Jury Theorems

Franz Dietrich Kai Spiekermann

preliminary version of May 2021 (the final version will appear in *Stanford Encyclopedia of Philosophy*, Winter 2021 Edition)

Abstract

Jury theorems are mathematical theorems about the ability of collectives to make correct decisions. Several jury theorems carry the optimistic message that, in suitable circumstances, 'crowds are wise': many individuals together (using, for instance, majority voting) tend to make good decisions, outperforming fewer or just one individual. Jury theorems form the technical core of epistemic arguments for democracy, and provide probabilistic tools for reasoning about the epistemic quality of collective decisions. The popularity of jury theorems spans across various disciplines such as economics, political science, philosophy, and computer science. This entry reviews and critically assesses a variety of jury theorems. It first discusses Condorcet's initial jury theorem, and then progressively introduces jury theorems with more appropriate premises and conclusions. It explains the philosophical foundations, and relates jury theorems to diversity, deliberation, shared evidence, shared perspectives, and other phenomena. It finally connects jury theorems to their historical background and to democratic theory, social epistemology, and social choice theory.

Contents

1	Introduction			
	1.1	What are jury theorems?	3	
	1.2	A broad notion of decision	4	
2	Thr	ree jury theorems	5	
	2.1	Condorcet's Jury Theorem	6	
	2.2	The problem of common causes of votes $\ldots \ldots \ldots \ldots \ldots \ldots$	8	
	2.3	Partial solution: the Conditional Jury Theorem	9	
	2.4	The fundamental tension between independence and competence	11	

	2.5	The Competence-Sensitive Jury Theorem	12			
	2.6	The three jury theorems compared $\ldots \ldots \ldots \ldots \ldots \ldots \ldots$	14			
3	Jury theorems and diversity					
	3.1	Diversity versus competence heterogeneity	15			
	3.2	Competence heterogeneity threatens the wisdom of crowds $\ . \ . \ .$	16			
	3.3	Diversity as judgmental independence	18			
	3.4	Diversity affects both decision-making stages: deliberation and ag- gregation	19			
4	Problems and questions					
	4.1	Do correctness facts exist?	20			
	4.2	Are correctness facts underdetermined?	21			
	4.3	Unlimited evidence generation?	22			
	4.4	The problem of incredible truths	23			
	4.5	Deliberation: independence under miner or competence booster? $\ . \ .$	24			
	4.6	Group-dependent individual competence	25			
	4.7	Epistemic-strategic voting	26			
5	Other types of jury theorems 2					
	5.1	The epistemic aggregation problem more generally	28			
	5.2	Social choice between multiple alternatives	29			
	5.3	Aggregating estimates	29			
	5.4	Aggregating evaluations or grades	30			
	5.5	Judgment aggregation over interconnected propositions $\ldots \ldots \ldots$	31			
	5.6	Jury theorems for voters tracking individual facts	32			
6	Background and selected applications					
	6.1	Jury theorems and democratic theory	34			
	6.2	Jury theorems and social epistemology	36			
	6.3	Jury theorems and epistemic social choice theory $\hfill \ldots \ldots \ldots$	36			
7	Bib	Bibliography 3				

1 Introduction

1.1 What are jury theorems?

How many individuals should be entrusted with a given decision? Epistemically speaking, what matters is how the correctness or quality of decisions depends on the number of participants, for instance, how the correctness of jury verdicts ('innocent' or 'guilty') depends on jury size, how the quality of parliamentary decisions depends on parliament size, how the quality of election outcomes depends on the size of the electorate, or how the quality of expert advice depends on the size of the advisory body. Jury theorems address this problem. Their perspective is epistemic rather than procedural: the sole aim is to reach 'good' or 'correct' decisions by an external standard rather than to respect individual participation rights or ensure democratically legitimate decisions. The reliance on correctness facts is the central philosophical commitment of jury theorems (see Section 4.1).

This entry refers to the set of individuals participating in the decision as the *group* or *(decision) body*, and to the larger set from which this group is selected as the *population*. Examples are electoral bodies selected from the population of citizens, and scientific advisory bodies selected from the population of scientists. In real life, some decision bodies are very large: electoral bodies often contain millions of citizens. Other bodies have medium size: juries in court often involve 12 members. Yet other bodies have few or just one member: authoritarian states are ruled by small cliques or just one person.

Jury theorems are mathematical theorems of the following general structure:

Generic Jury Theorem: Given a choice problem of type X, a voting procedure of type Y, and premises about individuals Z, the epistemic performance of group decisions depends on group size in ways W.

This schematic statement contains several parameters, filled in differently by different jury theorems. The first and simplest jury theorem goes back to French enlightenment philosopher and mathematician Marie Jean Antoine Nicolas de Caritat, Marquis de Condorcet. In Condorcet's (1785) theorem,

- the choice problem is binary, like the convict-or-acquit problem of juries;
- the voting procedure is majority voting;
- the premises on individuals are competence and independence assumptions of particularly simple kinds;
- the conclusion is that 'crowds are wise' in the twofold sense that the probability of a correct group decision (i) increases in group size and (ii) converges to one as group size approaches infinity.

These two Condorcetian 'wisdom of crowds' conclusions have considerably

shaped the field, because they reappear (partly or fully) in many other jury theorems, whilst being controversial. The research programme of the jury literature is partly an attempt to place one or both Condorcetian conclusions on improved premises, and partly an attempt to derive jury theorems reaching rival conclusions, such as the optimality of some finite and possibly small group size (against Condorcet's first conclusion) or the fallibility of asymptotically large bodies (against Condorcet's second conclusion).

Taken in themselves, the Condorcetian conclusions represent two controversial wisdom-of-crowds hypotheses, which can be stated more generally (Dietrich and Spiekermann 2020):

The Growing-Reliability Hypothesis: Larger bodies are better truth-trackers. They make epistemically better decisions than smaller bodies or single individuals.

The Infallibility Hypothesis: Huge bodies are infallible truth-trackers. They make epistemically optimal decisions in the limit as group size tends to infinity.

Upon deeper analysis, the infallibility hypothesis is untenable, and cannot even serve as an approximation, idealisation or paradigm of how large-scale democracy performs (see Section 2). The fact that many jury theorems reach this overoptimistic conclusion has not always been helpful, by threatening the credibility of the jury literature as a whole. The right response is to revise the premises of jury theorems such that they become justified and imply plausible conclusions, such as growing reliability. Jury theorems with more plausible premises and conclusions are indeed achievable (see Section 2).

The growing-reliability hypothesis thus appears to be the more appropriate rendition of the wisdom of crowds. The infallibility hypothesis should arguably give way, as will emerge in Sections 2 and 4.2 (see Section 4.3 for a potential exception). Jury theorems reaching the infallibility conclusion can be mathematically interesting; philosophically, they might be seen asarguments against their own premises.

1.2 A broad notion of decision

Jury theorems address collective 'decisions' in a broad sense. Decisions could, for instance, stand for collective *beliefs* or collective *actions or choices*. In the legal arena, jury verdicts represent beliefs (about guilt or innocence), whereas court rulings represent actions (of convicting or acquitting defendants). In the political arena, parliamentary decisions and referendum decisions usually represent choices. Depending on its mandate, an ethical or scientific commission either produces collective beliefs (which serve as advice), or performs actions, e.g., sets a budget or institutes ethical standards.

But there is one key condition: decisions must be *correctness-apt*, i.e., be either factually correct or incorrect (or, more generally, of some correctness degree). Which facts determine correctness? Belief-type decisions are correct if the belief has true content. Action-type decisions are correct depending on some state; e.g., court rulings are correct depending on past actions of the defendant.

Collective beliefs are examples of collective *attitudes*. Other possible collective attitudes are collective values, desires, preferences, or intentions. They too can be formed through aggregation: think of preference aggregation rules generating collective preferences. Any collective attitude whose possession can be factually correct or incorrect is potentially an object of jury theorems. Whether attitudes can be correct is, of course, more controversial for non-belief attitudes (cf. Section 4.1).

In sum, jury theorems address collective decisions of any correctness-apt kind, including collective actions, beliefs, and other attitudes.

2 Three jury theorems

This section states and discusses three jury theorems about majority decisions between two alternatives. They assume a population of individuals, labelled 1,2,..., representing the possible members of the decision-making group. The population might contain, for example, all citizens of a state (for a political decision) or all scientists (for a scientific decision). It is infinite – an idealisation needed for considering arbitrarily large decision bodies. The group has any finite size $n \ge 1$ and consists of the first n individuals 1, 2, ..., n. Each member votes for one alternative. The alternative receiving a majority wins (if n is even, there can be a tie without collective decision). Following the epistemic approach, one alternative is *correct* (right, better, etc.) and the other is *incorrect* (wrong, worse, etc.), independently of the decision process. This makes any individual judgment and any group decisioncorrect or incorrect.

Jury theorems operate in a probabilistic framework: anything unknown is modelled as a random event or variable (defined relative to some background probability space¹).

¹A probability space is a mathematical representation of what could happen, and with what probability. Formally, it is a structure (Ω, \mathcal{E}, P) with the following components. Firstly, Ω is a non-empty set of *(possible) worlds*. Secondly, \mathcal{E} is a set of *events* $E \subseteq \Omega$, representing propositions (such as the deffendant is guility or individual 3 votes correctly or the majority votes correctly in the group of size 9). \mathcal{E} might contain all subsets of Ω ; in practice it contains at least those subsets which represent interesting propositions. Technically, \mathcal{E} must be a σ -

Notation: For any jury theorem of the entry, the probability function (of the underlying probability space) is denoted P. Where needed by a jury theorem, the event that individual $i \ (= 1, 2, ...)$ is correct is denoted R_i , and for each possible group size $n \in \{1, 2, ...\}$ the event that a majority is correct (i.e., that more than half of $R_1, ..., R_n$ hold) is denoted Maj_n .²

2.1 Condorcet's Jury Theorem

Condorcet's (1785) jury theorem will be discussed first on grounds of simplicity and historic importance, setting aside concerns about its premises and conclusions.Condorcet's text does not follow modern mathematics, but many later authors have stated his theorem formally (e.g., Grofman 1975; Grofman, Owen and Feld 1983).³ The theorem operates in a simple formal framework, with only the following ingredients:

Primitives of Condorcet's Jury Theorem: correctness events $R_1, R_2, ...,$ defined relative to a probability space.

In these primitives (and those of the two later jury theorems), the correctness events could be replaced by more basic primitives, namely the following random variables: votes (judgments) of the individuals 1, 2, ..., whose values are alternatives, and a 'state' variable, whose value is, or more generally determines, the correct alternative. Each correctness event R_i is then defined as the event that *i*'s vote equals the correct alternative, and the whole analysis stays unchanged. For instance, in a court's decision problem, the votes of judges could take values in {convict, acquit} and the state could take values in {guilty, innocent}, where convict is correct if the state is guilty, and acquit is correct if the state is innocent. Equivalently, and more parsimoniously, votes and state could all take values in {0,1}, where R_i is the event that *i*'s vote equals the state.⁴

algebra: it contains the tautology $(\Omega \in \mathcal{E})$, is closed under negation $(E \in \mathcal{E} \Rightarrow \Omega \setminus E \in \mathcal{E})$, and is closed under countable disjunction $(E_1, E_2, ... \in \mathcal{E} \Rightarrow E_1 \cup E_2 \cup \cdots \in \mathcal{E})$. Thirdly, P is a probability measure on \mathcal{E} , i.e., a function assigning a probability $P(E) \in [0, 1]$ to each event $E \in \mathcal{E}$, subject to the probability axioms: $P(\Omega) = 1$ and, for all disjoint events $E_1, E_2, ... \in \mathcal{E}$, $P(E_1 \cup E_2 \cup ...) = P(E_1) + P(E_2) + \cdots$.

²Formally, $Maj_n = \bigcup_{I \subseteq \{1,...,n\}: |I| > n/2} \cap_{i \in I} R_i$, the event that, for some subset I containing a majority, all members of I are correct.

³In the literature, the term 'Condorcet jury theorem' sometimes refers to slightly different theorems, or to a vague type of theorem. The most basic rendition is chosen here.

