deliberation is more than information pooling," *Social Choice and Welfare*.
https://doi.org/10.1007/s00355-022-01392-9.

Hong, Lu, and Scott E Page. 2004. "Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers." *Proceedings of the National Academy of Sciences of the United States of America* 101 (46): 16385–89.

Hong, Lu, and Scott E Page. 2012. "Some Microfoundations of Collective Wisdom." In *Collective Wisdom: Principles and Mechanisms*, edited by Hélène Landemore and Jon Elster, 56–71. Cambridge: Cambridge University Press.

Knight, Jack, and James Johnson. 1997. "What Sort of Political Equality Does Deliberative Democracy Require?" In *Deliberative Democracy: Essays on Reason and Politics*, edited by James Bohman and William Rehg, 279–319. Cambridge MA: MIT Press.

Ladha, Krishna K. 1992. "The Condorcet Jury Theorem, Free Speech, and Correlated Votes." *American Journal of Political Science* 36 (3): 617–34.

Landa, Dimitri, and Adam Meirowitz. 2009. "Game Theory, Information, and Deliberative Democracy." *American Journal of Political Science* 53 (2): 427–44. https://doi.org/10.1111/j.1540-5907.2009.00379.x.

List, Christian, and Robert E. Goodin. 2001. "Epistemic Democracy: Generalizing the Condorcet Jury Theorem." *Journal of Political Philosophy* 9 (3): 277–306.

Manin, Bernard. 1987. "On Legitimacy and Political Deliberation". *Political Theory* 15 (3): 338–68. https://doi.org/10.1177/0090591787015003005 .

McLean, Iain, and Fiona Hewitt, eds. 1994. *Condorcet: Foundations of Social Choice and Political Theory.* Aldershot, Hants, England ; Brookfield, VT: Elgar.

Min, John B., and James K. Wong. 2018. "Epistemic Approaches to Deliberative Democracy." *Philosophy Compass* 13 (6): e12497. https://doi.org/10.1111/phc3.12497.

Müller, Julian F. 2018. "Epistemic Democracy: Beyond Knowledge Exploitation." *Philosophical Studies* 175 (5): 1267–88. https://doi.org/10.1007/s11098-017-0910-9.

Page, Scott E. 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies.* Princeton (NJ): Princeton University Press.

Perote-Peña, Juan, and Ashley Piggins. 2015. "A Model of Deliberative and Aggregative Democracy." *Economics & Philosophy* 31 (1): 93–121. https://doi.org/10.1017/S0266267114000418.

Pivato, Marcus. 2016. "Statistical Utilitarianism." In *The Political Economy of Social Choices*, edited by Maria Gallego and Norman Schofield, 187–204. Studies in Political Economy. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-40118-8_8.

Pivato, Marcus. 2017. "Epistemic Democracy with Correlated Voters." *Journal of Mathematical Economics* 72: 51–69. https://doi.org/10.1016/j.jmateco.2017.06.001.

Rafiee Rad, Soroush, and Olivier Roy. 2021. "Deliberation, Single-Peakedness, and Coherent Aggregation." *American Political Science Review* 115 (2): 629–48. https://doi.org/10.1017/S0003055420001045.

Rousseau, Jean-Jacques. 1972. *Du Contrat Social.* Oxford: Oxford University Press.

Young, Iris Marion. 2002. *Inclusion and Democracy.* Oxford: Oxford University Press.

Appendix

A Proof of Theorems 1 and 2 about rationality

We first prove Theorem 1, in fact generalised to *quasi-simple* opinion structures $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$. Such opinion structures are defined exactly like simple ones except that Independent Sources is weakened to *Quasi-Independent Sources*, the condition that the source-access events '$s \in \mathbf{S}_i$' (where $s \in S$ and $i \in N$) are independent across sources s and jointly independent of the state and the evidences, i.e., of $(\mathbf{x}, (\mathbf{e}_s))$. This condition weakens Independent Sources by no longer requiring independence across *persons* of the source-access events. The generalisation ensures that the theorem also captures post-deliberation opinions. Indeed, a share-absorb process transforms a *simple* opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ into a *quasi-simple* one $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i^+))$ – 'quasi' because of interpersonal source dependencies.

We begin by proving an astonishing fact about Gaussian evidences:

Lemma 1 *If an opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ satisfies Simple Gaussian Evidence (e.g., is quasi-simple), then each evidence is proportional to its own log-likelihood-ratio, more precisely*

$$\mathbf{e}_s = \frac{\sigma^2}{2} \log \frac{f(\mathbf{e}_s|1)}{f(\mathbf{e}_s|-1)} \text{ for each } s \in S,$$

where $f(\cdot|x)$ denotes the Gaussian density ('likelihood') function of each evidence $\mathbf{e}_s$ ($s \in S$) given state x ($\in \{\pm 1\}$).

Proof. Let $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ satisfy Simple Gaussian Evidence. Let $s \in S$. Conditional on a state x ($\in \{\pm 1\}$), $\mathbf{e}_s$ is normally distributed with mean x and variance σ^2, hence has a Gaussian density function given by

$$f(e|x) = \frac{1}{\sigma(2\pi)^{1/2}} e^{-\frac{1}{2}\left(\frac{e-x}{\sigma}\right)^2} \text{ for all } e \in \mathbb{R}.$$

We have $\mathbf{e}_s = \frac{\sigma^2}{2} \log \frac{f(\mathbf{e}_s|1)}{f(\mathbf{e}_s|-1)}$ because, for all values $e \in \mathbb{R}$ of $\mathbf{e}_s$,

$$\log \frac{f(e|1)}{f(e|-1)} = \log \frac{\exp\left(-\frac{1}{2}\left(\frac{e-1}{\sigma}\right)^2\right)}{\exp\left(-\frac{1}{2}\left(\frac{e+1}{\sigma}\right)^2\right)} = \log\left[\exp\left(\frac{1}{2}\left(\frac{e+1}{\sigma}\right)^2 - \frac{1}{2}\left(\frac{e-1}{\sigma}\right)^2\right)\right]$$
$$= \frac{1}{2\sigma^2}\left[\left(e^2+2e+1\right) - \left(e^2-2e+1\right)\right] = \frac{2}{\sigma^2}e. \blacksquare$$

Proof of Theorem 1 generalised to quasi-simple opinion structures. Fix a quasi-simple opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$, a person i, and a possible opinion of i, i.e., a random variable $\mathbf{o}$ generating values in $\{1, 0, -1\}$ based on i's evidence bundle $(\mathbf{e}_s)_{s \in \mathbf{S}_i}$. We must show that $\mathbb{E}(u(\mathbf{o}_i, \mathbf{x})) \geq \mathbb{E}(u(\mathbf{o}, \mathbf{x}))$, where the utility of any opinion-state pair (o, x) in $\{1, 0, -1\} \times \{1, -1\}$ is the correctness level, given by

$$u(o, x) = \begin{cases} 1 & \text{if } o = x \text{ (correct opinion)} \\ 0 & \text{if } o = -x \text{ (false opinion)} \\ \frac{1}{2} & \text{if } o = 0 \text{ (neutral opinion).} \end{cases}$$

We prove this by showing that $\mathbb{E}(u(\mathbf{o}_i,\mathbf{x})|(\mathbf{e}_s)_{s\in\mathbf{S}_i}) \geq \mathbb{E}(u(\mathbf{o},\mathbf{x})|(\mathbf{e}_s)_{s\in\mathbf{S}_i})$.[30] So, we fix any value $(e_s)_{s\in S_i}$ of $(\mathbf{e}_s)_{s\in\mathbf{S}_i}$ and, writing o_i (resp. o) for the value of $\mathbf{o}_i$ (resp. $\mathbf{o}$) under $(e_s)_{s\in S_i}$, we must prove that

$$\mathbb{E}(u(o_i,\mathbf{x})|(e_s)_{s\in S_i}) \geq \mathbb{E}(u(o,\mathbf{x})|(e_s)_{s\in S_i}). \tag{8}$$

We prove this in two steps.

Claim 1: $Pr(\mathbf{x}=1|(e_s)_{s\in S_i}) > (<,=)\ \frac{1}{2} \Leftrightarrow \sum_{s\in S_i} e_s > (<,=)\ 0$.

We only prove the equivalence for '>', as those for '<' and '=' are analogous. For any state $x \in \{\pm 1\}$, writing $f(\cdot|x)$ for the Gaussian density function on $\mathbb{R}$ of any $\mathbf{e}_s$ given x, and $g(\cdot|x)$ for the $|S_i|$-dimensional Gaussian density function on $\mathbb{R}^{S_i}$ of the vector $(\mathbf{e}_s)_{s\in S_i}$, we have

$$\begin{aligned} Pr(\mathbf{x}=1|(e_s)_{s\in S_i}) > \tfrac{1}{2} \quad &\Leftrightarrow \quad \frac{g((e_s)_{s\in S_i}|1)}{g((e_s)_{s\in S_i}|-1)} > 1 \\ &\Leftrightarrow \quad \prod_{s\in S_i} \frac{f(e_s|1)}{f(e_s|-1)} > 1 \\ &\Leftrightarrow \quad \sum_{s\in S_i} \log \frac{f(e_s|1)}{f(e_s|-1)} > 0 \\ &\Leftrightarrow \quad \sum_{s\in S_i} e_s > 0. \end{aligned}$$

Here, the first equivalence follows easily from Bayes' rule, using that $Pr(\mathbf{x}=1) = Pr(\mathbf{x}=-1)$ and also that i's source set is independent of the state and the evidences.[31] The second equivalence holds by state-conditional independence of the evidences. The third equivalence holds by applying the logarithm on both sides of the previous inequality. The fourth equivalence holds by Lemma 1. Q.e.d.

Claim 2: The inequality (8) holds (completing the proof).

We proceed case by case.

