Social Epistemology

Franz Dietrich, Kai Spiekermann

Social Epistemology

Franz Dietrich[1] and Kai Spiekermann[2]

May 2019

Abstract

Social epistemology studies knowledge in social contexts. Knowledge is 'social' when its holder communicates with or learns from others (Epistemology *in* groups), or when its holder is a group as a whole, literally or metaphorically (Epistemology *of* groups). Group knowledge can emerge explicitly, through aggregation procedures like voting, or implicitly, through institutions like deliberation or prediction markets. In the truth-tracking paradigm, group beliefs aim at truth, and group decisions at 'correctness' – in virtue of external facts that are empirical or normative, real or constructed, universal or relativistic, etc. Procedures and institutions are evaluated by epistemic performance: Are they truth-conducive? Do groups become 'wiser' than their members? We review several procedures and institutions, discussing epistemic successes and failures. *Jury theorems* provide formal arguments for epistemic success. Some jury theorems misleadingly conclude that 'huge groups are infallible' – an artefact of inappropriate premises. Others have defensible premises, and still conclude that groups outperform individuals, without being infallible.

Social Epistemology is the branch of Epistemology which studies knowledge in social contexts. In this review, we first set the stage by introducing clarifications, distinctions, and applications (Section 1). We then discuss formal procedures (Section 2) and informal institutional arrangements (Section 3) which generate collective beliefs or decisions that 'track the truth', or instead lead to epistemic failures (Section 4).

1 Scope and problems of social epistemology

1.1 Epistemology in groups versus epistemology of groups

A first distinction pertains to the knowledge holder, who is either a group member or the group as a whole. Epistemology *in* groups might be regarded

[1] Paris School of Economics & CNRS
[2] London School of Economics

as a species of individual epistemology, though its focus lies in how individual knowledge depends on social inputs; typical questions are rational responses to testimony and peer disagreement, and 'irrational' or sub-personal responses such as belief contagion and belief change through mass media. By contrast, epistemology *of* groups ascribes knowledge to the group as a whole; let us talk here of *social* knowledge or belief (List and Pettit 2011; cf. Goldman and Blanchard 2018, Schmitt 1999). The meaning and status of social knowledge or belief is controversial. The group in question can be more or less structured and cohesive, with implications for whether the group can qualify as an *agent*, and for what it means for the group to believe or know something, ranging from literal to purely metaphorical meanings. For instance, a committee, firm or state (if treated as an agent) can believe something literally, while a prediction market or the random group participating in a survey can believe something only in some metaphorical sense. List (2014) distinguishes between three types of social beliefs: *aggregate* beliefs are mere summaries of the beliefs of group members; *common* beliefs are ultimately individual beliefs held by each group member, with a common awareness of these beliefs; and *corporate* beliefs are beliefs of the group in a literal sense, which presupposes that the group qualifies as an agent. Theoretical distinctions aside, it is evident that we routinely invoke social knowledge or beliefs, for instance when saying that *we* know something or that prediction markets "knew" in advance that Obama would become president of the United States. In a formal analysis, our distinction between epistemology *in* groups and *of* groups is sharp: the former is studied using belief revision models, the latter using aggregation models. In this review, we focus on epistemology *of* groups. We refer to Rott [to be added] for individual belief revision in groups.

1.2 Social knowledge versus social decisions

Narrowly construed, social epistemology addresses social knowledge or belief, not social action: it addresses how a legal court learns whether the defendant is guilty, not how it decides whether to convict him; how a community of physicists discovers a law, not how it decides whether to perform an experiment; etc. Broadly construed, however, social epistemology also addresses social decisions, like courts passing sentences, scientists choosing an experiment, parliaments passing laws, governments setting goals, commissions working out ethical standards, etc. For a social decision to qualify as an epistemic matter, it must have the property of being (in some relevant sense) *correct or incorrect* in virtue of some fact or truth. A court sentence may well be an epistemic matter, being correct or incorrect depending on whether the crime has been committed

and merits the sentence. Selecting Oscar winners *might* be regarded as an epistemic matter. One might adopt this inclusive notion of social epistemology because correct decisions often are or can be rationalized as the result of true beliefs.[3]