⁴Formal models often define the state as an alternative (the 'correct' one), so that the votes and the state all take values in the same set (the set of alternatives) and R_i is the event that *i*'s vote equals the state. This makes the model more parsimonious. But philosophically it lumps together different objects. In a highly explicit but unparsimonious model of the court's decision between *convict* and *acquit*, the state is neither the correct alternative, nor *quilty* or *innocent*,

Condorcet's theorem assumes that correctness is independent across individuals and has (the same) probability greater $\frac{1}{2}$ for each individual. Formally:

Unconditional Independence (UI). The individual correctness events $R_1, R_2, ...$ are (unconditionally) independent.

Unconditional Competence (UC). The (unconditional) individual correctness probability $P(R_i)$ – the general competence – exceeds $\frac{1}{2}$ and is the same for all individuals *i*.

Condorcet's Jury Theorem: Assuming UI and UC, the probability of majority correctness, $P(Maj_n)$, increases in (odd⁵) group size n and converges to 1.

This theorem paints an optimistic picture of the wisdom of crowds, by concluding that majority outcomes are not only growingly reliable as voters are added, but also infallible in the limit. Figure 1 illustrates growing majority reliability, for different levels of individual competence $p = P(R_i)$. Note the fast convergence to one, even for just slightly competent individuals.

Why does this theorem hold? Intuitively, UI and UC imply that the individual judgments behave like independent tosses of the same coin biased towards the truth. The more often the coin is tossed, the more likely the majority of the tosses are 'correct', and this likelihood converges to one.

Technically, the growing-reliability conclusion follows via a non-trivial combinatorial argument. The infallibility conclusion has an easy proof, which can be sketched here. By UC, general competence $p = P(R_i)$ is voter-independent. The group's proportion of correct votes is a random variable that converges stochastically to p by the law of large numbers, using UI. So, as $p > \frac{1}{2}$ by UC, the probability that this proportion exceeds $\frac{1}{2}$, which is precisely $P(Maj_n)$, converges to 1 as $n \to \infty$.

Although 'increases' should be read as 'weakly increases' in Condorcet's and the other two jury theorems, the stronger conclusion of strictly increasing $P(Maj_n)$

but a complex fact about the defendant's actions. Here the state takes values in a large set of possible complex facts, and the function mapping each possible state value to the corresponding correct attentive is many-to-one.

⁵The growing-reliability conclusion is restricted to bodies of odd size: the probability of majority correctness $P(Maj_n)$ increases as group size n moves from 1 to 3, to 5, etc. Excluding even-sized bodies is necessary because there can be ties in such bodies, reducing the probability of a correct majority. For instance, $P(Maj_n)$ falls from n = 1 to n = 2, because majority correctness in the 1-member group only requires individual 1 to be correct (the event R_1) whereas majority correctness in the 2-member group requires two individuals to be correct (the event $R_1 \cap R_2$). The restriction to odd n could be lifted if ties are broken by an independent toss of a fair coin.

Dropbox/Franz Dietrich/R-SocChWelf/JuryTheoremsETC/JuryThms-SEP/Figures/Fig1.png



Figure 1: The probability of majority correctness for different body sizes and competence levels.

holds in all non-degenerate cases of these theorems.⁶

2.2 The problem of common causes of votes

A tempting mistake is to believe that votes are probabilistically independent as soon as they are *causally* independent, i.e., do not affect one another. Votes are often causally independent: secret voting prevents someone's vote from influencing someone else's vote. Sometimes not only the votes (or voting acts), but even the individuals more broadly (their reasonings, perspectives, knowledge, votes, etc.) are causally independent. This happens if individuals do not deliberate or otherwise interact – a rare, but possible scenario. Deliberation immediately creates causal dependence between individuals, but not between their (secret) votes.

Causal independence (between votes or even individuals) by no means makes votes *probabilistically* independent, since common causes create correlations. This fact will be obvious to empirical scientists, statisticians, or causal-network theorists; see for instance Reichenbach (1956), Pearl (2000) and Hitchcock and Redei (2020) on common causes in general and Dietrich and List (2004) and Dietrich and Spiekermann (2013b) on common causes in jury theorems. Common causes of votes are factors influencing two or more votes. They exist in abundance. Voters are usually exposed to

⁶For instance, in Condorcet's Jury Theorem, $P(Maj_n)$ grows strictly *except* if individuals are infallible, i.e., if $P(R_i) = 1$.

- shared evidence, such as witness reports;
- shared perspectival influences, coming from a shared language, shared concepts, a shared methodology, or shared hypotheses;
- shared contextual influences that are unrelated to the decision task which affect judgment skills, such as noise or heat.

Unfortunately, the resulting correlations between votes – and ultimately between correctness events R_1, R_2, \ldots – are usually positive, which undermines diversity and reinforces tendencies and errors. Positive correlation between correctness events means: given that someone votes correctly, others are more likely to vote correctly. Why is the correlation positive? A correct judgment by, say, individual 1 (the event R_1) raises the probability that the common causes take a truth-conducive form (i.e., that evidence is non-misleading, contextual influences support judgmental skills, etc.), which in turn raises the probability that other individuals are correct (the events R_2, R_3, \ldots). Conversely, an error by individual 1 raises the probability of misleading common causes, which raises the probability of errors by others.

In sum, voting and correctness are interpersonally (positively) correlated through common causes, against Condorcet's independence assumption.

2.3 Partial solution: the Conditional Jury Theorem

Dependencies between votes can be 'conditionalised away'. *Given* the common causes – the shared evidence, etc. – votes no longer exhibit dependence, as they no longer carry new information about common causes, hence about one another. Conditionalisation blocks the information flow through common causes. These considerations apply a well-known principle: causally independent phenomena are probabilistically independent conditional on their common causes (Reichenbach 1956).

The conditionalisation strategy assumes that votes do not influence one another, i.e., are *causally* independent, as in secret voting. Otherwise information can flow between votes *directly*, not just via common causes, and conditionalising on common causes fails to make votes probabilistically independent.

To implement conditionalisation, introduce a random variable \mathbf{x} , called the *facts*, on which to conditionalise. One could interpret the facts as the totality of common causes, but simpler interpretations are possible. The facts \mathbf{x} could be some proxy of common causes. Most commonly, \mathbf{x} is the fact determining the correct alternative. This correct-making fact is standardly called the *state (of the world)*. In a court's convict-or-acquit decision, its possible values could be 'guilty' and 'innocent', which make 'convict' or 'innocent' correct, respectively. Sections 2.1, 4.1 and 5.1 discuss the state.

The new jury theorem will therefore rest on the following primitives.

Primitives of the Conditional Jury Theorem: correctness events $R_1, R_2, ...$ and an arbitrary random variable \mathbf{x} (the facts variable, representing for instance the state or the common causes), all defined relative to a probability space.

Convention: Definitions involving \mathbf{x} will assume that each value of \mathbf{x} has positive probability (which facilitates conditionalising, but restricts attention to discrete rather than continuous \mathbf{x}). All definitions could be generalised.⁷

These are the revised premises and theorem.

Conditional Independence (CI). The individual correctness events $R_1, R_2, ...$ are independent given any value x of the facts x.

Conditional Competence (CC). For any value x of the facts **x**, the conditional correctness probability $P(R_i|x)$ – the specific competence on x – exceeds $\frac{1}{2}$ and is the same for all individuals *i*.

Conditional Jury Theorem: Assuming CI and CC, the probability of majority correctness, $P(Maj_n)$, increases in (odd) group size n and converges to 1.

This theorem reaches the same optimistic conclusions as Condorcet's theorem, but based on new premises. The new independence assumption is more plausible to the extent that the facts \mathbf{x} include common causes or proxies thereof. If the facts are the state of the world – a *very* rough proxy of common causes – one obtains the literature's most common form of conditionalisation and the classical *State*-Conditional Jury Theorem (e.g., Austen-Smith and Banks 1996, Dietrich and List 2004). The Conditional Jury Theorem generalises this familiar jury theorem to an arbitrary target of conditionalisation \mathbf{x} . Replacing state-conditionalisation by arbitrary conditionalisation makes the jury theorem more flexible, opening the door to genuinely plausible conditional-independence assumptions.

How has the competence assumption evolved? Conditional Competence is logically stronger than Unconditional Competence: individuals must outperform fair coins not just globally, but given any facts. For instance, in a court's decision

⁷For general \mathbf{x} , the probability of an event A conditional on \mathbf{x} , $P(A|\mathbf{x})$, remains definable (see standard probability textbooks). But $P(A|\mathbf{x})$ is a *non-unique* function of \mathbf{x} if some or all values x of \mathbf{x} have probability P(x) = 0. Still any two versions of $P(A|\mathbf{x})$ are essentially identical: they coincide outside a zero-probability event; equivalently, they coincide except if \mathbf{x} takes a value in some zero-probability set of values. To restate CI and CC generally, after 'any value x of the facts \mathbf{x} ' add 'except possibly from a zero-probability set of values'. Later in Section 2.5, to restate TC generally, after 'is the same for all individuals i' add 'almost surely, i.e., except possibly in a zero-probability event' (the generalized definition of 'tending to exceed $\frac{1}{2}$ ' is given in Dietrich and Spiekermann 2013a).

to convict or acquit, where the facts are the state of guilt or innocence, each individual (judge) must be more likely correct than incorrect given guilt and also given innocence.

An individual *i*'s general competence is expressible as his (probability-weighted) average specific competence across possible facts, i.e., possible values x of \mathbf{x} :

$$P(R_i) = \sum_{x} P(R_i|x)P(x).$$
(1)

This shows that CC implies UC.

Unfortunately, the Conditional Jury Theorem maintains the unrealistic conclusion of asymptotic infallibility. Conditionalisation on facts can repair the independence premise, but another problem in premises remains, as will now be explained.

2.4 The fundamental tension between independence and competence

One might have hoped that the 'conditional' premises CI and CC are jointly justified for at least *some* conditionalisation, i.e., *some* suitably designed facts variable \mathbf{x} . Then the Conditional Jury Theorem would rest on justified premises, and we could trust its conclusions. But a dilemma arises:

- Conditional Independence is only plausible if one packs many facts into **x** (ideally all common causes).
- Conditional Competence is only plausible if one packs few facts into **x** (ideally *no* facts, so that CC reduces to UC, the maximally plausible instance of CC).

Therefore, CI and CC require different conditionalisation. They are almostnever justified *jointly*, but only justified *individually* for different facts variables \mathbf{x} (Dietrich 2008; Dietrich and Spiekermann 2020). This is the fundamental tension between independence and competence. This tension lies not in a logical contradiction between both, but in the nature of realistic decision problems (for a potential exception, see Section 4.3).

Why does this tension exist? Formal theorizing aside, it is obvious that whether someone is competent – i.e., more often right than wrong – depends on the type (reference class) of judgment tasks considered. Plausibly, people are more often right among *all* conceivable judgment tasks (the maximal reference class), and presumably also among many large reference classes, such as all guilty-or-innocent judgment tasks. But someone is presumably incompetent – more often wrong – among some very specific types of judgment tasks, such as all *confusing* tasks, and all guilty-or-innocent judgment tasks *with seemingly honest but lying witnesses*. Each instance x of the facts \mathbf{x} defines a type (reference class) of tasks: those tasks in which the facts are x. Some instances of the facts \mathbf{x} make it easy to form correct judgments: instances x where the evidence is transparent, witnesses are honest, laboratory tests are correct, etc. Here individuals i are competent: $P(R_i|x) > \frac{1}{2}$. Perhaps some other instances of the facts \mathbf{x} make it hard to form correct judgments: instances x with misleading evidence, etc. Here voters are not competent: $P(R_i|x) \leq \frac{1}{2}$. Whether misleading instances exist – i.e., whether CC fails – depends on the facts variable \mathbf{x} used. If all common causes are packed into \mathbf{x} (presumably to ensure Conditional Independence CI), then inevitably certain instances x of \mathbf{x} represent misleading circumstances, in which voters i have low competence $P(R_i|x) < \frac{1}{2}$. Ironically, if one packs much less information into the facts \mathbf{x} , possibly even reducing \mathbf{x} to the state (often a variable with just two instances), then one may well rehabilitate Conditional Competence, whilst sacrificing Conditional Independence. This is the dilemma.

2.5 The Competence-Sensitive Jury Theorem

The dilemma can be escaped with a jury theorem with more plausible premises and conclusions. Think of the facts \mathbf{x} as being rich enough to justify Conditional Independence; so \mathbf{x} contains the common causes, or a good enough proxy or substitute for them.