Case 1: $o_i = 1$, i.e., $\sum_{s\in S_i} e_s > 0$. Then $\mathbb{E}(u(o_i,\mathbf{x})|(e_s)_{s\in S_i}) = Pr(\mathbf{x}=1|(e_s)_{s\in S_i}) > \frac{1}{2}$, where the inequality holds by Claim 1 as $\sum_{s\in S_i} e_s > 0$.

Subcase 1.1: $o = 1$. Then (8) holds (with '=') because $o_i = o$.

Subcase 1.2: $o = -1$. Then (8) holds (with '>') because $\mathbb{E}(u(o,\mathbf{x})|(e_s)_{s\in S_i}) = Pr(\mathbf{x}=-1|(e_s)_{s\in S_i}) = 1 - Pr(\mathbf{x}=1|(e_s)_{s\in S_i}) < \frac{1}{2}$, where the last inequality holds as $Pr(\mathbf{x}=1|(e_s)_{s\in S_i}) > \frac{1}{2}$.

Subcase 1.3: $o = 0$. Then (8) holds (with '>') because $\mathbb{E}(u(o,\mathbf{x})|(e_s)_{s\in S_i}) = \frac{1}{2}$.

Case 2: $o_i = -1$, i.e., $\sum_{s\in S_i} e_s < 0$. Then $\mathbb{E}(u(o_i,\mathbf{x})|(e_s)_{s\in S_i}) = Pr(\mathbf{x}=-1|(e_s)_{s\in S_i}) = 1 - Pr(\mathbf{x}=1|(e_s)_{s\in S_i}) > \frac{1}{2}$, where the inequality holds because $Pr(\mathbf{x}=1|(e_s)_{s\in S_i}) < \frac{1}{2}$ by Claim 1 as $\sum_{s\in S_i} e_s < 0$. An argument similar to that in Case 1 then implies (8).

Case 3: $o_i = 0$, i.e., $\sum_{s\in S_i} e_s = 0$. Then $\mathbb{E}(u(o_i,\mathbf{x})|(e_s)_{s\in S_i}) = \frac{1}{2}$. Further, $Pr(\mathbf{x}=1|(e_s)_{s\in S_i}) = Pr(\mathbf{x}=-1|(e_s)_{s\in S_i}) = \frac{1}{2}$, by Claim 1 as $\sum_{s\in S_i} e_s = 0$.

Subcase 3.1: $o = 0$. Then (8) holds (with '=') because $o_i = o$.

[30] Strictly speaking, we must show that this inequality holds *for some versions* of the conditional expectations on both sides. This qualification is necessary because conditional expectations are random variables that are not unique, but still 'essentially unique' in that any two versions of a conditional expectation coincide outside a zero-probability event.

[31] How is this independence condition used here? Informally, it ensures that the identity of the source set S_i is of no extra information, i.e., that only the evidences from those sources carry information. This explains why the numerator and denominator of the likelihood-ratio each features the joint (state-conditional) likelihood of the evidences only, not of the evidences *and* the source set S_i. To be slightly more explicit, conditionalising on the evidence bundle $(e_s)_{s\in S_i}$ is equivalent to conditionalising first on the source set S_i and then on the evidences from these sources; which however reduces to conditionalising only on the evidences, by the independence condition.

Subcase 3.2: $o = 1$. Then (8) holds (with '=') because $\mathbb{E}(u(o, \mathbf{x})|(e_s)_{s\in S_i}) = Pr(\mathbf{x} = 1|(e_s)_{s\in S_i}) = \frac{1}{2}$.

Subcase 3.3: $o = -1$. Then (8) holds (with '=') because $\mathbb{E}(u(o, \mathbf{x})|(e_s)_{s\in S_i}) = Pr(\mathbf{x} = -1|(e_s)_{s\in S_i}) = \frac{1}{2}$. ■

We now prove Theorem 2, again starting with a lemma.

Lemma 2 *If an opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ satisfies Generalised Gaussian Evidence, then for each $\varnothing \neq S' \subseteq S$ the vector $(\mathbf{e}_s)_{s\in S'}$ has a log-likelihood-ratio given by*

$$\log \frac{f((e_s)_{s\in S'}|1)}{f((e_s)_{s\in S'}|-1)} = 2g((e_s)_{s\in S_i}) \text{ for all } (e_s)_{s\in S'} \in \mathbb{R}^{S'},$$

where $f(\cdot|x)$ denotes the density function of the vector $(\mathbf{e}_s)_{s\in S'}$ given state x ($\in \{\pm 1\}$), i.e., the multivariate Gaussian density function with mean $x\mu|_{S'}$ and covariance matrix $\Sigma|_{S'}$.

Proof. Let $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ satisfy Generalised Gaussian Evidence. Fix $\varnothing \neq S' \subseteq S$. Write μ' for $\mu|_{S'}$ and Σ' for $\Sigma|_{S'}$. Given any $x \in \{\pm 1\}$, $(\mathbf{e}_s)_{s\in S'}$ has an $N(x\mu', \Sigma')$ distribution, with density function given by:

$$f(e|x) = \frac{1}{(\det \Sigma')^{1/2}(2\pi)^{|S'|/2}} \exp\left(-\frac{1}{2}(e - x\mu')^T \Sigma'^{-1}(e - x\mu')\right) \text{ for all } e \in \mathbb{R}^{S'}.$$

So, for all $e \in \mathbb{R}^{S'}$,

$$\begin{aligned}
\log \frac{f(e|1)}{f(e|-1)} &= \log \frac{\frac{1}{(\det \Sigma')^{1/2}(2\pi)^{|S'|/2}} \exp\left(-\frac{1}{2}(e - \mu')^T \Sigma'^{-1}(e - \mu')\right)}{\frac{1}{(\det \Sigma')^{1/2}(2\pi)^{|S'|/2}} \exp\left(-\frac{1}{2}(e + \mu')^T \Sigma'^{-1}(e + \mu')\right)} \\
&= \log \exp\left(\frac{1}{2}\left[(e + \mu')^T \Sigma'^{-1}(e + \mu') - (e - \mu')^T \Sigma'^{-1}(e - \mu')\right]\right) \\
&= \frac{1}{2}\left[e^T \Sigma'^{-1} e + 2\mu'^T \Sigma'^{-1} e + \mu'^T \Sigma'^{-1}\mu' - (e^T \Sigma'^{-1} e - 2\mu'^T \Sigma'^{-1} e + \mu'^T \Sigma'^{-1}\mu')\right] \\
&= 2\mu'^T \Sigma'^{-1} e = 2g(e). \blacksquare
\end{aligned}$$

Proof of Theorem 2 generalised to generalised quasi-simple opinion structures. Let $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ be generalised quasi-simple, i.e., generalised simple except that 'Independent Sources' is weakened to 'Quasi-Independent Sources'. Let $\mathbf{o}$ be a possible opinion of a given person i, i.e., a random variable into $\{1, 0, -1\}$ based on $(\mathbf{e}_s)_{s\in \mathbf{S}_i}$. We show $\mathbb{E}(u(\mathbf{o}_i, \mathbf{x})) \geq \mathbb{E}(u(\mathbf{o}, \mathbf{x}))$, with u defined as before. As in the proof of Theorem 1, it suffices to fix a value $(e_s)_{s\in S_i}$ of $(\mathbf{e}_s)_{s\in \mathbf{S}_i}$, and, writing o_i and o for the value of $\mathbf{o}_i$ resp. $\mathbf{o}$ under $(e_s)_{s\in S_i}$, to prove that

$$\mathbb{E}(u(o_i, \mathbf{x})|(e_s)_{s\in S_i}) \geq \mathbb{E}(u(o, \mathbf{x})|(e_s)_{s\in S_i}). \tag{9}$$

First assume $S_i = \varnothing$. Then (9) reduces to $\mathbb{E}(u(o_i, \mathbf{x})) \geq \mathbb{E}(u(o, \mathbf{x}))$, which holds with '=' because, firstly, $\mathbb{E}(u(o_i, \mathbf{x})) = \mathbb{E}(u(0, \mathbf{x})) = \mathbb{E}(0) = 0$, and secondly, $\mathbb{E}(u(o, \mathbf{x})) = \frac{1}{2}(u(o, 1) + u(o, 0))$, which equals 0 regardless of the value of o, since $o = 1 \Rightarrow [u(o, 1) = 1$ & $u(o, -1) = -1]$, $o = -1 \Rightarrow [u(o, 1) = -1$ & $u(o, -1) = 1]$, and $o = 0 \Rightarrow u(o, 1) = u(o, -1) = 0$.

From now on let $S_i \neq \varnothing$. The proof proceeds in two steps.

Claim 1: $Pr(\mathbf{x} = 1|(e_s)_{s\in S_i}) > (<, =)\ \frac{1}{2} \Leftrightarrow g((e_s)_{s\in S_i}) > (<, =)\ 0$.

We only show the equivalence for '>', as that for '<' or '=' is analogous. For any state $x \in \{\pm 1\}$, writing $f(\cdot|x)$ for the Gaussian density function on $\mathbb{R}^{S_i}$ of the vector $(\mathbf{e}_s)_{s\in S_i}$, we have

$$\begin{aligned} Pr(\mathbf{x}=1|(e_s)_{s\in S_i}) > \tfrac{1}{2} \quad &\Leftrightarrow \quad \frac{f((e_s)_{s\in S_i}|1)}{f((e_s)_{s\in S_i}|-1)} > 1 \\ &\Leftrightarrow \quad \log\frac{f((e_s)_{s\in S_i}|1)}{f((e_s)_{s\in S_i}|-1)} > 0 \\ &\Leftrightarrow \quad g((e_s)_{s\in S_i}) > 0. \end{aligned}$$

Here, the first equivalence holds in analogy to the corresponding equivalence in the proof of Theorem 1. The second equivalence holds trivially. The third equivalence holds Lemma 2. Q.e.d.

Claim 2: The inequality (9) holds (completing the proof).