1.3 The truth to be tracked

One can distinguish between three different types of facts that make social beliefs true or false, or social decisions correct or incorrect: logical, empirical, and normative facts. For example:

- The mathematical community forms mathematical beliefs. Correctness depends on logical facts.
- The monetary policy committee of a central bank predicts whether inflation will rise. Correctness depends on empirical facts.
- A group of doctors decides whether to shift funds from medical research to patient treatment. Correctness depends on normative and empirical facts.

Social epistemology with normative facts is controversial, since both the existence and the nature of normative facts is debatable. Some would therefore prefer a proceduralist justification of social decisions, at least for decisions of a normative kind. But this conclusion may be premature, since truth-tracking is meaningful regardless of whether normative facts are real or constructed (e.g., socially constructed); universal or relativistic (as in cultural relativism); or natural or non-natural. The fact should, however, be suitably stable and procedure-independent; otherwise there is no stable truth to be tracked.

---

[3] While the content of a social belief is true or false, *holding* it is correct or incorrect, depending on the content's truth value. If groups are construed as intentional agents, they also hold other attitudes like desires, preferences, or intentions. This immediately raises the question of whether holding such other attitudes is also 'correct' or 'incorrect' in some relevant sense (possibly just for certain attitudes, e.g., for Scanlon's judgment-sensitive attitudes; 1998, p. 20). The correctness to hold, say, an intention might depend on whether its content is morally permissible (rather than true). For instance, it might be correct for a firm to intend to maximize shareholder value if maximizing it is morally permissible. Like social decisions with a correctness standard, social attitudes with a correctness standard *might* be regarded as a matter of social epistemology (broadly construed), especially where attitudes can be rationalized by true beliefs about relevant facts. Whether or not one adopts this inclusive view, some of the technical machinery of social epistemology – like jury theorems – can be used to study when group attitudes become correct.

1.4 Social knowledge, justification, and reliability

Most well-known problems about individual knowledge carry over to social knowledge. For instance, it is controversial whether knowledge is justified true belief, a thesis threatened by so-called Gettier cases, in which a subject (for us: a group) forms justified true beliefs out of sheer luck (Gettier 1963). It is also controversial whether justification to hold a belief comes from having sufficient evidence or rather from using a reliable procedure (see Goldman and Beddor 2016 on evidentialism versus reliabilism). Despite their familiarity, such 'knowledge problems' can take a distinctive form for social rather than individual knowledge, notably because groups use different procedures than individuals to acquire beliefs (e.g., voting procedures). Nonetheless, we shall set aside the question of what exactly social knowledge means and zoom in on social *beliefs* and the question of which institutions, social arrangements and aggregation procedures *track the truth*, i.e., produce true social beliefs or correct social decisions (Goldman 1999).

1.5 Social epistemology applied to epistemic democracy

Social epistemology is partly a positive theory of which institutions and procedures *do* track the truth. By contrast, the epistemic theory of democracy is a thoroughly normative theory, arguing that (democratic) institutions or procedures *should* track the truth to be legitimate. The rival normative theory, proceduralism, maintains that (democratic) institutions or procedures should not track any independent truth, but be procedurally fair, which often means that voters should have an equal say and alternatives should get an equal chance. The difference is fundamental: if there is an Oracle of Delphi that always tells the truth, then the procedure of blindly implementing what the oracle recommends for society is epistemically good (it tracks the truth), but procedurally bad (it is totally undemocratic). The debate between proponents of epistemic and procedural theories of democracy is ongoing. For epistemic theories of democracy, social epistemology is a highly relevant enterprise (Estlund 2008). For a hybrid position, see Peter (2007).