Although voters will be incompetent under some (unfortunate) instances xof the facts \mathbf{x} , they will plausibly be competent most of the time, i.e., under most facts instances. This suggests replacing Conditional Competence with the weaker premise of *Tendency to Competence* (Dietrich and Spiekermann 2013a; 2020). What is this axiom? It says that competence is more often high than low. More precisely, competence exceeds 0.5 by some amount at least as often as it falls below 0.5 by this amount, for any positive amount. So, competence is 0.51 at least as often as 0.49; it is 0.55 at least as often as 0.45; etc. Formally, an individual i's specific competence $P(R_i|\mathbf{x})$ depends on the (random) facts \mathbf{x} , hence is itself random. If its value exceeds $\frac{1}{2}$, the facts can be called *truth-conducive* or easy for i; if its value is below $\frac{1}{2}$, the facts can be called *misleading* or *difficult* for i; if its value is $\frac{1}{2}$, the facts can be called *neutral* for *i*, as they do not push in any direction. A discrete random number (such as specific competence $P(R_i|\mathbf{x})$) tends to exceed $\frac{1}{2}$ if, for each $\epsilon > 0$, it equals $\frac{1}{2} + \epsilon$ at least as likely as $\frac{1}{2} - \epsilon$. Figure 2 gives an example where specific competence (with values $0, \frac{1}{8}, \frac{1}{4}, \frac{3}{8}, ..., 1$) tends to exceed $\frac{1}{2}$: competence is $\frac{5}{8}$ with higher probably than $\frac{3}{8}$; it is $\frac{3}{4}$ with higher probably than $\frac{1}{4}$; etc. Visually, this makes the distribution lean to the right.

This leads to the following jury theorem (Dietrich and Spiekermann 2013a).

Primitives of the Competence-Sensitive Jury Theorem: same as for the Conditional

Dropbox/Franz Dietrich/R-SocChWelf/JuryTheoremsETC/JuryThms-SEP/Figures/Fig2.png



Figure 2: A competence distribution that tends to exceed $\frac{1}{2}$.

Jury Theorem.

Tendency to Competence (TC): The conditional correctness probability $P(R_i|\mathbf{x})$ – the specific competence – tends to exceed $\frac{1}{2}$ and is the same for all individuals *i*.

Competence-Sensitive Jury Theorem: Assuming CI and TC, the probability of majority correctness, $P(Maj_n)$, increases in (odd⁸) group size n and converges to a value which is below 1 unless CC holds.

What is the limit of majority correctness under CI and TC? It is the probability that the facts (including the shared evidence) are truth-conducive plus half the probability that the facts are neutral.⁹ Loosely speaking, group performance in the limit is as good as the facts. Unsurprisingly, the group cannot beat the facts. It should already count as a success to match the facts asymptotically.

The twofold evolution from the Conditional Jury Theorem is that the competence premise is weakened to TC and the infallibility conclusion is reversed. Therefore, majorities can be wrong even in asymptotically large bodies – unless CC holds, i.e., unless voters are *always* competent. 'Unless' can be read strongly in the theorem, as meaning 'if *and only if* it is not the case that'.

⁸The remark in Footnote 5 applies analogously.

⁹Formally, $\lim_{n\to\infty} Pr(Maj_n) = Pr(\mathbf{p} > \frac{1}{2}) + \frac{1}{2}Pr(\mathbf{p} = \frac{1}{2})$, where **p** denotes specific competence, $Pr(R_i|\mathbf{x})$, which does not depend on *i* by TC.

2.6 The three jury theorems compared

	Condorcet's	Conditional	Competence- $Sensitive$
	Jury Theorem	Jury Theorem	Jury Theorem
Dromino 1	Unconditional	Conditional	Conditional
1 Tennise 1	Independence	Independence	Independence
Dromino 0	Unconditional	Conditional	Tendency to
1 Temuse 2	Competence	Competence	Competence
Conclusion 1	Growing	Growing	Growing
Conclusion 1	Reliability	Reliability	Reliability
Conclusion 9	Asymptotic	Asymptotic	Asymptotic
	Infallibility	Infallibility	Fallibility

Table 1 summarizes the three jury theorems. All theorems yield growing reliability,

Table 1: Three jury theorems compared

but only the first two yield asymptotic infallibility. The independence assumption evolves from unconditional to conditional. Figure 3 illustrates how the competence



Figure 3: The three competence axioms UC, CC and TC

Competence

1.0

0.0

0.5

1.0

0.0

assumption evolves:

0.5

1.0

0.0

0.5

0.0

0.0

- a. Condorcet's Jury Theorem focuses on general competence, i.e., the fixed probability $P(R_i)$, which must exceed $\frac{1}{2}$ (see Plot a).
- b. The Conditional Jury Theorem focuses on *specific* competence, i.e., the factsdependent probability $P(R_i|\mathbf{x})$, which must always exceed $\frac{1}{2}$. If one takes \mathbf{x} to be the state of the world, i.e., considers the *State*-Conditional Jury Theorem, then (assuming the state is binary) only two competence levels are possible, one for each state (see Plot b1). Richer facts allow for many possible competence levels (see Plot b2).
- c. The Competence-Sensitive Jury Theorem assumes, less demandingly, that specific competence *tends* to exceed $\frac{1}{2}$ (see Plot c).

Mathematically, the Competence-Sensitive Jury Theorem generalises the Conditional Jury Theorem, which generalises Condorcet's Jury Theorem. Why?

- The Competence-Sensitive Jury Theorem encompasses the Conditional Jury Theorem as a special case, the case when CC holds, because in this extreme case (of *always* competent voters) the theorem asserts asymptotic infallibility.¹⁰
- The Conditional Jury Theorem encompasses Condorcet's Jury Theorem as a special case, obtained when plugging in a trivial facts variable **x** with only one possible value. Here the premises CI and CC reduce to Condorcet's premises UI and UC, because conditionalising on facts that are certain is like not conditionalising at all.

3 Jury theorems and diversity

Intuitively, diversity – in backgrounds, perspectives, information, reasoning modes, skills, etc. – improves collective performance. Where can diversity be found in formal jury frameworks? Do the assumptions of jury theorems implicitly rule out diversity, or instead permit or even require diversity?

3.1 Diversity versus competence heterogeneity

Diversity is describable as plurality in sources and backgrounds: individuals reason differently, hold different perspectives, use different information, etc. Although diversity is a multi-dimensional and largely informal property, it manifests itself in jury models. As what? There are two natural candidates: competence heterogeneity and independence. The relation to competence heterogeneity will be clarified first (Section 3.3 turns to independence).

Diversity does *not* imply competence heterogeneity – fortunately, as otherwise the jury theorems above would exclude diversity in the homogeneity clause of their competence assumptions. Diversity does not imply differences in competence *levels*, but differences in competence *sources*. Even under high diversity, members could be correct with identical probabilities – for different reasons.

Conversely, competence heterogeity tends to imply diversity, i.e., non-diversity tends to imply competence homogeneity, because members with exactly same background, same information, same reasoning, etc., should be equally competent.

In sum, competence homogeneity is a systematic feature of non-diverse bodies, but not vice versa.

 $^{^{10}}$ Asymptotic infallibility follows by the 'unless' clause. Recall that 'unless CC holds' means 'if and only if CC does not hold'.

3.2 Competence heterogeneity threatens the wisdom of crowds

Differential competence is the main threat to the growing-reliability hypothesis. The political upshot is that the epistemic superiority of democracy is at stake – is it epistemically better for a society to be run by a few competent citizens? The wisdom-of-crowds conclusion indeed breaks down in all three theorems if one removes homogeneity from the competence premise. Larger bodies can then perform worse, because new members may be (much) less competent than existing members. Majority reliability can even become a *decreasing* function of group size, and can converge to a level as low as $\frac{1}{2}$ if competence of new members converges fast enough to $\frac{1}{2}$.

Finite optimal group rather than wise crowds? The question of whether decision bodies should be augmented when new members are less competent is not far-fetched, but a notorious issue in real-life committee design. This is the trade-off: adding less competent members, one the one hand, lowers average member competence, but, on the other hand, brings the advantage of size, namely that errors of individual members are more easily overruled ('washed out') by the majority. Depending on which effect is stronger, the probability of majority correctness increases or falls by adding less competent members.

Whether some (finite) group is optimal or 'bigger is always better' has been treated analytically. Answers depend on model assumptions (see Karotkin and Paroush 2003).

How could the wisdom of crowds be rehabilitated? Both wisdom-of-crowds hypotheses are discussed in turn.

Asymptotic infallibility despite heterogenous competence. The asymptotic conclusion of Condorcet's Jury Theorem remains true if UC is weakened to the existence of a fixed lower bound $\frac{1}{2} + \epsilon$ for everyone's competence (Paroush 1998). This assumption can be relaxed further: it suffices that the infinitely many individuals are competent on average, which allows for many incompetent individuals¹¹ (see Dietrich 2008; for related or more general results: Boland 1989; Berend and Paroush 1998; Owen et al. 1989). Unfortunately, these variants of Condorcet's Jury Theorem (and analogous variants of the Conditional Jury Theorem) retain only the infallibility conclusion, sacrificing growing reliability. The

¹¹That is, *infinite average competence* exceeds $\frac{1}{2}$. Infinite average competence is the limit of finite average competence, $\lim_{n\to\infty} \frac{1}{n} \sum P(R_i)$, or in full generally (since this limit need not exist) the 'limit inferior' or 'limiting lower bound' of finite average competence, $\lim_{m\to\infty} \inf_{n\geq m} \frac{1}{n} \sum P(R_i)$. A limit inferior always exists, and coincides with the ordinary limit when existent.

more problematic conclusion has been retained, one might complain.

Growing reliability despite heterogenous competence. There are broadly two approaches to rehabilitate the growing-reliability hypothesis and thereby defend large-scale democracy in the face of differential competence.

The first approach effectively denies that larger bodies have 'worse' members, by postulating a different procedure to determine group members: members are selected anonymously, rather than (for instance) by competence. Increasing group size no longer means keeping the 'old' members and adding 'new', often less competent, members, but it means drawing a fully new (larger) group from the population.¹² This procedure for selecting members differs significantly from that where the members must be the first n individuals in a predefined 'priority order' of the individuals. The new procedure is anonymous, instead of effectively prioritising individual 1 over individual 2, individual 2 over individual 3, etc. Which procedure is more realistic is debatable and context-dependent.

A second approach assumes – quite realistically – that individual competence levels are unknown. Even if additional group members were *objectively* less competent, this fact would be insufficiently known or established (cf. Romeijn and David Atkinson 2011). The institutional designer can consider three kinds of justification for a large, inclusive group. An epistemic-democratic justification points out that any proxy used for competence might be misleading. For instance, it would seem unjustified to exclude citizens without university education because they might be just as competent or more competent; excluding them might lower collective competence. One procedural-democratic justification holds that it is unfair to exclude individuals from democratic decision-making even if their incompetence were established beyond reasonable doubt. Such procedural reasons for inclusion could, for example, be based on the democratic right to equal opportunity for influence. Another, more nuanced procedural view does not generally oppose voter exclusion, but requires the incompetence to be established to a high standard, as those exclude are owed reasons for being ruled by others.

A version of the last view is defended by Estlund (2008). He grants that some individuals can be more competent than others, thereby opening the door to arguments for 'epistocracy' (decision by small bodies of competent citizens)

¹²To model this, consider a large finite population \mathcal{I} of potential members, e.g., all living scientists in case of a scientific group. For each possible group size n (at most the population size $|\mathcal{I}|$), the *n*-member group is drawn randomly from the set of possible *n*-member bodies $\{N \subseteq \mathcal{I} : |N| = n\}$ following a uniform probability distribution. So decisions are doubly random: member judgments and member identity are both random. This restores the growing-reliability conclusion, despite competence heterogeneity (Berend and Sapir 2005). The finiteness of the population \mathcal{I} reflects reality (and makes the mentioned uniform distribution well-defined); but it sets an upper bound on group size, thereby excluding asymptotic considerations.

rather than democracy (decision by all or most citizens, directly or indirectly). However, Estlund rejects epistocracy on grounds that competence levels cannot be identified beyond reasonable doubt. Individuals excluded from decision-making could rightfully object to being ruled by others (Estlund 2008: ch. 11). Estlund's proposal has led to a lively debate on arguments for and against epistocracy and the role of uncertainty about competence (Lippert-Rasmussen 2012; Brennan 2018; Goodin and Spiekermann 2018, ch. 15; Gunn 2019).