This claim follows from Claim 1 through an argument analogous to that underlying Claim 2 in the proof of Theorem 1. ■

B Formal Analytics of Share-Absorb Processes

The definition of share-absorb processes has been stated informally. The formalisation is obvious. In short, given a simple opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ (we could have used a general opinion structure), the share-absorb process with parameters $(p_{s,i\rightarrow}, p_{s,i\leftarrow})_{s\in S, i\in N}$ assumes that there exist events 'i shares s' and 'i absorbs s' for any person $i \in N$ and source $s \in S$; that the new source set of any person i is $\mathbf{S}_i^+ = \mathbf{S}_i \cup \{s \in S : \text{'}i \text{ absorbs } s\text{'}\}$, the set of initially accessed or later absorbed sources; that, for any person i and source s, the probability of 'i shares s' given any initial source profile (S_j) is $p_{s,i\rightarrow}$ if $s \in S_i$ and 0 otherwise$_i$; that, for any person i and source s, the probability of 'i absorbs s' given any initial source profile (S_j) and any sharing profile is $p_{s,i\leftarrow}$ if [$s \notin S_i$ and someone shares s in the sharing profile] and 0 otherwise (where a 'sharing profile' is a combination of truth values of the sharing events across persons and sources); and, finally, that the access, sharing, and absorbing events are jointly independent of the state and the evidences.

How is the new source profile $(\mathbf{S}_i^+)$ distributed? And how is it distributed conditional on the initial evidence profile? We now answer both questions. The first answer completes the description of the new opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i^+))$, as we already know how $(\mathbf{x}, (\mathbf{e}_s))$ is distributed and that $(\mathbf{S}_i^+)$ is independent of $(\mathbf{x}, (\mathbf{e}_s))$. The second answer implies an alternative (and equivalent) definition of the share-absorb process as a deliberation process in the general sense of Section 7, i.e., a mapping D from initial evidence profiles to lotteries over new source profiles.

Fix a simple[32] opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ and a share-absorb process with sharing and absorbing probabilities $(p_{s,i\rightarrow}, p_{s,i\leftarrow})_{s\in S, i\in N}$, generating a new source profile $(\mathbf{S}_i^+)$. The probability of any new source profile (S_i^+) (any value of $(\mathbf{S}_i^+)$) is

$$Pr((S_i^+)) = \prod_{s\in S} \pi_s \tag{10}$$

where, for each source $s \in S$, π_s is the probability that the new set of owners of s is

[32] Simplicity could be weakened considerably, to Independent Sources.

$I_s = \{i : s \in S_i^+\}$, and equals

$$\pi_s = \left(\prod_{i \in I_s} p_{s\to i}\right)\left(\prod_{i \in \overline{I_s}} \overline{p_{s\to i}}\right)\left(\prod_{i \in I_s} \overline{p_{s,i\to}}\right)$$
$$+\left(\prod_{i \in \overline{I_s}} \overline{p_{s,i\leftarrow}}\right)\sum_{\varnothing \neq I \subseteq I_s}\left(\prod_{i \in I} p_{s\to i}\right)\left(\prod_{i \in \overline{I}} \overline{p_{s\to i}}\right)\left(\overline{\prod_{i \in I} \overline{p_{s,i\to}}}\right)\left(\prod_{i \in I_s \setminus I} p_{s,i\leftarrow}\right).$$

Further, the conditional probability of any new source profile (S_i^+) given any initial evidence profile $((e_s)_{s\in S_i})$, or given just (S_i), is

$$Pr((S_i^+)|((e_s)_{s\in S_i}) = Pr((S_i^+)|(S_i)) = \prod_{s\in S} \gamma_s \tag{11}$$

where, for each source $s \in S$, γ_s is the probability that the new set of owners of source s is $I_s = \{i : s \in S_i^+\}$ given that the initial one is $J_s = \{i : s \in S_i\}$, which equals

$$\gamma_s = \begin{cases} \left(\overline{\prod_{i\in J_s} \overline{p_{s,i\to}}}\right)\left(\prod_{i\in I_s\setminus J_s} p_{s,i\leftarrow}\right)\left(\prod_{i\in\overline{I_s}} \overline{p_{s,i\leftarrow}}\right) & \text{if } J_s \subsetneq I_s \\ \left(\overline{\prod_{i\in J_s} \overline{p_{s,i\to}}}\right)\left(\prod_{i\in\overline{J_s}} \overline{p_{s,i\leftarrow}}\right) + \prod_{i\in J_s} \overline{p_{s,i\to}} & \text{if } J_s = I_s \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Proof of (10). For each $s \in S$, define two random subgroups, the old set of owners $\mathbf{J}_s = \{i : s \in \mathbf{S}_i\}$ and the new one $\mathbf{I}_s = \{i : s \in \mathbf{S}_i^+\}$. Fix any (S_i^+), and define each I_s ($s \in S$) as above. Note that $(\mathbf{S}_i^+)$ takes the value (S_i^+) if and only if $(\mathbf{I}_s)$ $(= (\mathbf{I}_s)_{s\in S})$ takes the value (I_s) $(= (I_s)_{s\in S})$. Hence, $Pr((S_i^+)) = Pr((I_s))$. The sets $\mathbf{I}_s$ are independent across sources s. So, $Pr((I_s)) = \prod_{s\in S} Pr(I_s)$, and thus

$$Pr((S_i^+)) = \prod_{s\in S} Pr(I_s).$$

Now fix a source s. We calculate $Pr(I_s)$ $(= \pi_s)$. We do this under the assumption that all parameters $p_{s\to i}$ and $p_{s,i\to}$ are strictly between 0 and 1. This is sufficient since the formula generalises to extreme parameter values by a continuity argument.

First assume $I_s = \varnothing$. Note that $\mathbf{I}_s$ takes the value $\varnothing$ if and only if $\mathbf{J}_s$ takes the value $\varnothing$. The probability of the latter is $\prod_i \overline{p_{s\to i}}$. So, $Pr(I_s) = \prod_i \overline{p_{s\to i}}$. This is what had to be proved, since the claimed expression for $Pr(I_s)$ $(= \pi_s)$ indeed reduces to $\prod_i \overline{p_{s\to i}}$ if $I_s = \varnothing$.

From now on assume $I_s \neq \varnothing$. Then it is certain that $\mathbf{J}_s$ takes a value J_s that satisfies $\varnothing \neq J_s \subseteq I_s$, and each such value J_s has non-zero probability (as each $p_{s\to i}$ is strictly between 0 and 1), so can be conditionalised on. By implication,

$$Pr(I_s) = \sum_{\varnothing \subsetneq J_s \subseteq I_s} Pr(J_s)Pr(I_s|J_s). \tag{13}$$

In this expression, the term $Pr(J_s)$ can be written as

$$Pr(J_s) = \left(\prod_{i\in J_s} p_{s\to i}\right)\left(\prod_{i\in\overline{J_s}} \overline{p_{s\to i}}\right).$$

We now calculate $Pr(I_s|J_s)$. Denote by $!_s$ the event that at least someone shares s. Given the (non-empty) event J_s, each of $!_s$ and $\overline{!_s}$ has non-zero probability (as each $p_{s,i\rightarrow}$ is strictly between 0 and 1), so can be conditionalised on. Hence, $Pr(I_s|J_s)$ is writable as $Pr(!_s|J_s)Pr(I_s|!_s, J_s) + Pr(\overline{!_s}|J_s)Pr(I_s|\overline{!_s}, J_s)$, where $Pr(I_s|\overline{!_s}, J_s)$ is 0 if $J_s \neq I_s$ and 1 if $J_s = I_s$. So,

$$Pr(I_s|J_s) = \begin{cases} Pr(!_s|J_s)Pr(I_s|!_s, J_s) & \text{if } J_s \neq I_s \\ Pr(!_s|J_s)Pr(I_s|!_s, J_s) + Pr(\overline{!_s}|J_s) & \text{if } J_s = I_s. \end{cases}$$

Note that if $J_s = I_s$ then

$$Pr(\overline{I_s}|J_s) = \prod_{i \in I_s} \overline{p_{s,i\rightarrow}}.$$

Upon inserting the derived expressions into (13) and rearranging,

$$Pr(I_s) = \left(\prod_{i\in I_s} p_{s\rightarrow i}\right)\left(\prod_{i\in \overline{I_s}} \overline{p_{s\rightarrow i}}\right)\left(\prod_{i\in I_s} \overline{p_{s,i\rightarrow}}\right) + \sum_{\varnothing \subsetneq J_s \subseteq I_s} \left(\prod_{i\in J_s} p_{s\rightarrow i}\right)\left(\prod_{i\in \overline{J_s}} \overline{p_{s\rightarrow i}}\right) Pr(!_s|J_s)Pr(!_s|!_s, J_s).$$

In this,

$$Pr(!_s|J_s)Pr(I_s|!_s, J_s) = \left(\overline{\prod_{i\in J_s} \overline{p_{s,i\rightarrow}}}\right)\left(\prod_{i\in I_s\setminus J_s} p_{s,i\leftarrow}\right)\left(\prod_{i\in \overline{I_s}} \overline{p_{s,i\leftarrow}}\right).$$

So, after rearranging and relabelling the index 'J_s' into 'I',

$$Pr(I_s) = \left(\prod_{i\in I_s} p_{s\rightarrow i}\right)\left(\prod_{i\in \overline{I_s}} \overline{p_{s\rightarrow i}}\right)\left(\prod_{i\in I_s} \overline{p_{s,i\rightarrow}}\right) + \left(\prod_{i\in \overline{I_s}} \overline{p_{s,i\leftarrow}}\right) \sum_{\varnothing \neq I \subseteq I_s} \left(\prod_{i\in I} p_{s\rightarrow i}\right)\left(\prod_{i\in \overline{I}} \overline{p_{s\rightarrow i}}\right)\left(\overline{\prod_{i\in I} \overline{p_{s,i\rightarrow}}}\right)\left(\prod_{i\in I_s\setminus I} p_{s,i\leftarrow}\right). \blacksquare$$

Proof of (11). Fix any initial evidence profile $((e_s)_{s\in S_i})$ and new source profile (S_i^+). Notation is as above. By definition of share-absorb processes, $Pr((S_i^+)|((e_s)_{s\in S_i})) = Pr((S_i^+)|(S_i))$. I_s and J_s are instances of the random variables $\mathbf{I}_s$ and $\mathbf{J}_s$ defined in the proof of (10). Since the events $(\mathbf{S}_i^+) = (S_i^+)$ and $(\mathbf{I}_s) = (I_s)$ are equivalent, and the events $(\mathbf{S}_i) = (S_i)$ and $(\mathbf{J}_{s'}) = (J_{s'})$ are also equivalent,

$$Pr((S_i^+)|(S_i)) = Pr((I_s)|(J_{s'})) = \prod_{s\in S} Pr(I_s|(J_{s'})) = \prod_{s\in S} \underbrace{Pr(I_s|J_s)}_{\gamma_s},$$

where the second and third equalities hold by construction of share-absorb processes.