2 Aggregation procedures to track the truth

We now discuss prominent procedures generating social beliefs (Section 2.1) or truth-tracking decisions (Section 2.2), and then sketch how *jury theorems* help to establish the truth-conduciveness of such procedures (Section 2.3).

2.1 Forming social beliefs

Aggregating beliefs on a single proposition. A group of individuals is interested in whether a given proposition *p* is true. For instance, an economic panel is interested in whether inflation will rise, or a jury in whether the defendant has committed the crime. Each group member holds a belief about *p*, in the binary form of 'yes' or 'no'. An aggregation procedure takes the individual beliefs about *p* as input and returns a social belief about *p* as output. We allow social abstentions, so that society says neither yes nor no to *p* (but for simplicity we exclude individual abstentions). We now give examples. *Majority rule*: society believes what the majority believes (and abstains in case of a tie). *Asymmetric supermajority rules*: these make social belief of *p* harder to achieve than social disbelief, by requiring a supermajority support for a social 'yes' while otherwise opting for a social 'no'. *Symmetric supermajority rules*: these make social 'yes' and 'no' equally hard to achieve, by requiring the same supermajority support for each. This often results in social abstention, especially for high thresholds. Extreme cases of (asymmetric or symmetric) supermajority rules are (asymmetric or symmetric) unanimity rules, in which the supermajority threshold is a unanimity threshold. Later, in section 2.3, we analyze how likely it is that resulting social beliefs track the truth.

Aggregating beliefs about multiple propositions: judgment aggregation. Many groups are interested simultaneously in the truth of different propositions, something referred to as a *judgment-aggregation* problem (e.g., List and Pettit 2002, Dietrich 2007). For instance, a court may need collective beliefs about three propositions: the defendant has committed an act (*p*); such an act breaks the law (*q*); the defendant has broken the law (*r*). A judgment-aggregation problem does not simply reduce to several belief-aggregation problems about a single proposition each, because the propositions in question are typically logically interconnected. If in our court example the court comes to believe *p* and *q*, then it must believe *r*. Disbelieving *r* would be logically inconsistent. Abstaining on *r* (neither believing nor disbelieving *r*) would be deductively unclosed. The trouble is that voting on each relevant proposition in isolation often generates inconsistent and deductively unclosed social beliefs: even if each member of court holds consistent and complete (and thus deductively closed) beliefs, it may happen that a majority believes *p*, another majority believes *q*, and yet a majority disbelieves *r*. This phenomenon is referred to as the doctrinal paradox (Kornhauser and Sager 1986) or the discursive dilemma (Pettit 2001). The discursive dilemma generalizes far beyond propositionwise majority voting. General impossibility theorems establish that there exist *no* propositionwise judgment-aggregation procedures which are well-behaved in some sense, as soon as propositions are sufficiently interconnected; see

in particular List and Pettit's (2002) theorem and the Arrow-type theorem for judgment aggregation (Dietrich and List 2007 and Dokow and Holzman 2010, building on Nehring and Puppe 2010). So, whether a proposition is socially believed must depend not only on the individual beliefs about *this* proposition, but also on the individual beliefs about other propositions. This holistic nature of social beliefs makes social beliefs a more interesting and less transparent concept. The most famous holistic procedure to form social beliefs is the *premise-based procedure* (e.g., Pettit 2001, Dietrich and Mongin 2010). In our court example, this procedure determines the social belief on the propositions *p* and *q* (the *premise propositions*) through a majority vote on each of these propositions, while logically deducing the social belief about *r* (the *conclusion proposition*). So *r* is socially believed if and only if *p* and *q* are both socially believed. This is holistic since the social belief about *r* is no longer determined only by the individual beliefs about *r, potentially* overruling a majority belief about *r*. But the premise-based procedure comes with its own problems; for instance, it presupposes we can prioritize certain (premise) propositions over other (conclusion) propositions. Rival holistic procedures are distance-based rules (e.g., Miller and Osherson 2008), sequential rules (e.g., List 2004, Dietrich 2015), relevance-based rules (Dietrich 2015) and 'approximate majoritarian' rules (Nehring, Pivato and Puppe 2014).