To formally tackle limited knowledge about individuals, one could, for instance, introduce subjective probabilities over competence levels, i.e., over objective correctness probabilities. More simply, one can eliminate objective probabilities and competence altogether from the model and work solely with subjective probabilities over individual judgments (rather than over individual competence levels). If individuals are subjectively indistinguishable, i.e., 'look' the same, judgments can be assumed to be *interchangeable* in de Finetti's sense of permutation-invariance. This implies (with an additional assumption) that larger bodies are more reliable, without becoming asymptotically infallible, by a *jury theorem for interchangeable voters* (Dietrich and Spiekermann 2013; generalising Ladha 1993). Here, 'reliable' and 'infallible' are no longer objective features of the group, but features of knowledge about it.

This subjectivist approach begs the question of who should hold these credences. But at least it takes the problem of uncertainty about individuals seriously.

3.3 Diversity as judgmental independence

Diversity manifests itself as probabilistic *independence* between individual judgments. If diversity is small – i.e., individuals reason similarly, use similar evidence, etc. – then judgments are highly correlated. Without any diversity, judgments coincide, and the correctness events $R_1, R_2, ...$ are equivalent. If diversity is high, then judgments correlate less. Extreme diversity corresponds to full independence – or even to negative correlation, to mention a rare but interesting possibility (Hong and Page 2009; 2012).

Condorcet's Jury Theorem implicitly presupposes extensive diversity by assuming independence. The Conditional and the Competence-Sensitive Jury Theorem are less restrictive: they permit more or less diversity, because their *conditional* independence assumption is compatible with more or less *unconditional* independence.

Many models of voter dependence have been proposed. A classical model introduces an 'opinion leader' whose judgment each other individual copies with some probability (Boland et al. 1989). Instead of a single opinion leader, the individuals could influence one another, as captured by pairwise correlations (e.g., Ladha 1992). The dependence structure could become yet more complex and escape a description in terms of pairwise correlations (Kaniovski 2010; Pivato 2017).

Some authors focus on the merits of diversity, by presenting theorems in which more independence (diversity) leads to better majority decisions (e.g., Berg 1993; Ladha 1995). The mechanism at work is this: more independence makes it more likely that errors of voters are compensated by correctness of other voters. Many authors also present jury theorems, showing that, despite some dependence or non-diversity, crowds can be 'wise', in the sense of growing reliability (e.g., Berg 1993) or asymptotic infallibility (e.g., Pivato 2017). The infallibility conclusion remains objectionable (cf. Section 2).

Modelling voter dependence in adequate, transparent and tractable ways remains a central challenge for future jury theorem research.

3.4 Diversity affects both decision-making stages: deliberation and aggregation

Collective decision-making can be subdivided into a deliberation phase, in which the individuals form or revise their judgments, and an aggregation phase, in which the post-deliberation judgments are fed into a voting rule which returns the decision (Dryzek and List 2003). Diversity affects not only the judgment profile in the group at any given moment, and hence the aggregation process, but also the deliberation process. Deliberation owes its richness and fruitfulness largely to diversity. Non-diverse bodies gain little from deliberating: their members already reason similarly, take similar perspectives, know similar things, etc. In diverse bodies, deliberation can broaden the informational, perspectival, and methodological horizon. The more diversity, the more potential for judgment revision through deliberation.

Causal graphs help to understand how diversity affects deliberation. Think of the votes (judgments) as embedded in a causal graph. Each individual's vote is a (probabilistic) causal consequence of this individual's judgment sources: informational sources, intellectual influences, etc. The more the sources differ across individuals, the more diverse the group is. In deliberation, members share their sources. This changes the causal structure: whenever a source is shared with an individual i, a causal arrow is added from that source to i's vote. This enlarges a voter's spectrum of sources. Figure 4 gives an illustration with just two sources and three individuals. Source sharing makes the basis of judgments more similar across individuals, possibly letting judgments converge. Ideal deliberation achieves full source sharing, as in Figure 4. The more diverse the group was initially, the





Figure 4: Source sharing through deliberation.

more sources there are to share, hence the more deliberation can 'do'.

While deliberation reduces interindividual diversity by removing source asymmetries, it creates intraindividual diversity by widening the spectrum of someone's sources. Individuals 'internalise' diversity when deliberating.

4 Problems and questions

4.1 Do correctness facts exist?

The controversial philosophical premise of jury theorems is that alternatives are factually correct or incorrect (or of some degree of correctness). The (possibly composite) fact that determines correctness is called the *state of the world* or just *state*.¹³

To disambiguate, notice that formal models often use simplified notions of the state. Firstly, many models reduce the state to its unknown part; for instance, the state in a court decision might be modelled as the fact of whether a crime has been committed, treating all other determinants of correct verdicts (such as the right interpretation of the law, but also logical or physical facts) as known background facts not worth modelling. Secondly, and more radically, many models identify the state with the correct alternative itself; they might take the state in a court decision to *be* the correct verdict ('convict' or 'acquit').

Setting modelling practice aside, which (known or unknown) facts determine correctness, i.e., constitute the state? One can distinguish between logical, empirical and normative facts. When a group of logicians votes on whether an argument is valid, correctness depends on logical facts. When a panel of climate experts predicts the average global surface temperature in 2100, correctness also depends on empirical facts. When a parliament decides on whether to allocate more funding

 $^{^{13}\}mathrm{In}$ a more restricted sense, the 'state' consists only of the unknown determinants of correctness.

to the health system, correctness depends on normative facts, besides empirical (and logical) facts. In many social decisions, correctness depends both on unknown empirical facts and unknown normative facts, particularly for decisions on actions or on beliefs about what is valuable or ought to be done.

Critics of jury theorems often question the very notion of correctness. Some doubt the existence of normative facts necessary for correctness in political, moral, or other evaluative decisions. This critique is often voiced in the form that there is no 'truth' pertaining to decisions in a certain domain (e.g., Muirhead 2014 criticizing Landemore 2013; cf. Gaus 2011 for how Estlund 2008 stakes out the domain of truth-apt political claims). Certain decisions might indeed be nonepistemic. Judgments of taste, for instance, might express desires, or perhaps beliefs without correctness fact. For moral judgments, both the existence and the nature of correctness facts is controversial.

Note, however, that the existence of moral and other normative facts does not depend on meta-ethical realism, universalism, or naturalism. Constructivism about moral facts, for example, could provide an intersubjectively shared social fact that is perfectly suitable for jury theorems, as long as the construction predates decision-making, i.e., occurs previously in the larger population.

In sum, all that jury theorems need is the existence of correctness facts of *some* kind that are process-independent, i.e., independent of the choice of group (size), the deliberation, and the aggregation.

In political settings, the real problem often lies not in non-existent correctness facts, but in ambiguously defined decision problems. For example, in presidential elections, is the question who best promotes the public good, or who best satisfies the citizens' preferences? Ambiguous decision problems lead to ambiguous correctness standards. Jury theorems stay applicable *in principle*, subject to resolving ambiguities.

4.2 Are correctness facts underdetermined?

Many group decision problems display what might be called *Possible Underdetermination*: the true state of the world can (with non-zero probability) be objectively underdetermined by the totality of influences on (one or more individuals in) the population, which includes all evidence.¹⁴ A jury's choice between 'guilty' and 'innocent' verdicts normally displays Possible Underdetermination, since the

¹⁴More precisely, assume all individual judgments are functions of a variable \mathbf{t} , 'total influences', interpretable as the viector of all private or shared causes of judgments. Underdetermination is the event that \mathbf{t} takes a value conditional on which more than one state has non-zero probability. *Possible Underdetermination* means that underdetermination has positive probability ('possible' thus stands for probabilistic, not just logical, possibility).

total available evidence can be objectively inconclusive (despite supporting one of these states). Dietrich and List's (2004) jury model implicitly assumes Possible Underdetermination, because even the ideal interpretation of total evidence can be incorrect. Landemore (2013, 145) mentions settings in which the correctness of the decision remains objectively uncertain, since the truth is never revealed with certainty. This could be interpreted as a strong ('persisting') type of underdetermination: the truth is objectively underdetermined not just by the influences at the time of decision, but even by all future influences.

Possible Underdetermination implies that individual judgments, the group judgment, and even the asymptotic group judgment (as group size increases) are all fallible, i.e., incorrect with non-zero probability. The reason is simple: each of these judgments is determined by something (the total influences) that can underdetermine the state, hence does not always match the state.

Exceptions exist. When mathematicians vote on the truth of a mathematical conjecture, the correct decision is given by logical facts, hence is never underdetermined. Even then, individual judgments and (finite or asymptotic) group judgments can be incorrect, because of a positive probability that the objectively accessible truth is not subjectively recognized, say due to subjectively misleading evidence, intransparent logical facts, or distracting circumstances.

The lesson is that the infallibility hypothesis is untenable under Possible Underdetermination, and dubious even without Possible Underdetermination. The next subsection will, however, discuss a special scenario in which the group *is* asymptotically infallible.

4.3 Unlimited evidence generation?

Perhaps the only case once can try to make for asymptotic infallibility involves the idea of unlimited evidence generation (or 'generated signals' in the words of Hong and Page 2009). What does this mean, and how plausible is it?

Assume that each individual accesses independent private information, on which alone their judgment is based, without any information sharing or dependence on common influences (as in Dietrich and List 2004). For instance, doctors perform independent tests on one and the same patient to judge whether some virus is present, or chemists perform independent experiments to judge whether a liquid contains a given molecule. This private information is truth-conducive: given any true state, the information supports the true state with probability above $\frac{1}{2}$ – the same probability for each information (each individual). More problematically, private evidence is independent given the true state. This rules out, implausibly, that any common causes (except the state itself) affect different private information: for instance, the patient's physiology cannot affect virus

tests (possibly rendering all tests unreliable), and the liquid's acidity level cannot affect tests for the molecule (possibly rendering all tests unreliable). Under these questionable assumptions, majority judgments are asymptotically infallible, as the (State-)Conditional Jury Theorem applies.

Such a scenario of fully private and independent information escapes the problem of common causes (Section 2.2) and the dilemma between independence of competence (Section 2.4), but only by excluding hidden common influences, excluding informational exchange between individuals, postulating an ever extendable rather than fixed body of information, and thereby indirectly excluding underdetermination of the truth discussed in Section 4.2.¹⁵ An individual 'creates' new independent information (in the examples: through 'experiments') rather than facing the same shared information as others. Increasing the group thus adds information, not interpretations of information. This unlimited availability of independent information might be approximately realistic in decision problems where information is *produceable* (say, through experiments) rather than merely *observable*. But political and other real-life decision problems come with a limited and possibly difficult body of known or knowable facts, causing asymptotic fallibility.

4.4 The problem of incredible truths

Some truths are hard to recognize as they are highly specific, and thus unlikely to hold. Recognizing that global temperature will rise may be easy, but recognizing that it will rise by 2.3 Degrees Celsius is hard. Whenever some choice alternative (and thus its correctness) is highly specific, competence in jury models is threatened. Suppose a climate panel must form a belief about the proposition p that global temperature will rise by 2.3 degrees. Someone's competence given p (or given richer facts that entail p) can easily fall below $\frac{1}{2}$, because correctly recognizing p is hard. Indeed, the available evidence might be compatible with a temperature rise close to, but distinct from 2.3 degrees. The other effect of p's high specificity is that competence given not-p (or given richer facts that entail not-p) can become very high, because not-p is a very unspecific and thus 'credible' proposition. The lesson is that highly unbalanced choice alternatives – a highly specific alternative against a highly unspecific alternative – bring the competence assumption of the Conditional Jury Theorem to fall.

By contrast, the competence assumptions of Condorcet's and the Competence-Sensitive Jury Theorem, UC and TC, are not vulnerable to this problem. Why?

• UC pertains to the general ('unconditional') competence of an individual $i, P(R_i)$, not their specific competence given facts $x, P(R_i|x)$. General

¹⁵The infinite sequence of private evidence determines the truth with probability one, by the law of large numbers.

competence stays above $\frac{1}{2}$, because the (error-prone) event of the 'specific truth' is much less likely than the (truth-conducive) event of the 'unspecific truth'.¹⁶

• TC explicitly allows (specific) competence to sometimes fall below $\frac{1}{2}$.