Now fix a source $s \in S$. It remains to prove that $Pr(I_s|J_s)$ $(= \gamma_s)$ is given by (12). We do this under the assumption that each $p_{s,i\rightarrow}$ is strictly between 0 and 1. (The generalisation to extreme parameters follows by continuity.)

If $J_s = I_s = \varnothing$, then $Pr(I_s|J_s) = 1$, because if no one initially owns s, then certainly no one shares or absorbs s.

If $J_s = \varnothing$ and $I_s \neq \varnothing$, then $Pr(I_s|J_s) = 0$, because a source that no one owns is never shared, hence never acquired.

If $J_s \not\subseteq I_s$, i.e., if J_s is not a subset of I_s, then $Pr(I_s|J_s) = 0$, because during deliberation no one loses any initially held sources.

Now assume the remaining case that $\varnothing \neq J_s \subseteq I_s$. As in the proof of (10), denote by $!_s$ the event that at least someone shares s. Given (S_i), each of $!_s$ and $\overline{!_s}$ has non-zero probability (because the parameters $p_{s,i\rightarrow}$ are neither 0 nor 1, and in case of $!_s$ also because $J_s \neq \varnothing$). So we can conditionalise on $!_s$ and on $\overline{!_s}$, and write

$$Pr(I_s|J_s) = Pr(!_s|J_s)Pr(I_s|!_s, J_s) + Pr(\overline{!_s}|J_s)Pr(I_s|\overline{!_s}, J_s).$$

Hence, as $Pr(I_s|\overline{!_s}, J_s)$ is 0 if $J_s \neq I_s$ and 1 if $J_s = I_s$,

$$Pr(I_s|J_s) = \begin{cases} Pr(!_s|J_s)Pr(I_s|!_s, J_s) & \text{if } \varnothing \neq J_s \subsetneq I_s \\ Pr(!_s|J_s)Pr(I_s|!_s, J_s) + Pr(\overline{!_s}|J_s) & \text{if } \varnothing \neq J_s = I_s \end{cases}$$

In this,

$$\begin{aligned} Pr(\overline{!_s}|J_s) &= \prod_{i\in J_s} \overline{p_{s,i\rightarrow}} \\ Pr(!_s|J_s) &= \overline{\prod_{i\in J_s} \overline{p_{s,i\rightarrow}}} \\ Pr(I_s|!_s, J_s) &= \left(\prod_{i\in I_s\setminus J_s} p_{s,i\leftarrow}\right)\left(\prod_{i\in \overline{I_s}} \overline{p_{s,i\leftarrow}}\right). \end{aligned}$$

Here, $Pr(I_s|!_s, J_s)$ reduces to $\prod_{i\in\overline{J_s}} \overline{p_{s,i\leftarrow}}$ if $J_s = I_s$. In sum, we have shown that

$$Pr(I_s|J_s) = \begin{cases} 1 & \text{if } J_s = I_s = \varnothing \\ 0 & \text{if } \varnothing = J_s \subsetneq I_s \\ 0 & \text{if } J_s \not\subseteq I_s \\ \left(\overline{\prod_{i\in J_s} \overline{p_{s,i\rightarrow}}}\right)\left(\prod_{i\in I_s\setminus J_s} p_{s,i\leftarrow}\right)\left(\prod_{i\in\overline{I_s}} \overline{p_{s,i\leftarrow}}\right) & \text{if } \varnothing \neq J_s \subsetneq I_s \\ \left(\overline{\prod_{i\in J_s} \overline{p_{s,i\rightarrow}}}\right)\left(\prod_{i\in\overline{J_s}} \overline{p_{s,i\leftarrow}}\right) + \prod_{i\in J_s} \overline{p_{s,i\rightarrow}} & \text{if } \varnothing \neq J_s = I_s \end{cases}$$

Of these six cases, the first can be subsumed under the last, as the formula in the last reduces to 1 if $J_s = I_s = \varnothing$; and the second can be subsumed under the fourth, as the formula in the fourth reduces to 0 if $\varnothing = J_s$. This yields formula (12). ■

C The Jury Theorems: Proofs

Proof of equation (2). Under the given assumptions, $\mathbf{o}_{ideal} = \mathbf{x}$ holds if and only if total evidence $\sum_{s\in S} \mathbf{e}_s$ has the same sign as $\mathbf{x}$. The probability of this event equals the conditional probability that $\sum_{s\in S} \mathbf{e}_s > 0$ given $\mathbf{x} = 1$, by Simple Gaussian Evidence. Given $\mathbf{x} = 1$, $\sum_{s\in S} \mathbf{e}_s$ is the sum of $|S|$ independent Gaussian variables of mean 1 and variance σ^2, hence is itself a Gaussian variable, with mean $|S|$ and variance $|S|\,\sigma^2$. The probability that such a variable is positive equals the probability that a standard-Gaussian variable is below $\frac{\sqrt{|S|}}{\sigma}$, by a simple linear transformation. ■

Proof of the Pre-Deliberation Jury Theorem. Assume a simple opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ for an infinite population $N = \{1, 2, ...\}$. Notation is as usual.

(a) To prove the non-asymptotic claim, we fix a group size n and write $\mathbf{o}_{maj}$ for $\mathbf{o}_{maj,n}$. We first show that $p_{maj} \leq p_{ideal}$, i.e., that $Pr(\mathbf{o}_{maj} = \mathbf{x}) \leq Pr(\mathbf{o}_{ideal} = \mathbf{x})$. We begin by proving a general claim:

Claim: For every discrete random variable $\mathbf{z}$ that is independent of the state-evidence combination $(\mathbf{x}, (\mathbf{e}_s))$ (e.g., for $\mathbf{z} = (\mathbf{S}_i)$),

$$Pr(\mathbf{o}_{ideal} = \mathbf{x} | (\mathbf{e}_s), \mathbf{z}) > \frac{1}{2}$$

except in a zero-probability event (i.e., except if the combination $((\mathbf{e}_s), \mathbf{z})$ falls into a set into which it falls with zero probability).

To show the claim, note first that such a variable $\mathbf{z}$ is independent of the event $\mathbf{o}_{ideal} = \mathbf{x}$ conditional on $(\mathbf{e}_s)$, because $\mathbf{o}_{ideal}$ is a function of $(\mathbf{e}_s)$. So, $Pr(\mathbf{o}_{ideal} = \mathbf{x} | (\mathbf{e}_s), \mathbf{z})$ can be replaced by $Pr(\mathbf{o}_{ideal} = \mathbf{x} | (\mathbf{e}_s))$, which, by construction of the ideal opinion $\mathbf{o}_{ideal}$, indeed exceeds $\frac{1}{2}$, except in the zero-probability event that $\sum_s \mathbf{e}_s = 0$ (i.e., except if $\mathbf{o}_{ideal}$ is zero, hence certainly distinct from $\mathbf{x}$). Q.e.d.

Now choose $\mathbf{z} = (\mathbf{S}_i)$. Then

$$Pr(\mathbf{o}_{maj} = \mathbf{x} | (\mathbf{e}_s), \mathbf{z}) = \begin{cases} Pr(\mathbf{o}_{ideal} = \mathbf{x} | (\mathbf{e}_s), \mathbf{z}) & \text{if } \mathbf{o}_{maj} = \mathbf{o}_{ideal} \\ 1 - Pr(\mathbf{o}_{ideal} = \mathbf{x} | (\mathbf{e}_s), \mathbf{z}) & \text{if } \mathbf{o}_{maj} = -\mathbf{o}_{ideal} \\ 0 & \text{if } \mathbf{o}_{maj} = 0. \end{cases} \quad (14)$$

Here, 'if $\mathbf{o}_{maj} = \mathbf{o}_{ideal}$' of course means 'if $((\mathbf{e}_s), \mathbf{z})$ takes a value such that $\mathbf{o}_{maj} = \mathbf{o}_{ideal}$', which is indeed a well-defined condition because the value of $((\mathbf{e}_s), \mathbf{z})$ determines the values of $\mathbf{o}_{maj}$ and $\mathbf{o}_{ideal}$, hence determines whether $\mathbf{o}_{maj} = \mathbf{o}_{ideal}$. The meanings of 'if $\mathbf{o}_{maj} = -\mathbf{o}_{ideal}$' and 'if $\mathbf{o}_{maj} = 0$' are analogous.

The 'Claim' and (14) jointly imply that, still for $\mathbf{z} = (\mathbf{S}_i)$,

$$Pr(\mathbf{o}_{maj} = \mathbf{x} | (\mathbf{e}_s), \mathbf{z}) \leq Pr(\mathbf{o}_{ideal} = \mathbf{x} | (\mathbf{e}_s), \mathbf{z}) \quad (15)$$

with probability one. By taking expectations on both sides (thereby averaging out $(\mathbf{e}_s)$ and $\mathbf{z}$), we obtain $Pr(\mathbf{o}_{maj} = \mathbf{x}) \leq Pr(\mathbf{o}_{ideal} = \mathbf{x})$, i.e., $p_{maj} \leq p_{ideal}$.