The problem of incredible truths is a variant of Estlund's (2008: 232–4) 'disjunction problem' and of problems diagnosed in List (2005) and Dietrich and Spiekermann (2020).

4.5 Deliberation: independence underminer or competence booster?

Two intuitions compete: does deliberation primarily threaten collective epistemic success, by reducing judgmental independence, *or* primarily increase epistemic success, by raising individual competence?

The first intuition needs nuancing. Deliberation reduces *unconditional* judgmental independence, by adding common sources of votes; see Section 3.4 and Figure 4.¹⁷ Thus, deliberation further undermines Condorcet's (naive) Unconditional Independence axiom. But deliberation does not undermine the Conditional Independence axiom of the Conditional and the Competence-Sensitive Jury Theorem, because one can conditionalise on common causes.¹⁸

Turning to the second intuition, deliberation certainly affects competence. Since deliberation tends to widen the basis of judgments (see again Section 3.4), the effect on competence tends to be positive, as one may conjecture. All three competence axioms considered in Section 2 would then hold more easily with deliberation than without. This conjecture is illustrated in Figure 5. Plots a, b and c refer to the competence axiom of Condorcet's, the Conditional, and the Competence-Sensitive Jury Theorem, respectively. In all three plots, the competence axiom is violated pre-deliberation, and holds post-deliberation – as an extreme example where deliberation is of vital importance. In Plot a, deliberation lets general competence $P(R_i)$ move from 0.4 to 0.8 post-deliberation, so that

¹⁶As equation (1) shows, an individual *i*'s general competence $P(R_i)$ is the probability-weighted average of her specific competence $P(R_i|x)$ across facts x. This average can exceed $\frac{1}{2}$ even if $P(R_i|x) < \frac{1}{2}$ for some x of low probability.

 $^{^{17}}$ Deliberation lets individuals causally affect one another. Despite causation between individuals, there is normally no causation between *votes*, at least if voting is silent (cf. Section 2.2). This is why deliberation introduces common causes of, not causation between, votes.

¹⁸A group that has deliberated usually has more common causes of votes than a group that has not. So one may want to conditionalize on more facts than one would have in the absence of deliberation. This points towards a general issue: which sort of facts variable \mathbf{x} makes the independence premise CI of a jury theorem plausible depends on several concrete circumstances, including whether and which deliberation happens prior to voting. Fortunately, the Conditional and Competence-Sensitive Jury Theorems hold for *any* facts variable \mathbf{x} .

Dropbox/Franz



Figure 5: Three examples where deliberation improves competence. Grey: predeliberation. Black: post-deliberation.

Condorcet's competence axiom UC becomes satisfied. In Plots b and c, the distribution of specific competence across facts moves upwards, making the respective competence axiom satisfied.

To be clear, deliberation sometimes distorts judgments. Sometimes misleading evidence and other judgment-deteriorating sources are shared, opinion cascades occur, bad opinion leaders emerge, etc. (e.g., Sunstein and Hastie 2014). So the hypothesis that deliberation improves competence should be qualified: deliberation usually improves judgments, thereby helping competence axioms like UC, CC and TC become satisfied.

4.6 Group-dependent individual competence

The growing-reliability hypothesis is threatened not only if additional group members are less competent than existing members (Section 3.2), but also if *existing* members lose competence when the group grows too much. Firstly, the deliberation process, to which members partly owe their competence (Section 4.5), can become less fruitful: some bodies are too large for successful deliberation. Secondly, large 'anonymous' bodies can demotivate members, who feel individually less responsible and spend less effort on forming correct judgments. The opposite effect is also imaginable: adding members can improve deliberation and motivation. Bodies can be too small, not just too big (see the debates over optimal parliament size, e.g., Elster 2012).

All this puts a fundamental assumption of jury models into question: the assumption that someone's judgment and thus competence are group-independent, hence, for instance, are the same in a 3-member group (with small-scale deliberation) as in a 333-member group (with large-scale deliberation). This denies any effects of group dynamics on judgments.

To model group-dependence, the event R_i that an individual *i* is correct should be replaced by group-size-dependent correctness events $R_{i,n}$ for all group sizes *n* such that $n \ge i$ (i.e., such that the group includes *i*). An individual *i*'s competence, general or specific, becomes a group-size-dependent quantity, given by $P(R_{i,n})$ or $P(R_{i,n}|\mathbf{x})$, respectively. One might hypothesize that competence is an initially growing and subsequently falling function of *n*, which therefore peaks at some 'individually optimal' group size n_i . If one replaces the ordinary assumption of group-independent judgments by the latter hypothesis, then the growing-reliability conclusion of jury theorems will still hold when restricted to group sizes *n* of at most the individual optimum n_i of the group members i = 1, ..., n. The collectively optimal group size can exceed the individual optimum of all individuals *i*, because the merits of drawing on many minds can outweigh individual competence losses. But the collective optimum can become finite, against the growing-reliability hypothesis.

4.7 Epistemic-strategic voting

Conventional jury models implicitly assume that individuals share the same objective of a correct collective decision. Surprisingly, this assumption does not rule out strategic voting. To the contrary, voters may vote against what they believe to be correct *in order* to facilitate a correct aggregate decision.

This can be called *epistemic*-strategic voting, because voters strategise out of a shared epistemic objective rather than conflicting interests. How is this even possible? Suppose a 9-member jury reaches a guilty-or-innocent verdict by majority. A juror reasons: 'I believe in guilt. My vote only makes a difference if it is pivotal, i.e., if my jury colleagues are split 4:4. So, let me assume this split. But then four (competent) colleagues believe in innocence. So innocence is more likely than I thought. The higher probability of innocence justifies an 'innocent' verdict. So I shall vote *innocent*.' She votes on the basis, not of her beliefs, but her conditional beliefs assuming pivotality. If everyone reasons alike, no one reveals their genuine judgment. Votes reflect no private information or insights, and collective decision are arbitrary. Defendants are unanimously acquitted even when everyone believes in guilt.

Fortunately, this absurd situation is no more stable than sincere voting. Why? The strategic reasoning leading to non-sincere voting cannot be universalised: it assumes others vote sincerely. Indeed, our example juror assumes her colleagues are sincere in her 'But then' inference.

Epistemic-strategic voting has been addressed game-theoretically, using jury

models enriched by voters' private information; see Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1999), Peleg and Zamir (2012), Bozbay et al. (2014), and many others. The generic finding is that sincere voting by everyone is no (Nash) equilibrium, but that the rationality of sincere voting can be restored by carefully adjusting the aggregation rule, raising a problem of procedure ('mechanism') design rather than group design.¹⁹

But is epistemic-strategic reasoning and voting really plausible? There are three very different objections. Firstly, voters may be boundedly rational and unable (or unwilling) to strategise. Secondly, variations of the decision procedure, possibly introducing pre-voting deliberation can make strategic voting less likely (e.g., Coughlan 2000; Gerardi and Yariv 2007). Thirdly, and most interestingly, sincere voting becomes perfectly rational under a richer and arguably more realistic picture of voter motivation, where the narrowly consequentialist concern for correct collective outcomes is replaced or complemented by a concern about the act of voting itself, often an intrinsic concern for being sincere, or for expressing one's opinion, or for complying with norms of sincerity. The literature contrasts instrumental and expressive voting (Brennan and Lomasky 1993; Schuessler 2000). Why do expressive concerns (even if just small) easily crowd out instrumental concerns, so that sincerity becomes rational? The reason is that, in sufficiently large group s, the probability of being pivotal, i.e., of affecting the outcome, is small, so that the instrumental concern almost cancels out when solving the voter's optimisation problem. So, strategic voting may well be a game-theoretic artifact of ascribing purely instrumental preferences to voters (Goodin and Spiekermann 2018: 45–50; Dietrich and Spiekermann 2020).

5 Other types of jury theorems

The jury theorems discussed above apply to binary choices and majority rule. This section defines the epistemic aggregation problem more generally (Section 5.1), and then discusses jury theorems for aggregating votes over multiple alternatives (Section 5.2), aggregating estimations (Section 5.3), aggregating evaluations or grades (Section 5.4), aggregating judgments over interconnected propositions (Section

¹⁹One has to replace the majority threshold by an acceptance threshold such that pivotal voters prefer to tip the decision in the direction of their private belief, rather than ignoring private information. So, clever institutional design might prevent strategic voting. Yet there are obstacles. One obstacle is that the particular acceptance threshold that induces truthful voting is highly sensitive to parameters such as voters' prior correctness probabilities and strength of private information. Without knowing these parameters, the threshold cannot be set accordingly. Worse, the acceptance threshold that induces truthful voting may vary from voter to voter (for instance due to different prior beliefs), so that *no* acceptance threshold simultaneously prevents all voters from strategising.

5.5), and aggregating votes of voters tracking individual facts (Section 5.6). Many theorems reach the implausible infallibility conclusion; the critical analysis of this conclusion (Section 2) continues to apply.

5.1 The epistemic aggregation problem more generally

An epistemic aggregation problem has the following components.

Alternatives, votes, and aggregation. Let \mathcal{A} be a non-empty set of 'alternatives' in the most general sense – e.g., choice alternatives, belief or judgment sets, or value assignments to options. A decision-making group selects one alternative based on the members' votes, applying some aggregation rule F that maps each profile $(a_1, ..., a_n) \in \mathcal{A}^n$ of any number $n \geq 1$ of votes to a decision $F(a_1, ..., a_n) \in \mathcal{A}$ (or perhaps a subset $F(a_1, ..., a_n) \subseteq \mathcal{A}$, to allow for indeterminate outcomes).²⁰ For example, for majority rule (assuming \mathcal{A} contains just two alternatives), $F(a_1, ..., a_n)$ is the alternative in \mathcal{A} occurring more than n/2 times among $a_1, ..., a_n$.²¹

Aggregation problems in which voter inputs and collective decisions differ in type (rather than both belonging to the same set \mathcal{A}) can also be of interest.²² We set them aside here.

Correctness, state, and votes. Each alternative possesses a 'true' or 'objective' value, quality, or correctness level. If alternatives are beliefs or belief sets, value usually depends on truth; if alternatives are evaluations of objects, value might depend on distance to correct evaluations; etc. Value is determined by an unknown 'state'. Let S be the non-empty set of *possible states*. The votes of individuals 1, 2, ... and the state are interpreted as the outcomes of random variables $\mathbf{v}_1, \mathbf{v}_2, \ldots$ with range \mathcal{A} and \mathbf{s} with range \mathcal{S} , respectively. One can be more or less sophisticated:

• The 'simple' standard of correctness distinguishes between just two correctness levels, 'correct' and 'incorrect', and identifies the state with the (single) correct alternative: S = A. Epistemic performance is measured by correctness probability, i.e., probability that the judgment matches the state. So, individual *i*'s performance is measured by $P(\mathbf{v}_i = \mathbf{s})$, and collective performance by $P(F(\mathbf{v}_1, ..., \mathbf{v}_n) = \mathbf{s})$.

²⁰Formally, F is a function on $\cup_{n=1,2,\ldots} \mathcal{A}^n$.

²¹Ties, in which both alternatives occur n/2 times among $a_1, ..., a_n$ (for even n), could be handled in different ways. Either some tie-breaking rule selects a winning alternative. Or one allows outcomes to be sets, in which case either both alternatives are winners ($F(a_1, ..., a_n) = \mathcal{A}$) or no alternative is a winner ($F(a_1, ..., a_n) = \emptyset$).

²²For instance, voters might submit more complex inputs than the collective output, e.g., rankings or 'approval sets' over possible outputs.

Under a more general ('graded') standard of correctness, each pair (a, s) of an alternative and a state is assigned a number V(a, s), the value or correctness degree of a in state s. This defines a value function V from A × S to R. Epistemic performance is measured by expected decision value: E(V(v_i, s)) measures individual i's performance, E(V(F(v₁,...,v_n), s)) measures collective performance. This general case is more flexible. It can, for instance, distinguish between type-1 errors (like false convictions) and type-2 errors (like false acquittals). The simple standard of correctness is a special case.²³

The question. While much of epistemic social choice theory seeks to optimize the aggregation rule F given a fixed group size n, jury theorems are an exception. They fix F and vary n, addressing how collective epistemic performance – measured by correctness probability or more generally expected decision value – depends on group size. Notorious questions are: does performance increase in group size (growing reliability) or peak at some finite group size? And how well does the collective perform in the limit?

5.2 Social choice between multiple alternatives

One can state jury theorems that apply at once to several aggregation rules and choice problems, showing that almost nothing hinges on the classic focus on binary choice and majority rule (Pivato 2017). More concretely, List and Goodin (2001) analyse plurality rule over some finite set of alternatives \mathcal{A}^{24} Adopting the simple standard of correctness (see Section 5.1), they show that the plurality outcome is asymptotically infallible under largely classical assumptions of independence, homogeneity, and competence. Interestingly, given a correct alternative $a \in \mathcal{A}$, a voter's probability of picking a is not required to exceed 0.5, but only to exceed the probability of picking b for each incorrect alternative $b \in \mathcal{A} \setminus \{a\}$.

5.3 Aggregating estimates

Assume the group must estimate or predict some real-valued quantity, such as inflation, income inequality, or global temperature. Here, \mathcal{A} is a real interval, e.g., \mathbb{R} or [0, 1], containing the possible estimates. Let the group aggregate members' estimates by taking their average (rather than their median, as in Galton's 1907 classic paper on the wisdom of crowds). Formally, $F(x_1, ..., x_n) = \frac{1}{n} \sum_{i=1}^n x_i$ for any group size n and any individual estimates $x_1, ..., x_n \in \mathcal{A}$. The state \mathbf{s} is the

²³This is because expected decision value reduce to correctness probability if $\mathcal{A} = \mathcal{S}$ and moreover the value V(a, s) is 1 for correct decisions (a = s) and 0 for incorrect decisions $(a \neq s)$.

²⁴Plurality rule selects the alternative receiving maximally many votes. Ties could be broken in different ways, as for majority rule (see footnote 21).

true amount in $\mathcal{A} (= \mathcal{S})$. Assume that, conditional on the state **s**, the individual estimates $\mathbf{v}_1, \mathbf{v}_2, ...$ are (i) independent, (ii) identically distributed, and (iii) correct in expectation.²⁵ Condition (i) and (ii) reflect familiar ideas of voter independence and homogeneity. Condition (iii) captures competence as unbiasedness. Then, by Kolmogorov's (1930) celebrated (strong) Law of Large Numbers, the average estimate $F(\mathbf{v}_1, ..., \mathbf{v}_n)$ converges to the true state **s** with probability one as *n* tends to infinity.²⁶ This implies asymptotic infallibility.²⁷

This turns the Law of Large Numbers into a jury theorem for aggregating estimates. Its independence and homogeneity assumptions could be significantly weakened.²⁸ But the (suspicious) unbiasedness condition is essential, unfortunately.

5.4 Aggregating evaluations or grades

Some option, prospect, or other object is being evaluated ('graded') in terms of some criterion. A painting might be evaluated in terms of beauty: is it 'beautiful', 'neutral', or 'ugly'? An academic applicant might be graded in terms of research skills. Let \mathcal{A} be a finite set of possible grades, linearly ordered from 'highest' to 'lowest'. The state s is the true value or grade in $\mathcal{A} (= \mathcal{S})$. Each group member assigns a grade. The resulting grading profile (a_1, \ldots, a_n) is aggregated into the median grade $a = F(a_1, ..., a_n)$, i.e., the middle grade after putting $a_1, ..., a_n$ into a weakly increasing order $a'_1, ..., a'_n$ (to ensure medians exist, let n be odd). If, for instance, a painting is evaluated twice as 'beautiful' and once as 'ugly', where n=3, then the median evaluation is 'beautiful'. Given an instance s of the state, consider a voter *i*'s correctness probability $P(\mathbf{v}_i = s | s)$, over-valuation probability $P(\mathbf{v}_i > s | s)$, and under-valuation probability $P(\mathbf{v}_i < s | s)$. The absolute difference between the over- and under-valuation probability defines i's bias, i.e., tendency to over- or under-value. Morreau (2020) proves a 'grading-jury theorem': if, conditional on any state, the grades $\mathbf{v}_1, \mathbf{v}_2, \dots$ are (i) independent, (ii) homogeneously distributed, and (iii) sufficiently unbiased (i.e., with bias below the correctness

²⁵ i.e., each \mathbf{v}_i has a conditional expectation equal to \mathbf{s} .

²⁶More precisely, the assumptions (i)-(iii) guarantee state-conditional convergence of $F(\mathbf{v}_1, ..., \mathbf{v}_n)$ to **s** by applying the Law of Large Numbers state-by-state, which then implies unconditional convergence by taking the expectation.

²⁷More precisely, the expected collective decision value is asymptotically maximal, under various plausible (non-simple) standards of correctness, i.e., under various plausible definitions of decision value. For instance, decision value could be negative absolute distance to the truth: V(a, s) = -|a - s|. The simple standard of correctness (cf. Section 5.1) would be inappropriate here, because it treats all deviations from the truth – small or large – as equally bad. Indeed, for continuously distributed input variables $\mathbf{v}_1, \mathbf{v}_2, ...$, the mere probability of (exact) collective correctness is zero, however much the group is increased.

²⁸This follows from other statements of the Law of Large Numbers.

probability), then the correctness probability of the median evaluation increases in (odd) group size and converges to certainty.²⁹ The correctness probability can become very small, provided it exceeds the bias. Having small bias seems particularly difficult if the true value is very high or very low, since extreme values are easily under- or over-estimated.

5.5 Judgment aggregation over interconnected propositions

In judgment aggregation, a group forms yes/no judgments on several propositions (List and Pettit 2002; Dietrich 2007; Pigozzi 2016). One might judge whether economic growth will increase (proposition p), whether some climate goals will be missed (proposition q), and whether the former implies the latter (proposition $p \to q$). The three judgments are interconnected: one cannot consistently accept $p, p \to q$, and $\neg q$. Let alternatives be complete and consistent judgment sets; in the example, $\mathcal{A} = \{\{p, p \to q, q\}, \{\neg p, p \to q, q\}, ...\}$. The state **s** is the *correct* judgment set in $\mathcal{A} (= S)$, containing the true propositions. Someone's competence can vary across propositions. The propositions on which someone is competent, i.e., likely to judge the truth value correctly, form her *area of competence*.

Epistemically successful judgment-aggregation rules tend to rely on individual judgments where individuals are competent. How, in the example, should the group judge whether q, i.e., when should the collective judgment set contain q and when $\neg q$? The 'conclusion-based procedure' adopts the majority judgment on whether q. The 'premise-based procedure' adopts the logical implication of the majority judgments on whether p and whether $p \rightarrow q$.³⁰ Which of these two prominent procedures is more likely to find out the truth about q depends, roughly speaking, on whether individuals are primarily competent on q or primarily competent on p and $p \rightarrow q$. The situation gets complicated if individuals have different areas of competence.

Further complications arise if not only correctness on q, but also correctness on p and $p \to q$ matter. Technically, this refines the standard of correctness by which judgment sets in \mathcal{A} are evaluated, i.e., the epistemic objective.³¹ Being correct about p and $p \to q$ may matter in itself. Alternatively, judgments about p and

 $^{^{29}\}mathrm{This}$ holds conditional on any state, and thus unconditionally.

 $^{^{30}}$ It is set aside here what happens in the absence of any implication.

³¹A purely conclusion-oriented standard of correctness can be captured by a value function V assigning to a judgment-set/state pair (J, s) the value 1 or 0 depending on whether or not the judgment on the conclusion proposition q is correct, i.e., whether or not $q \in J \Leftrightarrow q \in s$. A refined standard of correctness that also cares about correctness on the premise propositions p and $p \to q$ can be captured by a value function V that is also sensitive to correctness on premise propositions. For instance, V(J, s) could be defined as $|J \cap s|$, the number of true judgments; this correctness standard cares equally about correctness on the conclusion and correctness about each premise proposition.

 $p \rightarrow q$ may count as reasons for the judgment about q, and it may matter to be right about q for correct reasons. Reasons or justifications matter in many real examples. For example, courts should not only make correct sentences (convict or acquit), but also justify them correctly in terms of facts and laws.

Jury theorems capturing these and other insights are developed in List (2005) and Bovens and Rabinowicz (2006).

5.6 Jury theorems for voters tracking individual facts

Ordinary jury theorems assume the classic paradigm of epistemic democracy: collective decisions should be correct, and individuals vote for what they believe to be correct. But collective correctness can emerge even if voters do not pursue it. For instance, decisions can serve the collective interest even if voters vote for what they believe serves their personal interest. This is the message of *jury theorems for voters tracking individual facts*. They follow a semi-epistemic paradigm of democracy: collective decisions should still be correct, but voters express judgments about individual facts.

Within this paradigm, what vote is correct depends on the individual. The rationale could be that votes express judgments about agent-indexed facts, for instance about which alternative is best *for oneself*, i.e., serves *one's own* interest. Obviously, correctness of judgments like 'option X is in my interest' is voter-dependent. An entirely different rationale for voter-dependent correctness is that correctness facts are subjective – a thesis sometimes advanced for normative facts. Either way, an individual might not know with certainty which alternative is individually correct (e.g., serves their own interest). Individual competence, in this framework, is the ability to track one's own correctness fact.

Which truth does *the group* track in its aggregate decision? The collective correctness standard combines the individual standards. Collective correctness could, for instance, be defined as correctness *for most individuals.*³² The natural procedure to track this collective truth is majority voting (assuming there are just two alternatives). Still, majority outcomes may be collectively incorrect, as individuals may cast individually incorrect votes. Yet, under some stylized assumptions, including broadly Condorcetian competence and independence assumptions, majority voting can generate growingly reliable and possibly asymptotically infallible outcomes (see theorems conjectured by Miller 1986 and List and Spiekermann 2016).

 $^{^{32}}$ Here an alternative is collectively correct if and only if the number of group members for whom it is individually correct is maximal. The collectively correct alternative need not be unique: just imagine two alternatives are each individually correct for exactly n/2 members, which can happen for even group size n.

But this standard of collective correctness can be questioned. It rests solely on how often an alternative is individually correct. Yet it might also matter how strongly someone's correctness standard (e.g., her welfare) is satisfied or violated. This becomes evident in concrete examples.³³ To account for degrees, let a correctness standard be given by a value function V, assigning a value or correctness degree to each alternative, formally $V : \mathcal{A} \to \mathbb{R}$. If V_1, V_2, \dots are the (unknown) value functions of the individuals 1, 2, ..., respectively, how valuable or correct are alternatives *collectively*? Under an additive approach, the collective value function is $V_1 + \cdots + V_n$. If individual value represents individual welfare or interest, then such additive collective value represents utilitarian collective welfare; but nothing hinges on a welfarist interpretation of value. An alternative in \mathcal{A} is now correct (individually or collectively) if it maximises the relevant value function. Here collective correctness no longer reduces to correctness for *most* individuals. Pivato's (2016) statistical utilitarianism theorem – initially a theorem about making approximately utilitarian-optimal decisions – can be interpreted as a jury theorem for individual correctness of this graded kind and collective correctness of additive kind.³⁴

Jury theorems with voters tracking individual facts depart from ordinary jury frameworks and classic epistemic democracy in two fundamental ways. First, different voters track different 'truths', against the epistemic conception of democracy. Second, the collectively correct alternative depends on group size n, hence is a moving target that changes as voters are added. Still, such theorems make a remarkable point: voting in pursuit of the individual good creates the collective good, as if through an invisible hand. Democracy finds the collective truth even if voters search for different, purely individual, truths. But does democracy find the collective truth more easily if voters express judged individual interest or judged collective interest? This is an open question for epistemic democracy.

 $^{^{33}}$ Consider a distributional choice between keeping the status-quo distribution and taking all goods from the poorest citizen and redistributing them equally between everyone else. The redistribution seems incorrect according to egalitarian and utilitarian standards; but it benefits all but one individual, and hence counts as collectively correct if collective correctness means correctness for *most* individuals (and if moreover individual correctness is given by personal interest). The problem is the neglect of *how much* someone benefits. Degrees matter.

³⁴Unconventionally, the inputs of aggregation are not alternatives $a_1, ..., a_n$, i.e., not individual estimates of what is individually best, but entire value functions $W_1, ..., W_n$, i.e., individual estimates of the true individual value functions $V_1, ..., V_n$, respectively. So, everyone submits their judgment of how individually valuable or correct each alternative is. Let the submitted estimates be aggregated additively; the result $W_1 + \cdots + W_n$ is an estimate of the true collective value function $V_1 + \cdots + V_n$. The group then chooses an alternative maximising $W_1 + \cdots + W_n$. Under some probabilistic assumptions about the quality and independence of individual estimates (and a finiteness assumption on \mathcal{A}), the probability that the collective choice is optimal, i.e., maximises true collective value, tends to one as n grows (Pivato 2016).