Finally, assume Imperfect Access. Then with non-zero probability the variable $\mathbf{z} = (\mathbf{S}_i)$ takes a value such that some source is not accessed by anyone, hence not accessed by a majority. This easily implies that with non-zero probability the second or third case in (14) applies. So, in (15) the '$\leq$' is a '$<$' with non-zero probability. Hence, taking the expectation on both sides of (15) now yields $Pr(\mathbf{o}_{maj} = \mathbf{x}) < Pr(\mathbf{o}_{ideal} = \mathbf{x})$, i.e., $p_{maj} < p_{ideal}$.

(b) We now show the convergence claim, assuming Access Competence. By this assumption, there is an $\epsilon > 0$ such that $p_{s \to i} \geq 2^{-1/|S|} + \epsilon$ for all s and i. Consider a person i. The probability of having full source set S satisfies $Pr(\mathbf{S}_i = S) \geq \frac{1}{2} + \epsilon^{|S|}$, because

$$Pr(\mathbf{S}_i = S) = \prod_{s \in S} p_{s \to i} \geq \prod_{s \in S} (2^{-1/|S|} + \epsilon) = \left(2^{-1/|S|} + \epsilon\right)^{|S|} \geq (2^{-1/|S|})^{|S|} + \epsilon^{|S|} = \frac{1}{2} + \epsilon^{|S|}.$$

Since the full-access events '$\mathbf{S}_i = S$' ($i = 1, 2, ...$) are mutually independent (by Independent Sources) and each of probability at least $\frac{1}{2} + \epsilon^{|S|}$, the probability that the proportion of members with full access exceeds $\frac{1}{2}$ (the event $\frac{\#\{i\in\{1,...,n\}:\mathbf{S}_i=S\}}{n} > \frac{1}{2}$) tends to one as $n \to \infty$, by the law of large numbers. In other words, the probability of a majority with full access (the event $\#\{i \in \{1,...,n\} : \mathbf{S}_i = S\} > \frac{n}{2}$) tends to 1 as $n \to \infty$. Meanwhile, full access implies an ideal opinion (i.e., $\mathbf{S}_i = S$ implies $\mathbf{o}_i = \mathbf{o}_{ideal}$). So a majority with full access implies a majority with the ideal opinion (i.e., $\#\{i \in \{1,...,n\} : \mathbf{S}_i = S\} > \frac{n}{2}$ implies $\mathbf{o}_{maj,n} = \mathbf{o}_{ideal}$). Hence, also the probability of an ideal majority opinion converges to one: $Pr(\mathbf{o}_{maj,n} = \mathbf{o}_{ideal}) \to 1$. This implies that $Pr(\mathbf{o}_{maj,n} = \mathbf{x}) \to Pr(\mathbf{o}_{ideal} = \mathbf{x})$, i.e., that $p_{maj,n} \to p_{ideal}$. ■

Proof of the Post-Deliberation Jury Theorem. Assume a simple opinion structure $(\mathbf{x}, (\mathbf{e}_s), (\mathbf{S}_i))$ and a share-absorb process, both for an infinite population $N = \{1, 2, ...\}$. The usual notation applies.

(a) The non-asymptotic claim holds by a version of the proof of part (a) of the Pre-Deliberation Jury Theorem. One should substitute $\mathbf{o}^+_{maj}$ for $\mathbf{o}_{maj}$, and apply the 'Claim' with $\mathbf{z} = (\mathbf{S}^+_i)$ rather than $\mathbf{z} = (\mathbf{S}_i)$, which is possible since also $(\mathbf{S}^+_i)$ is independent of $(\mathbf{x}, (\mathbf{e}_s))$.

(b) We now turn to the asymptotic claim. We shall face the difficulty of interpersonal correlations between post-deliberation source sets. The weak law of large numbers in Pivato's (2017) version for correlated variables will ultimately come to help, but first several claims must be established. We assume Acquisition Competence (needed only from Claim b5) and Non-Vanishing Participation (needed only from Claim b4).

Claim b1: For any source $s \in S$, group size $n \in \{1, 2, ...\}$, and group member $i \in \{1, ..., n\}$, the probability that another member shares s is

$$p_{s,i,n} = \overline{\prod_{j\in\{1,...,n\}\backslash\{i\}} \overline{p_{s\to j}p_{s,j\to}}}. \tag{16}$$

The probability is given by (16) because it equals the probability that it is *not* the case that each other member j does *not* share s, where j shares s with probability $p_{s\to j}p_{s,j\to}$, the product of the probabilities of accessing s and of sharing an accessed s. Q.e.d.

Claim b2: For any $s \in S$, $n \in \{1, 2, ...\}$, and $i \in \{1, ..., n\}$, the probability that some member other than i shares s and then i absorbs s, given that i has not accessed s initially, is $p_{s,i,n}p_{s,i\leftarrow}$.

The claim holds because the relevant probability is the product of the probability that someone else shares s, i.e., $p_{s,i,n}$ by Claim b1, and the probability that i absorbs a shared source s, i.e., $p_{s,i\leftarrow}$. Q.e.d.

Claim b3: For any $s \in S$, $n \in \{1, 2, ...\}$, and $i \in \{1, ..., n\}$, $Pr(s \in \mathbf{S}^+_{i,n}) = \overline{\overline{p_{s\to i}} \times \overline{p_{s,i,n}p_{s,i\leftarrow}}}$.

This holds because i does *not* hold s post-deliberation if and only if i does not initially access s (probability: $\overline{p_{s\to i}}$) and i does not absorb s (probability: $\overline{p_{s,i,n}p_{s,i\leftarrow}}$). Q.e.d.

Claim b4: For any $s \in S$ and $i \in \{1, 2, ...\}$, $P(s \in \mathbf{S}^+_{i,n}) \to \overline{\overline{p_{s\to i}} \times \overline{p_{s,i\leftarrow}}}$ as $n \to \infty$.

Fix s and i. By Claim b3, we just show $p_{s,i,n} \to 1$, i.e., $\prod_{j\in\{1,...,n\}\backslash\{i\}} \overline{p_{s\to j}p_{s,j\to}} \to 0$. By Non-Vanishing Participation, $p_{s\to j}p_{s,j\to} \not\to 0$ as $j \to \infty$, and hence $\overline{p_{s\to j}p_{s,j\to}} \not\to 1$ as $j \to \infty$. In consequence, $\prod_{j\in\{1,...,n\}\backslash\{i\}} \overline{p_{s\to j}p_{s,j\to}} \to 0$ as $n \to \infty$. Q.e.d.

Claim b5: For any $i \in \{1, 2, ...\}$, the full-access probability $Pr(\mathbf{S}^+_{i,n} = S)$ converges to a value of at least $\frac{1}{2} + \epsilon^{|S|}$ as $n \to \infty$, where $\epsilon > 0$ is the threshold in Acquisition Competence (which is independent of i).

Fix a person i. We have $Pr(\mathbf{S}^+_{i,n} = S) = \prod_{s\in S} Pr(s \in \mathbf{S}^+_{i,n})$, because the access events '$s \in \mathbf{S}^+_{i,n}$' are independent across sources s as a consequence of the fact that the pre-deliberation access events '$s \in \mathbf{S}_i$' are independent across sources (by Source Independence) and the share-absorb process operates independently across sources. So, by Claim b4, $Pr(\mathbf{S}_{i,n}^{+} = S) \to \prod_{s\in S} \overline{\overline{p_{s\to i}} \times \overline{p_{s,i\leftarrow}}}$ as $n \to \infty$. Now choose $\epsilon > 0$ as in Acquisition Competence. Then, for all s, $\overline{p_{s\to i}} \times \overline{p_{s,i\leftarrow}} \leq 1 - 2^{-1/|S|} - \epsilon$, i.e., $\overline{\overline{p_{s\to i}} \times \overline{p_{s,i\leftarrow}}} \geq 2^{-1/|S|} + \epsilon$. So,

$$\prod_{s\in S} \overline{\overline{p_{s\to i}} \times \overline{p_{s,i\leftarrow}}} \geq \left(2^{-1/|S|} + \epsilon\right)^{|S|} \geq (2^{-1/|S|})^{|S|} + \epsilon^{|S|} = \frac{1}{2} + \epsilon^{|S|}.$$

Hence $\lim_{n\to\infty} Pr(\mathbf{S}_{i,n}^{+} = S) \geq \frac{1}{2} + \epsilon^{|S|}$. Q.e.d.