6 Background and selected applications

6.1 Jury theorems and democratic theory

Whether masses can be wise and should be entrusted with important decisions is a question with long philosophical pedigree. The Athenian democracy increased the decision body compared to its oligarchic competitors, demonstrating that such an enlargement can engender collective epistemic success (Ober 2008; 2013). Against this backdrop, the philosophical debate about the epistemic performance of democracy began. Aristotle, for instance, despite being generally sceptical of the rule by the people, concedes the possibility of group wisdom (Aristotle, *Politics*: 1281a11; Waldron 1995), though this reading is contested (Cammack 2013; Lane 2013).

While the epistemic advantages of drawing on many minds had been debated long before Condorcet, he is the first to develop a probabilistic framework akin to modern jury theorems and to focus explicitly on voting mechanisms. Condorcet's remarkable advance may also have rubbed off on his contemporaries. Jean-Jacques Rousseau's argument for popular rule was influenced by Condorcet's thought and resembles an informal statement of his theorem, according to an influential article by Grofman and Feld (1988) (but see Wyckoff 2011 for a critical take on that historical speculation).

While Condorcet's jury theorem largely falls into oblivion after his death, the mathematical take on aggregation which he had initiated remains. Francis Galton, for example, stumbles over useful data for testing the epistemic performance of belief aggregation when visiting a cattle fair in 1906. Galton observes a weight-judging competition: participants were asked to guess the slaughtering weight of an ox, submitting their guesses in writing. To Galton's surprise, the median guess was within 0.8% of the actual weight, beating by far the individual judgments. While Galton does not directly invoke a jury theorem, he notes that 'the middlemost estimate expresses the *vox populi*, every other estimate being condemned as too low or too high by a majority of the voters' (Galton 1907).

Condorcet is a founding father of social choice theory, not only because of his jury theorem. When Duncan Black (1958) famously rediscovers his work, his jury theorem receives new attention in political science and democratic theory (Grofman 1975 seems to have coined the term 'Condorcet Jury Theorem'). However, the initial reception is not friendly: Black (1958: 163) dismisses the idea of an independent standard of correctness in elections and John Rawls (1971 [1999]: 315) doubts that votes could be sufficiently independent. Brian Barry (1965: 205-6) is perhaps the first 20th-century political theorist to affirm the importance of the theorem for democratic theory.

Contemporary democratic theory remains divided on jury theorems. Broadly

speaking, the field distinguishes between instrumental and procedural arguments for democracy (Anderson 2009; Dworkin 1987-88; List and Goodin 2001). Recent years have seen increasing interest in instrumental, and especially epistemic arguments. But even among epistemic democrats (early proponents are Cohen 1986; Coleman and Ferejohn 1986; see Schwartzberg 2014 for an overview), the merits of jury theorems are controversial. Some epistemic democrats regard jury theorems as central tools for justifying democracy (e.g., List and Goodin 2001; Landemore 2013; Dietrich and Spiekermann 2013a; Goodin and Spiekermann 2018). Others disagree. Henry S. Richardson (2002), for example, rejects Condorcet's independence assumption. David Estlund instead rejects Condorcet's competence assumption (Estlund 2008: ch. 12) and dismisses Condorcet's Jury Theorem as providing 'too shaky a basis' (Estlund 2008, 223). Elizabeth Anderson also objects to Condorcet's implausible assumptions and conclusions (Anderson 2006).

Such debates in political philosophy were hampered by a focus on Condorcet's initial version of the theorem with its problematic assumptions and conclusions (see Section 2); these objections often lose weight for other jury theorems.

Grofman and Feld fan worries about deliberation by attributing to Rousseau the view that 'each voter is polled about his or her independently reached choice, without any group deliberation' (1988: 570). Unsurprisingly, since many democratic theorists see public deliberation as instrumentally or intrinsically valuable, the proposal to sacrifice deliberation for independence was not met with enthusiasm. Recent debates in democratic theory, however, have turned to conditional notions of independence which can better accommodate deliberation and fruitful exchange, thereby also avoiding the controversial infallibility conclusion (Ladha 1992; Dietrich and List 2004; Dietrich and Spiekermann 2013a,b; Goodin and Spiekermann 2018; cf. Sections 2 and 4.5).

Some epistemic democrats worry that an exaggerated focus on jury theorems (and aggregation more generally) suppresses other important epistemic mechanisms, such as democratic experimentation (Anderson 2006; Fuerstein 2008), individual learning as opposed to aggregation (Müller 2018), distributed search, deliberation, and learning. Certainly, jury theorems are only *one* building block; a comprehensive epistemic analysis of collective decision-making also needs models of such other mechanisms. Nevertheless, jury theorems can assist the epistemic analysis of institutions and help assess political representation, bicameralism, epistemic division of labour, political cue taking, the merits of diversity, and the dangers of excessive bias, to name just a few applications (Goodin and Spiekermann 2018). Jury theorems have also been used to analyse legal institutions, including the importance of legal precedent and, of course, the (in)correctness of jury verdicts (e.g., Vermeule 2009; Feddersen and Pesendorfer 1998; Coughlan 2000).

6.2 Jury theorems and social epistemology

Social epistemology can be roughly divided into epistemology of group s and epistemology *in*groups (Dietrich and Spiekermann forthcoming). Jury theorems matter to the former, and indirectly also to the latter.

Epistemology of groups analyses how groups can be knowers and how they perform as knowers. A lively debate among applied social epistemologists explores how (and how well) different institutions acquire knowledge. Alvin Goldman's seminal *Knowledge in a Social World* (1999), for example, draws on Condorcet's Jury Theorem in his social-epistemological analysis of democracy. The debate extends to courts (Spiekermann and Goodin 2012; Sunstein 2009; Vermeule 2009), to the internet (Masterton, Olsson and Angere 2016), and even to knowledge of the scientific community and its perception by the wider public (Hahn, Harris and Corner 2016). Stretching the boundaries of social epistemology further, one could even say that animal decision-making can be more or less conducive to group knowledge (e.g., Conradt and List 2009).

Jury theorems matter more indirectly to social epistemology *in* groups, especially belief revision under peer disagreement. How ought one revise one's beliefs when learning the beliefs of others? Many have argued that the number of peers, their competence and their independence all matter (Elga 2010; Kelly 2011; Lackey 2013). Barnett (2019), for example, discusses peer disagreement with a 'group representative' who adopts the group's majority view. Jury theorems can help approach such questions systematically.

The jury literature also offers analyses of opinion independence in social contexts, distinguishing between different notions of independence: *causal* independence, *probabilistic* ('statistical', 'evidential') independence, and *reasoning* independence (a form of autonomy of agents); see Dietrich and Spiekermann (2013b), Lackey (2013), and Section 2.2. None of these notions entails another, but they are often unhelpfully mixed up.

6.3 Jury theorems and epistemic social choice theory

If viewed mathematically, jury theorems belong to epistemic social choice theory, a branch of social choice theory (see List 2013). Social choice theory studies aggregation procedures in groups. Like democratic theory, social choice theory has a procedural and an epistemic branch. Procedural social choice theorists want procedures to be intrinsically 'democratic' or 'fair' to individuals; they formulate axiomatic requirements on procedures capturing normative principles (such as anonymity or Paretianism) and ask which procedure(s) satisfy them.³⁵ Epistemic social choice theorists instead want procedures to 'track the truth'; they formulate some epistemic criterion on outcomes (e.g., high probability of correctness) and analyse procedures in terms of how well their outcomes satisfy the criterion.

Jury theorems fall into the epistemic camp. What distinguishes them within epistemic social choice is the focus on decision-bodysize rather than more conventional procedural parameters, such as the acceptance threshold in a supermajority rule or the individual weights in a weighted voting rule. Jury theorems seek to optimize group size *given* a procedure (e.g., given majority rule), rather than optimizing the procedure *given* a group size. Procedure optimisation is sometimes possible analytically. The generic finding is that, under a classical independence assumption, the epistemically optimal procedure in binary choices is a weighted (simple, super- or sub-) majority rule in which a voter's weight is a well-defined increasing function of competence (which becomes negative for competence below $\frac{1}{2}$); see e.g., Nitzan and Paroush (1982), Ben-Yashar and Nitzan (1997), Dietrich (2006), and Pivato (2013). Simple majority rule can be far from optimal; it is optimal if voters are *equally* competent and options are initially *equally* likely correct.

7 Bibliography

- Anderson, Elizabeth, 2006, "The Epistemology of Democracy," Episteme, 3 (1–2): 8–22.
 - ——, 2009, "Democracy: Instrumental vs. Non-Instrumental Value," in Thomas Christiano and John Christman (eds.), Contemporary Debates in Political Philosophy, Chichester: Wiley Blackwell. pp. 213–227.
- Aristotle, 1988, The Politics, Cambridge Texts in the History of Political Thought, Cambridge: Cambridge University Press.
- Austen-Smith, D., and J. S Banks, 1996, "Information Aggregation, Rationality, and the Condorcet Jury Theorem," American Political Science Review, 90 (1): 34–45.
- Barnett, Zach, 2019, "Belief Dependence: How Do the Numbers Count?," Philosophical Studies, 176 (2): 297–319.
- Barry, Brian, 1965, Political Argument. Routledge Revivals. Abingdon, Oxon; New York, NY: Routledge.

³⁵The class of solutions may be empty – the case of impossibility theorems like Arrow's Theorem. Here axioms are mutually inconsistent and must be relaxed. The class may be singleton – the ideal case. Here a single procedure emerges as admissible, and the theorist has finished her work. Or the class may contain many procedures – the underdetermination case. Here stronger axioms are needed for choosing a procedure.

- Ben-Yashar, Ruth C., and Shmuel I. Nitzan, 1997. "The Optimal Decision Rule for Fixed-Size Committees in Dichotomous Choice Situations: The General Result," International Economic Review, 38 (1): 175.
- Berend, Daniel, and Jacob Paroush, 1998, "When Is Condorcet's Jury Theorem Valid?," Social Choice and Welfare, 15: 481–488.
- Berend, Daniel, and Luba Sapir, 2005, "Monotonicity in Condorcet Jury Theorem," Social Choice and Welfare, 24 (1): 83–92.
- Berg, Sven, 1993, "Condorcet's Jury Theorem, Dependency among Voters," Social Choice and Welfare 10: 87–95.
- Black, Duncan, 1958, The Theory of Committees and Elections, Cambridge: Cambridge University Press.
- Boland, Philip J., 1989, "Majority Systems and the Condorcet Jury Theorem," Journal of the Royal Statistical Society. Series D (The Statistician), 38 (3): 181–189.
- Boland, Philip J, Frank Proschan, and Y L Tong, 1989, "Modelling Dependence in Simple and Indirect Majority Systems," Journal of Applied Probability 26 (1): 81–88.
- Bovens, Luc, and Wlodek Rabinowicz, 2006, "Democratic Answers to Complex Questions - an Epistemic Perpective," Synthese, 150: 131–153.
- Bozbay, Irem, Franz Dietrich, and Hans Peters, 2014, "Judgment Aggregation in Search for the Truth," Games and Economic Behavior, 87 (September): 571–90.
- Brennan, Jason, 2018, "Does the Demographic Objection to Epistocracy Succeed?," Res Publica, 24 (1): 53–71.
- Brennan, Geoffrey and Loren E. Lomasky, 1993, "Democracy and Decision: the pure theory of electoral preference," Cambridge: Cambridge University Press.
- Cammack, Daniela, 2013, "Aristotle on the Virtue of the Multitude," Political Theory, 41 (2): 175–202.
- Cohen, Joshua, 1986, "An Epistemic Conception of Democracy," Ethics, 97 (1): 26–38.
- Coleman, Jules, and John Ferejohn, 1986, "Democracy and Social Choice," Ethics, 97 (1): 6–25.
- 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.
- Conradt, Larissa, and Christian List, 2009, "group Decisions in Humans and Animals: A Survey," Philosophical Transactions of the Royal Society B: Biological Sciences, 364 (1518): 719–742.
- Coughlan, Peter J., 2000, "In Defense of Unanimous Jury Verdicts: Mistrials, Communication, and Strategic Voting," American Political Science Review, 94 (2): 375–93.