Claim b6: For all $n \in \{1, 2, ...\}$ and distinct $i, j \in \{1, ..., n\}$, the covariance between i's and j's full access satisfies

$$Cov(\mathbf{S}_{i,n}^{+} = S, \mathbf{S}_{j,n}^{+} = S) \leq \prod_{s\in S} \prod_{k=i,j} \overline{\overline{p_{s\to k}} \times \overline{p_{s,k\leftarrow}}} - \prod_{s\in S} \prod_{k=i,j} \overline{\overline{p_{s\to k}} \times \overline{p_{s,k,n} p_{s,k\leftarrow}}}$$

Fix $n \in \{1, 2, ...\}$ and distinct $i, j \in \{1, ..., n\}$. Then

$$\begin{aligned} Cov(\mathbf{S}_{i,n}^{+} = S, \mathbf{S}_{j,n}^{+} = S) &= Pr(\mathbf{S}_{i,n}^{+} = S, \mathbf{S}_{j,n}^{+} = S) - \prod_{k=i,j} Pr(\mathbf{S}_{k,n}^{+} = S) \\ &= \prod_{s\in S} Pr(s \in \mathbf{S}_{i,n}^{+}, s \in \mathbf{S}_{j,n}^{+}) - \prod_{s\in S} \prod_{k=i,j} Pr(s \in \mathbf{S}_{k,n}^{+}), \end{aligned}$$

where the second equality holds by independence across sources of the access events. Since $Pr(s \in \mathbf{S}_{k,n}^{+}) = \overline{\overline{p_{s\to k}} \times \overline{p_{s,k,n} p_{s,k\leftarrow}}}$ by Claim b3, it suffices to show that

$$Pr(s \in \mathbf{S}_{i,n}^{+}, s \in \mathbf{S}_{j,n}^{+}) \leq \prod_{k=i,j} \overline{\overline{p_{s\to k}} \times \overline{p_{s,k\leftarrow}}} \text{ for all } s \in S.$$

This holds since, letting E be the event that s is shared by someone in $\{1, ..., n\}$,

$$\begin{aligned} Pr(s \in \mathbf{S}_{i,n}^{+}, s \in \mathbf{S}_{j,n}^{+}) &\leq Pr(s \in \mathbf{S}_{i,n}^{+}, s \in \mathbf{S}_{j,n}^{+} | E) \\ &= \prod_{k=i,j} Pr(s \in \mathbf{S}_{k,n}^{+} | E) = \prod_{k=i,j} \overline{\overline{p_{s\to k}} \times \overline{p_{s,k\leftarrow}}}, \end{aligned}$$

where the first equality holds by independence between '$s \in \mathbf{S}_{i,n}^{+}$' and '$s \in \mathbf{S}_{i,n}^{+}$' given E, and the second equality holds because s is held ex-post if and only if it is *not* the case that s is *not* accessed ex-ante (probability: $\overline{p_{s\to k}}$) and *not* absorbed ex-post (probability: $\overline{p_{s,i\leftarrow}}$). Q.e.d.

Claim b7: : $\min_{s\in S, k\leq n} p_{s,k,n} \to 1$ as $n \to \infty$.

For all s and n, pick $i_{s,n} \in \{1, ..., n\}$ with $p_{s\to i_{s,n}} p_{s,i_{s,n}\to} = \max_{k\leq n} p_{s\to k} p_{s,k\to}$. By construction,

$$\min_{k\leq n} p_{s,k,n} = \overline{\prod_{j\in\{1,...,n\}\setminus\{i_{s,k}\}} \overline{p_{s\to j} p_{s,j\to}}}.$$

By Non-Vanishing Participation, $p_{s\to j} p_{s,j\to} \not\to 0$. So $\prod_{j\in\{1,...,n\}\setminus\{i_{s,k}\}} \overline{p_{s\to j} p_{s,j\to}} \to 0$. Hence, $\min_{k\leq n} p_{s,k,n} \to 1$. So, as $|S|$ is finite, $\min_{s\in S, k\leq n} p_{s,k,n} \to 1$. Q.e.d.

Claim b8: $\delta_n \equiv \max_{s\in S, k\leq n} \left(\overline{\overline{p_{s\to k}} \times \overline{p_{s,k\leftarrow}}} - \overline{\overline{p_{s\to k}} \times \overline{p_{s,k,n} p_{s,k\leftarrow}}}\right) \to 0$ as $n \to \infty$.

For all $s \in S$, $n \in \{1, 2, ...\}$ and $k \leq n$, we have

$$\begin{aligned}
\overline{\overline{p_{s\to k}} \times \overline{p_{s,k\leftarrow}}} - \overline{\overline{p_{s\to k}} \times \overline{p_{s,k,n}p_{s,k\leftarrow}}} &= \overline{p_{s\to k}} \times \overline{p_{s,k,n}p_{s,k\leftarrow}} - \overline{p_{s\to k}} \times \overline{p_{s,k\leftarrow}} \\
&= \overline{p_{s\to k}} \left(\overline{p_{s,k,n}p_{s,k\leftarrow}} - \overline{p_{s,k\leftarrow}}\right) \\
&= \overline{p_{s\to k}} \left(p_{s,k\leftarrow} - p_{s,k,n}p_{s,k\leftarrow}\right) \\
&= \overline{p_{s\to k}} p_{s,k\leftarrow} \left(1 - p_{s,k,n}\right) \\
&\geq 1 - p_{s,k,n}.
\end{aligned}$$

This lower bound implies the desired convergence via Claim b7. Q.e.d.

Claim b9: The average covariance of full access between group members converges to zero, i.e.,

$$\frac{1}{n^2} \sum_{i,j=1}^{n} Cov(\mathbf{S}^+_{i,n} = S, \mathbf{S}^+_{j,n} = S) \to 0 \text{ as } n \to \infty.$$

It suffices to show that $Cov(\mathbf{S}^+_{i,n} = S, \mathbf{S}^+_{j,n} = S) \leq 1$ whenever $i = j$ $(\leq n)$ and that $\max_{i,j \leq n, i \neq i} Cov(\mathbf{S}^+_{i,n} = S, \mathbf{S}^+_{j,n} = S) \to 0$, by a simple argument (which uses that each $Cov(\mathbf{S}^+_{i,n} = S, \mathbf{S}^+_{j,n} = S)$ is positive). The former is obvious. We now show the latter. By Claim b6 and the positivity of all the covariances, it is enough to prove that, for any distinct i, j,

$$\prod_{s\in S} \prod_{k=i,j} a_{s,k} - \prod_{s\in S} \prod_{k=i,j} a_{s,k,n} \to 0 \text{ as } n \to \infty,$$

where $a_{s,k} = \overline{\overline{p_{s\to k}} \times \overline{p_{s,k\leftarrow}}}$ and $a_{s,k,n} = \overline{\overline{p_{s\to k}} \times \overline{p_{s,k,n}p_{s,k\leftarrow}}}$. Fix distinct i, j. Note that $a_{k,s} = (a_{s,k} - a_{s,k,n}) + a_{s,k,n} \leq \delta_n + a_{s,k,n}$, by Claim b8. So it suffices to show that

$$\prod_{s\in S} \prod_{k=i,j} (\delta_n + a_{s,k,n}) - \prod_{s\in S} \prod_{k=i,j} a_{s,k,n} \to 0 \text{ as } n \to \infty. \tag{17}$$

By developing the product $\prod_{s\in S} \prod_{k=i,j}(\delta_n + a_{s,k,n})$, check that it equals a polynomial in δ_n (of order $2|S|$) whose constant term is '$+\prod_{s\in S} \prod_{k=i,j} a_{s,k,n}$'. This constant term cancels out against '$-\prod_{s\in S} \prod_{k=i,j} a_{s,k,n}$', so that the expression in (17) is a polynomial in δ_n with *zero* constant term. As $n \to \infty$, δ_n converges to 0 (by Claim b8), and so any polynomial in δ_n with zero constant term also converges to 0. This proves (17). Q.e.d.

Claim b10: $p^+_{maj,n} \to p_{ideal}$. (This completes the proof.)

Since every person i's full-access event $\mathbf{S}^+_{i,n} = S$ has probability converging to $\frac{1}{2} + \epsilon^{|S|}$ as $n \to \infty$ by Claim b5, while the average covariance of these events tends to zero by Claim b9, the probability that the proportion of members with full access exceeds $\frac{1}{2}$ (the event $\frac{\#\{i \in \{1,...,n\}: \mathbf{S}^+_{i,n} = S\}}{n} > \frac{1}{2}$) tends to one, by the weak law of large numbers in Pivato's (2017) version for correlated variables.[33] Equivalently, the probability of a majority with full access (the event $\#\{i \in \{1, ..., n\} : \mathbf{S}^+_{i,n} = S\} > \frac{n}{2}$) tends to 1. Since (a majority with) full access implies (a majority with) an ideal opinion, also the probability of an ideal majority opinion converges to 1: $Pr(\mathbf{o}^+_{maj,n} = \mathbf{o}_{ideal}) \to 1$. So, $Pr(\mathbf{o}^+_{maj,n} = \mathbf{x}) \to Pr(\mathbf{o}_{ideal} = \mathbf{x})$, i.e., $p^+_{maj,n} \to p_{ideal}$. ■

[33] This version of the law follows from the proof of Pivato's Theorem 5.2 (i.e., from Claim 2 in that proof, combined with Chebyshev's Inequality). A closely related result is Proposition A2 in Pivato (2016).

D Proof of Propositions 2–4

Notation for the proofs of Propositions 2–4: The values of $\mathbf{o}_{maj}$, $\mathbf{SI}$, $\mathbf{II}$, $\mathbf{N}_s$ ($s \in S$) and $\mathbf{E}_i$ ($i \in N$) at an evidence profile $((e_s)_{s\in S_i})$ are denoted o_{maj}, SI, II, N_s resp. E_i; and their value for an evidence profile denoted using prime(s) (such as $((e'_s)_{s\in S'_i})$ or $((e'_s)_{s\in S_i})$ or $((e_s)_{s\in S'_i})$) are denoted o'_{maj}, SI', II', N'_s resp. E'_i.

Proof of Proposition 2. Fix a simple opinion structure $(\mathbf{x}, (\mathbf{e}_s)_{s\in S}, (\mathbf{S}_i)_{i\in N})$ such that $|S|, |N| \geq 2$, and $p_{s\to i} \neq 0$ for all $s \in S$ and $i \in N$.

(a) Pick distinct sources $s_1, s_2 \in S$. Let $((e_s)_{s\in S_i})$ be an evidence profile with source sets $S_1 = \{s_1, s_2\}$ and $S_2 = \cdots = S_n = \{s_2\}$ and with evidences $e_{s_1} = -2$ and $e_{s_2} = 1$. At $((e_s)_{s\in S_i})$, all persons i have the same evidence strength of $|E_i| = 1$, so that $II = 0$, whereas $SI \neq 0$ as $\#N_{s_1} = 1$ while $\#N_{s_2} = n \neq 1$. Finally, the majority opinion o_{maj} is inefficient, because the efficient opinion is -1 (as $\sum_{s\in\cup_i S_i} e_s = e_{s_1} + e_{s_2} = -1 < 0$ so that $Pr(\mathbf{x} = 1|(e_s)_{s\in\cup_i S_i}) < \frac{1}{2}$) but the majority opinion satisfies $o_{maj} \geq 0$ as only person 1 votes for -1 while the other $n - 1$ (≥ 1) persons vote for 1.

(b) Let $m = |S|$, say $S = \{s_1, ..., s_m\}$. Consider two cases.

Case 1: $m \geq n$. Let $((e_s)_{s\in S_i})$ be the evidence profile with $S_i = \{s_i\}$ for all $i = 1, ..., n-1$, $S_n = \{s_n, ..., s_m\}$, $e_{s_1} = \cdots = e_{s_{m-1}} = 1$ and $e_{s_m} = -m$. At $((e_s)_{s\in S_i})$, each source is accessed by exactly one person, so that $SI = 0$, whereas $II \neq 0$ because persons $i = 1, ..., n-1$ have evidence strength $|E_i| = 1$ while person n has $|E_n| = m-(m-n)\times 1 = n \neq 1$. The majority opinion o_{maj} is inefficient, because the efficient opinion is -1 (as $\sum_{s\in\cup_i S_i} e_s = (m-1)\times 1 - m < 0$ so that $Pr(\mathbf{x} = 1|((e_s)_{s\in S_i})) < \frac{1}{2}$) but the majority opinion satisfies $o_{maj} \geq 0$ as $n - 1$ (≥ 1) persons vote for 1 while only person n votes for -1.

Case 2: $m < n$. Let $((e_s)_{s\in S_i})$ be evidence profile with $S_i = \{s_i\}$ for all $i = 1, ..., m$, $S_{m+1} = \cdots = S_n = \varnothing$, $e_{s_1} = \cdots = e_{s_{m-1}} = 1$, and $e_{s_m} = -m$. At $((e_s)_{s\in S_i})$, each source is accessed by exactly one person, so that $SI = 0$, whereas $II \neq 0$ because persons $i = 1, ..., m-1$ have evidence strength $|E_i| = 1$ while person m has $|E_m| = m > 1$. The majority opinion o_{maj} is again inefficient, since it satisfies $o_{maj} \geq 0$ while the efficient opinion is -1, as can be checked (the argument uses that the voters $i = m+1, ..., n$ vote 0, not affecting the majority outcome). ■

Proof of Proposition 3 in its strengthened form with $\Delta = \frac{6|S|-1}{5|S|-5}$ *and* $\Delta' = \frac{4}{n+2}$. Let the assumptions hold.

(a) Consider an evidence profile $((e_s)_{s\in S_i})$ at which $\mathbf{SI} > \frac{6|S|-1}{5|S|-5}$. Write the spread of a source $s \in S$ as $x_s = \#N_s$, and let $\tilde{S} = \{s : x_s \neq 0\}$ and $j = \#\tilde{S}$. We first establish the following fact:

Claim: There are sources $s_+, s_- \in \tilde{S}$ such that $x_{s_+} > 2x_{s_-}$.

Assume for a contradiction that $x_s \leq 2x_t$ for all $s, t \in \tilde{S}$. We write

$$SI = \frac{1}{|S|(|S|-1)} \sum_{(s,t)\in S^2:s\neq t} \frac{|x_s - x_t|}{\frac{1}{2}(x_s + x_t)} = \frac{1}{|S|(|S|-1)} (\Sigma_0 + \Sigma_1 + 2\Sigma_2)$$

where we define

$$\Sigma_0 = \sum_{(s,t)\in(S\backslash\tilde{S})^2:s\neq t} \frac{|x_s - x_t|}{\frac{1}{2}(x_s + x_t)}, \Sigma_1 = \sum_{(s,t)\in\tilde{S}^2:s\neq t} \frac{|x_s - x_t|}{\frac{1}{2}(x_s + x_t)}, \Sigma_2 = \sum_{(s,t)\in\tilde{S}\times(S\backslash\tilde{S})} \frac{|x_s - x_t|}{\frac{1}{2}(x_s + x_t)}.$$

In Σ_0, each term $\frac{|x_s-x_t|}{\frac{1}{2}(x_s+x_t)}$ is 0; so $\Sigma_0=0$. In Σ_1, each term satisfies $\frac{|x_s-x_t|}{\frac{1}{2}(x_s+x_t)}\leq\frac{2}{3}$, as can be shown using that $x_s\leq 2x_t$ and $x_t\leq 2x_s$; so $\Sigma_1\leq\frac{2}{3}\#\{(s,t)\in\tilde{S}^2: s\neq t\}=\frac{2}{3}j(j-1)$. Finally, in Σ_2 each term satisfies $\frac{|x_s-x_t|}{\frac{1}{2}(x_s+x_t)}=2$, since $x_s>0$ and $x_t=0$; so $\Sigma_2=2|\tilde{S}\times(S\backslash\tilde{S})|=2|\tilde{S}||S\backslash\tilde{S}|=2j(|S|-j)$. It follows that

$$SI\leq\frac{1}{|S|\,(|S|-1)}\left[0+\frac{2}{3}j(j-1)+4j(|S|-j)\right]=\frac{1}{|S|\,(|S|-1)}\left[-\frac{10}{3}j^2+(4\,|S|-\frac{2}{3})j\right].$$

By basic algebra, the second-order polynomial '$-\frac{10}{3}j^2+(4\,|S|-\frac{2}{3})j$' in j (regarded as a real number) takes a maximum at $j^*=\frac{6|S|-1}{10}$, where its value is $\frac{6}{5}\,(|S|-1/6)^2$. Since $j^*\notin\mathbb{N}$ but the actual j belongs to $\mathbb{N}$, the maximum is not reached. Thus, $-\frac{10}{3}j^2+(4\,|S|-\frac{2}{3})j<\frac{6}{5}\,(|S|-1/6)^2$, and so

$$SI<\frac{6}{5}\times\frac{(|S|-1/6)^2}{|S|\,(|S|-1)}\leq\frac{6}{5}\times\frac{|S|-1/6}{|S|-1}=\frac{6\,|S|-1}{5\,|S|-5}.$$

This contradicts the assumption that $SI\geq\frac{6|S|-1}{5|S|-5}$. Q.e.d.

Now fix $s_+,s_-\in\tilde{S}$ as in the 'Claim'. We construct another evidence profile $((e'_s)_{s\in S_i})$ with same source sets S_i such that monotonicity in evidence is violated for $((e_s)_{s\in S_i})$ and $((e'_s)_{s\in S_i})$.

Case 1: $o_{maj}\geq 0$. Choose $e'_{s_+}=-1$, $e'_{s_-}>\max\left\{\sum_{s\in\cup_i S_i}e_s,0\right\}+1$, and $e'_s=0$ for all $s\in(\cup_i S_i)\backslash\{s_+,s_-\}$. Then $o'_{maj}=-1$, because at $((e'_s)_{s\in S_i})$ option 1 receives only $\#N_{s_-}=x_{s_-}$ votes while option -1 receives $\#(N_{s_+}\backslash N_{s_-})\geq x_{s_+}-x_{s_-}>2x_{s_-}-x_{s_-}=x_{s_-}$ votes. So $o_{maj}>o'_{maj}$. This implies a violation of (3) for the pair of profiles $((e_s)_{s\in S_i})$ and $((e'_s)_{s\in S_i})$, because $Pr(\mathbf{x}=1|(e_s)_{s\in\cup S_i})<Pr(\mathbf{x}=1|(e'_s)_{s\in\cup S_i})$ as $\sum_{s\in\cup_i S_i}e'_s>-1+\max\left\{\sum_{s\in\cup_i S_i}e_s,0\right\}+1\geq\sum_{s\in\cup_i S_i}e_s$.

Case 2: $o_{maj}\leq 0$. Choose $e'_{s_+}=1$, $e'_{s_-}<\min\left\{\sum_{s\in\cup_i S_i}e_s,0\right\}-1$, and $e'_s=0$ for all $s\in(\cup_i S_i)\backslash\{s_+,s_-\}$. Arguments analogous to those under Case 1 show that $o'_{maj}=1$ $(>o_{maj})$ and that (3) is violated for the pair of profiles $((e'_s)_{s\in S_i})$ and $((e_s)_{s\in S_i})$ (in this order).

(b) Let $((e_s)_{s\in S_i})$ be an evidence profile with a minority source $t\in S$. We construct an evidence profile $((e'_s)_{s\in S_i})$ with same source sets S_i such that monotonicity in evidence is violated for $((e_s)_{s\in S_i})$ and $((e'_s)_{s\in S_i})$ and $II'>\frac{4}{n+2}$.