- Dietrich, Franz, 2006, "General Representation of Epistemically Optimal Procedures," Social Choice and Welfare, 26 (2): 263–83.
 - —, 2007, "A Generalised Model of Judgment Aggregation," Social Choice and Welfare, 28 (4): 529–565.
 - —, 2008, "The Premises of Condorcet's Jury Theorem Are Not Simultaneously Justified," Episteme, 5 (1): 56–73.
- Dietrich, Franz, and Christian List, 2004, "A Model of Jury Decisions Where All Jurors Have the Same Evidence," Synthese, 142: 175–202.
- Dietrich, Franz, and Kai Spiekermann, forthcoming, "Social Epistemology," in Markus Knauff and Wolfgang Spohn (eds), The Handbook of Rationality, Cambridge (MA): MIT Press.
 - —, 2013a, "Epistemic Democracy with Defensible Premises," Economics and Philosophy, 29 (1): 87–120.
 - —, 2013b, "Independent Opinions?," Mind, 122 (487): 655–685.
- ——, 2020, "Jury Theorems," in Miranda Fricker, Peter J. Graham, Peter Henderson, and Nikolai J.L.L. Pedersen (eds.), The Routledge Handbook of Social Epistemology, New York and London: Routledge.
- Dryzek, John, and Christian List, 2003, "Social Choice Theory and Deliberative Democracy: A Reconciliation," British Journal of Political Science, 33 (1): 1–28.
- Dworkin, Ronald, 1987, "What Is Equality Part 4: Political Equality MarshallP. Madison Lecture," University of San Francisco Law Review, 22: 1–30.
- Elga, Adam, 2010, "How to Disagree about How to Disagree," in Richard Feldman and Ted A Warfield (eds.), Disagreement, Oxford: Oxford University Press, pp. 175–86.
- Elster, Jon, 2012, "The Optimal Design of a Constituent Assembly," in Hélène Landemore and Jon Elster (eds.), Collective Wisdom: Principles and Mechanisms, Cambridge: Cambridge University Press, pp. 148–172.
- Estlund, David M., 2008, Democratic Authority: A Philosophical Framework, Princeton: Princeton University Press.
- Feddersen, Timothy, and Wolfgang Pesendorfer, 1998, "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," American Political Science Review, 92 (1): 23–35.
 - —, 1999, "Elections, Information Aggregation, and Strategic Voting," Proceedings of the National Academy of Sciences of the United States of America, 96 (19): 10572–74.
- Fuerstein, Michael, 2008, "Epistemic Democracy and the Social Character of Knowledge," Episteme, 5 (1): 74–93.
- Galton, Francis, 1907, "Vox Populi," Nature, 75 (1949): 450–51.
- Gaus, Gerald. 2011, "On Seeking the Truth (Whatever That Is) through Democracy: Estlund's Case for the Qualified Epistemic Claim," Ethics, 121 (2):

270 - 300.

- Gerardi, Dino and Leeat Yariv, 2007, "Deliberative Voting," Journal of Economic Theory, 134 (1): 317-338.
- Goldman, Alvin I., 1999, Knowledge in a Social World, Oxford: Clarendon.
- Goodin, Robert E, and Kai Spiekermann, 2018, An Epistemic Theory of Democracy, Oxford: Oxford University Press.
- Grofman, Bernard, 1975, "A Comment on 'Democratic Theory: A Preliminary Mathematical Model," Public Choice, 21: 99–103.
- Grofman, Bernard, and Scott L Feld, 1988, "Rousseau's General Will: A Condorcetian Perspective," American Political Science Review, 82 (2): 567–576.
- Grofman, Bernard, Guillermo Owen, and Scott L Feld, 1983, "Thirteen Theorems in Search of the Truth," Theory and Decision, 15: 261–278.
- Gunn, Paul, 2019, "Against Epistocracy," Critical Review, 31 (1): 26-82.
- Hahn, Ulrike, Adam J. L. Harris, and Adam Corner, 2016, "Public Reception of Climate Science: Coherence, Reliability, and Independence," Topics in Cognitive Science, 8 (1): 180–95.
- Hitchcock, Christopher, and Miklós Rédei, 2020, "Reichenbach's Common Cause Principle," in Edward N. Zalta (ed.), The Stanford Encyclopedia of Philosophy (Spring 2020 edition). URL= https://plato.stanford.edu/archives/spr2020/entries/p Rpcc/:>.
- Hong, Lu, and Scott Page, 2009, "Interpreted and Generated Signals," Journal of Economic Theory, 144 (5): 2174–2196.
- Hong, Lu, and Scott E Page, 2012, "Some Microfoundations of Collective Wisdom," in Hélène Landemore and Jon Elster (eds.), Collective Wisdom: Principles and Mechanisms, Cambridge: Cambridge University Press.
- Kaniovski, Serguei, 2010, "Aggregation of Correlated Votes and Condorcet's Jury Theorem," Theory and Decision, 69 (3): 453–468.
- Karotkin, Drora, and Jacob Paroush, 2003, "Optimum Committee Size: Quality-versus-Quantity Dilemma," Social Choice and Welfare, 20 (3): 429–441.
- Kelly, Thomas, 2011, "Peer Disagreement and Higher Order Evidence," in Alvin Goldman and Dennis Whitcomb (eds.), Social Epistemology: Essential Readings, New York: Oxford University Press, pp. 183–217.
- Kolmogorov, Andrey N, 1930, "Sur La Loi Forte Des Grands Nombres," Comptes Rendus de l'Académie Des Sciences, 191: 910–912.
- Lackey, Jennifer, 2013, "Disagreement and Belief Dependence Why Numbers Matter," in David Christensen and Jennifer Lackey (eds.), The Epistemology of Disagreement: New Essays, Oxford: Oxford University Press. pp. 243–68.
- Ladha, Krishna K., 1992, "The Condorcet Jury Theorem, Free Speech, and Correlated Votes," American Journal of Political Science, 36 (3): 617–634.
 - -, 1993, "Condorcet's Jury Theorem in Light of de Finetti's Theorem,"

Social Choice and Welfare, 10 (1): 69–85.

- ——, 1995, "Information Pooling through Majority-Rule Voting: Condorcet's Jury Theorem with Correlated Votes," Journal of Economic Behavior & Organization, 26 (3): 353–372.
- Landemore, Hélène, 2013, Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many, Princeton, N.J.: Princeton University Press.
- Lane, Melissa, 2013, "Claims to Rule: The Case of the Multitude," in Marguerite Deslauriers and Pierre Destrée (eds.), The Cambridge Companion to Aristotle's Politics, Cambridge: Cambridge University Press. pp. 247–74.
- Lippert-Rasmussen, Kasper, 2012, "Estlund on Epistocracy: A Critique," Res Publica, 18 (3): 241–258.
- List, Christian, 2005, "group Knowledge and group Rationality: A Judgment Aggregation Perspective," Episteme, 2 (1): 25–38.
 - ——, 2013, "Social Choice Theory," in Edward N. Zalta (ed.), The Stanford Encyclopedia of Philosophy (Winter 2013), URL= https://plato.stanford.edu/archives/win20 choice/>.
- List, Christian, and Robert E. Goodin, 2001, "Epistemic Democracy: Generalizing the Condorcet Jury Theorem," Journal of Political Philosophy, 9 (3): 277–306.
- List, Christian, and Philip Pettit, 2002, "Aggregating Sets of Judgments: An Impossibility Result," Economics and Philosophy, 18 (1): 89–110.
- List, Christian, and Kai Spiekermann, 2016, "The Condorcet Jury Theorem and Voter-Specific Truth," in Hilary Kornblith and Brian McLaughlin (eds.), Goldman and His Critics, Chichester: John Wiley & Sons, pp. 219–231.
- Masterton, George, Erik J. Olsson, and Staffan Angere, 2016, "Linking as Voting: How the Condorcet Jury Theorem in Political Science Is Relevant to Webometrics," Scientometrics, 106 (3): 945–66.
- Miller, Nicholas R., 1986, "Information, Electorates, and Democracy: Some Extentions and Interpretations of the Condorcet Jury Theorem," in Bernard Grofman and Guillermo Owen (eds.), Information Pooling and group Decision Making, Greenwich, Conn.: JAI Press, pp. 173–192.
- Morreau, Michael, 2020, "Democracy without Enlightenment: A Jury Theorem for Evaluative Voting," Journal of Political Philosophy, Online first.
- Muirhead, Russell, 2014, "The Politics of Getting It Right," Critical Review, 26 (1–2): 115–28.
- Müller, Julian F., 2018, "Epistemic Democracy: Beyond Knowledge Exploitation," Philosophical Studies, 175 (5): 1267–88.
- Nitzan, Shmuel, and Jacob Paroush. 1982. "Optimal Decision Rules in Uncertain Dichotomous Choice Situations." International Economic Review, 23 (2): 289–297.
- Ober, Josiah, 2008, Democracy and Knowledge: Innovation and Learning in

Classical Athens, Princeton, Princeton, N.J.: Princeton University Press.
—, 2013, "Democracy's Wisdom: An Aristotelian Middle Way for Collective Judgment," American Political Science Review, 107 (1): 104–22.

- Owen, Guillermo, Bernard Grofman, and Scott L Feld, 1989, "Proving a Distribution-Free Generalization of the Condorcet Jury Theorem," Mathematical Social Sciences, 17: 1–16.
- Paroush, Jacob, 1998, "Stay Away from Fair Coins: A Condorcet Jury Theorem," Social Choice and Welfare, 15: 15–20.
- Pearl, Judea, 2000, Causality: Models, Reasoning and Inference, Cambridge: Cambridge University Press.
- Peleg, Bezalel, and Shmuel Zamir, 2012, "Extending the Condorcet Jury Theorem to a General Dependent Jury," Social Choice and Welfare, 39 (1): 91–125.
- Pigozzi, Gabriella, 2016, "Belief Merging and Judgment Aggregation," in Edward N. Zalta (ed.). The Stanford Encyclopedia of Philosophy (Winter 2016), Available at URL = <https://plato.stanford.edu/archives/win2016/entries/beliefmerging/>.
- Pivato, Marcus, 2013, "Voting Rules as Statistical Estimators," Social Choice and Welfare, 40 (2): 581–630.
 - —, 2016, "Statistical Utilitarianism," in Maria Gallego and Norman Schofield (eds.), The Political Economy of Social Choices, Studies in Political Economy. Cham: Springer International Publishing. pp. 187–204.
 - —, 2017, "Epistemic Democracy with Correlated Voters," Journal of Mathematical Economics, 72: 51–69.
- Rawls, John, 1971, A Theory of Justice, Rev. ed. 1999, Oxford: Oxford University Press.
- Reichenbach, Hans, 1956, The Direction of Time, Berkeley: University of California Press.
- Richardson, Henry S., 2002, Democratic Autonomy: Public Reasoning about the Ends of Policy, New York: Oxford University Press.
- Romeijn, Jan-Willem, and David Atkinson, 2011, "Learning Juror Competence: A Generalized Condorcet Jury Theorem," Politics, Philosophy & Economics, 10 (3): 237–62.
- Schuessler, Alexander A., 2000, "Expressive Voting", Rationality and Society, 12 (1): 87-119.
- Schwartzberg, Melissa, 2014, "Epistemic Democracy and Its Challenges," Annual Review of Political Science, 18 (1).
- Spiekermann, Kai, and Robert E. Goodin, 2012, "Courts of Many Minds," British Journal of Political Science, 42 (3): 555–71.
- Sunstein, Cass R., 2009, A Constitution of Many Minds: Why the Founding Document Doesn't Mean What It Meant Before, Princeton, N.J: Princeton University Press.

- Sunstein, Cass R., and Reid Hastie, 2014, Wiser: Getting Beyond group think to Make groups Smarter, Boston, Massachusetts: Harvard Business Review Press.
- Vermeule, Adrian, 2009, "Many Minds Arguments in Legal Theory," Journal of Legal Analysis, 1 (1): 1–45.
- Waldron, Jeremy, 1995, "The Wisdom of the Multitude: Some Reflections on Book 3, Chapter 11 of Aristotle's Politics," Political Theory, 23 (4): 563–84.
- Wyckoff, Jason, 2011, "Rousseau's General Will and the Condorcet Jury Theorem," History of Political Thought, 32 (1): 49–62.