Case 1: $o_{maj}\geq 0$. Choose $e'_s=-1$ for all $s\in(\cup_i S_i)\backslash\{t\}$ and $e'_t=(n+2)\,|S|+\max\left\{\sum_{s\in\cup S_i}e_s,0\right\}$. Then $o'_{maj}=-1$ $(<o_{maj})$, because at $((e'_s)_{s\in S_i})$ the option 1 receives $\#\{i:t\in S_i\}$ votes and the option -1 receives $\#\{i:t\notin S_i\neq\varnothing\}$ votes, where $\#\{i:t\in S_i\}<\#\{i:t\notin S_i\neq\varnothing\}$ as t is a minority source. It follows that (3) is violated for the pair of profiles $((e_s)_{s\in S_i})$ and $((e'_s)_{s\in S_i})$, because $Pr(\mathbf{x}=1|(e_s)_{s\in\cup S_i})<Pr(\mathbf{x}=1|(e'_s)_{s\in\cup S_i})$ as

$$\sum_{s\in\cup_i S_i}e'_s=(-1)\underbrace{|\cup_i S_i)\backslash\{t\}|}_{\leq|S|}+(n+2)\,|S|+\max\left\{\sum_{s\in\cup S_i}e_s,0\right\}>\sum_{s\in\cup S_i}e_s.$$

Write a person i's evidence strength at $((e'_s)_{s\in S_i})$ as $y_i=|E'_i|$. For all $j\in N\backslash N_t$ we have $y_j=|(-1)\,|S_j||=|S_j|<|S|$, while for all $j\in N_t$, writing z for $\max\left\{\sum_{s\in\cup S_i}e_s,0\right\}$, we have

$$y_j=(-1)\,|S_j\backslash\{t\}|+(n+2)\,|S|+z>z+(n+1)\,|S|\,.$$

Now for all $(i,j) \in N_t \times (N \backslash N_t)$, since $y_i > z + (n+1)|S| > |S| > y_j \geq 0$, we have

$$\frac{|y_i - y_j|}{\frac{1}{2}(y_i + y_j)} = 2\frac{y_i - y_j}{y_i + y_j} > 2\frac{z + (n+1)|S| - |S|}{z + (n+1)|S| + |S|)} > 2\frac{n|S|}{(n+2)|S|} = \frac{2n}{n+2}.$$

In $II' = \frac{1}{n(n-1)} \sum_{(i,j)\in N^2 : i \neq j} \frac{|y_i - y_j|}{\frac{1}{2}(y_i+y_j)}$, the term $\frac{|y_i - y_j|}{\frac{1}{2}(y_i+y_j)}$ is 0 if $(i,j) \in N^2$ or $(i,j) \in (N \backslash N_t)^2$, and it exceeds $\frac{2n}{n+2}$ if $(i,j) \in N_t \times (N \backslash N_t)$ or $(i,j) \in (N \backslash N_t) \times N_t$. So,

$$II' > \frac{1}{n(n-1)} 2 \sum_{(i,j)\in N_t \times (N \backslash N_t)} \frac{2n}{n+2} = \frac{4|N_t||N \backslash N_t|}{(n-1)(n+2)} \geq \frac{4 \times 1 \times (n-1)}{(n-1)(n+2)} = \frac{4}{n+2}.$$

Case 2: $o_{maj} \leq 0$. Here choose $e'_s = 1$ for all $s \in (\cup_i S_i) \backslash \{t\}$ and $e'_t = -(n+2)|S| + \min\{\sum_{s\in\cup S_i} e_s, 0\}$. For reasons like those under Case 1, $o'_{maj} = 1$ $(> o_{maj})$, (3) is violated for the pair of profiles $((e'_s)_{s\in S_i})$ and $((e_s)_{s\in S_i})$ (in this order), and $II' > \frac{2n}{n+2}$. ■

Proof of Proposition 4 in its strengthened form with $\Delta = 4\frac{n-2}{n|S|}$ *and* $\Delta' = (3 - 2\sqrt{2})\frac{n-1}{n}$. Let the assumptions hold. For the proof of both parts, we fix an evidence profile $(e_s)_{s\in S_i})$ such that $|\{s \in \cup S_i : e_s > 0\}|, |\{s \in \cup S_i : e_s < 0\}| \geq 2$, and construct another profile $((e_s)_{s\in S'_i})$ with $e'_s = e_s$ for all $s \in (\cup_i S_i) \cap (\cup_i S'_i)$ such that monotonicity in sources is violated for this pair of profiles and $SI' \geq \frac{4(n-2)}{|S|n}$ resp. $II' > (3 - 2\sqrt{2})\frac{n-1}{n}$. The proof assumes w.l.o.g. that $o_{maj} \geq 0$ (an analogous proof works if $o_{maj} \leq 0$). To construct $((e_s)_{s\in S'_i})$, we choose a subgroup $M \subseteq N$ such that $0 < |M| < \frac{n}{2}$ (it exists as $n \geq 3$) and a source set $T \neq \varnothing$ that is a strict subset of $\{s \in \cup_i S_i : e_s < 0\}$ (it exists as $|\{s \in \cup S_i : e_s < 0\}| \geq 2$), and we define

$$S'_i = \begin{cases} \{s \in \cup_i S_i : e_s > 0\} & \text{if } i \in M \\ T & \text{if } i \in N \backslash M. \end{cases}$$

Then $o'_{maj} = -1$, as all $i \in N \backslash M$ vote for -1 and $|N \backslash M| > \frac{n}{2}$. Monotonicity in sources is violated, because $o_{maj} > o'_{maj}$ although option 1 has gained evidence support as $\sum_{s\in\cup_i S'_i} e_s = \sum_{s\in\cup_i S_i : e_s > 0} e_s + \sum_{s\in T} e_s > \sum_{s\in\cup_i S_i} e_s$. From here on, the proof of parts (a) and (b) diverge.

Remaining proof for (a). For (a), choose $M = \{m\}$ and $T = \{s \in \cup_i S_i : e_s < 0\} \backslash \{t\}$ for some $m \in N$ and some $t \in \cup_i S_i$ with $e_t < 0$. Write the spread of a source $s \in S$ as $x_s = \# N_s$. Since $x_s = n - 1$ if $s \in T$ and $x_s \leq 1$ if $s \in S \backslash T$ (note that $x_s = 0$ if $s \notin \cup_i S'_i$), the spread imbalance between any $s \in T$ and any $t \in S \backslash T$ satisfies $\frac{|x_s - x_t|}{\frac{1}{2}(x_s + x_t)} = 2\frac{x_s - x_t}{x_s + x_t} \geq 2\frac{(n-1)-1}{(n-1)+1} = 2\frac{n-2}{n}$.

Now recall that $SI' = \frac{1}{|S|(|S|-1)} \sum_{(s,t)\in S^2 : s \neq t} \frac{|x_s - x_t|}{\frac{1}{2}(x_s + x_t)}$. Here, each $\frac{|x_s - x_t|}{\frac{1}{2}(x_s+x_t)}$ is 0 if $(s,t) \in T^2$ or $(s,t) \in (S \backslash T)^2$ and is at least $2\frac{n-2}{n}$ if $(s,t) \in T \times (S \backslash T)$ or $(s,t) \in (S \backslash T) \times T$. So, $SI' \geq \frac{1}{|S|(|S|-1)} 2|T||S \backslash T| \times 2\frac{n-2}{n}$. Hence, as $|T||S \backslash T| = |T|(|S| - |T|) \geq |S| - 1$, we have $SI' \geq 4\frac{n-2}{n|S|}$. Q.e.d.

Remaining proof for (b). For (b), let M be a maximal minority $M \subseteq N$, i.e., $|M| = \frac{n-2}{2}$ if n is even and $|M| = \frac{n-1}{2}$ if n is odd. Pick distinct sources $t, t' \in \cup_i S_i$ with $e_{t'} \leq e_t < 0$ (they exist by assumption) and let

$$T = \begin{cases} \{t\} & \text{if } |e_t| \leq \frac{1}{\sqrt{2}} \sum_{s\in\cup_i S_i : e_s > 0} e_s \\ \{t, t'\} & \text{if } |e_t| > \frac{1}{\sqrt{2}} \sum_{s\in\cup_i S_i : e_s > 0} e_s. \end{cases}$$

Any person i's evidence strength is written $y_i = |E'_i| = \left|\sum_{s \in S'_i} e_s\right|$; it equals $\sum_{s \in \cup_i S_i : e_s > 0} e_s$ if $i \in M$ and equals $|e_t|$ or $|e_t + e_{t'}|$ if $i \in N \backslash M$. Now, the interpersonal imbalance between an $i \in M$ and a $j \in N \backslash M$ satisfies

$$\frac{|y_i - y_j|}{\frac{1}{2}(y_i + y_j)} \geq 2(3 - 2\sqrt{2}). \tag{18}$$

Why? If $|e_t| \leq \frac{1}{\sqrt{2}} y_i$, then $T = \{t\}$ and so $y_j = |e_t| \leq \frac{1}{\sqrt{2}} y_i$; thus $\frac{|y_i - y_j|}{\frac{1}{2}(y_i + y_j)} = 2\frac{y_i - y_j}{y_i + y_j}$, where

$$\frac{y_i - y_j}{y_i + y_j} \geq \frac{y_i - \frac{1}{\sqrt{2}} y_i}{y_i + \frac{1}{\sqrt{2}} y_i} = \frac{\sqrt{2} - 1}{\sqrt{2} + 1} = \frac{(\sqrt{2} - 1)^2}{(\sqrt{2} + 1)(\sqrt{2} - 1)} = 3 - 2\sqrt{2}. \tag{19}$$

In instead $|e_t| > \frac{1}{\sqrt{2}} y_i$, then $T = \{t, t'\}$, and so $y_j = |e_t + e_{t'}| = |e_t| + |e_{t'}| \geq 2|e_t| > \frac{2}{\sqrt{2}} y_i = \sqrt{2} y_i$, an thus $y_i < y_j / \sqrt{2}$; so, $\frac{|y_i - y_j|}{\frac{1}{2}(y_i + y_j)} = 2\frac{y_j - y_i}{y_i + y_j}$, where $\frac{y_j - y_i}{y_i + y_j} > 3 - 2\sqrt{2}$ by a calculation analogous to (19).

Recall that $II' = \frac{1}{n(n-1)} \sum_{(i,j) \in N^2 : i \neq j} \frac{|y_i - y_j|}{\frac{1}{2}(y_i + y_j)}$. Here, each $\frac{|y_i - y_j|}{\frac{1}{2}(y_i + y_j)}$ is 0 if $(i, j) \in M^2$ or $(i, j) \in (N \backslash M)^2$, and it is at least $2(3 - 2\sqrt{2})$ if $(i, j) \in M \times (N \backslash M)$ or $(i, j) \in (N \backslash M) \times M$. So, $II' \geq \frac{1}{n(n-1)} 2|M||N \backslash M| \times 2(3 - 2\sqrt{2})$. The definition of M and the fact that $n \geq 3$ imply that $|M|(n - |M|) > \frac{(n-1)^2}{4}$. Therefore,

$$II' > \frac{1}{n(n-1)} \left[2\frac{(n-1)^2}{4} \times 2(3 - 2\sqrt{2}) \right] = (3 - 2\sqrt{2}) \frac{n-1}{n}. \blacksquare$$