Probabilistic opinion pooling. If we adopt the Bayesian paradigm, social beliefs should come in degrees rather than in binary yes/no form. What probabilities should society assign to propositions, given the probabilities assigned by the individuals? This is the so-called *opinion-pooling* problem – the probabilistic counterpart of the judgment-aggregation problem (for reviews, see Genest and Zidek 1986; Dietrich and List 2016). Once beliefs are probabilistic rather than binary, coherence of beliefs consists in respecting principles of probability theory like additivity, rather than logical principles like logical consistency and deductive closure. The picture reverses entirely: propositionwise ('local') aggregation of beliefs no longer runs into trouble. The social probability of any proposition can simply be the average (or a weighted average) of the individual probabilities, which guarantees coherent social beliefs as long as individual beliefs are coherent. Such 'linear averaging' procedures have been characterized axiomatically by McConway (1981), Wagner (1982), Mongin (1995), and Dietrich and List (2017). Although social beliefs produced by linear averaging are coherent, they are not fully Bayesian. Bayesianism is usually taken to require two things: holding beliefs in probabilistic form ("probabilism"), and revising beliefs through Bayes' rule ("conditionalization"). Social beliefs generated by linear averaging violate

the second requirement: they fail to be revised via Bayes' rule. Why? Assume some proposition is learnt, so that every individual revises his or her beliefs (via Bayes' rule, assuming that individuals are Bayesian). If at any time the current social beliefs are the average of the current individual beliefs, then the post-information social beliefs usually differ from the initial social beliefs updated via Bayes' rule. In short, this social belief revision is non-Bayesian. Opinion pooling through *geometric* rather than linear averaging repairs this flaw, i.e., produces dynamically rational social beliefs. In fact, geometric opinion pooling is the *only* way to guarantee 'Bayesian' group beliefs of a suitably well-behaved sort, as has only recently been shown (Russell et al. 2015, Dietrich forth.).

Aggregating evaluations – absolute versus ordinal approach. Certain objects – e.g., wines, political candidates, holiday destinations – must be evaluated in terms of some criterion, e.g., moral value, well-being, aesthetic value, or size. Let this be an epistemic problem: evaluations express *beliefs* about facts of some kind – which is plausible for size evaluations and more controversial for other evaluations. A first question is whether value is measured in absolute or ordinal terms. Consider the set X of objects evaluated. *Ordinal* evaluations are captured by a binary relation $\succcurlyeq$ on X, where '$x \succcurlyeq y$' means that x is at least as valuable as y (w.r.t. the relevant criterion)'. *Absolute* evaluations are captured by a function assigning to each object in X a value from a set V of possible values, e.g., from {very good, good, ...} (moral evaluation), or {beautiful, ugly, …} (aesthetic evaluation), or {large, medium, …} (size evaluation), or a set of numbers (numerical evaluations, e.g., of size). What should be the *social* evaluations of objects, given the evaluations by its members?

- In the ordinal case, this problem is structurally the notorious preference-aggregation problem, reinterpreting preference relations as value-judgment relations. One of many proposals is to use pairwise majority voting: object x is socially ranked over object y if and only if more individuals rank x over y than y over x. This procedure can lead to social cycles: some object x is majority-ranked over another y, which is majority-ranked over another z, which is majority-ranked over x, Condorcet's *voting paradox*. It generalizes into Arrow's impossibility theorem (Arrow 1963): very roughly, not just pairwise majority voting, but *any* pairwise aggregation procedure is flawed.
- The picture is brighter when aggregating *absolute* evaluations – an aggregation problem called the *social grading problem* by Balinski and Laraki (2011). In case of just two possible values – 'approved' and 'non-approved' – the most natural procedure is 'approval voting': an object is

socially approved if the number of individuals approving it is at least as high as for each other object, and socially non-approved otherwise (Brams and Fishburn 2007). Balinski and Laraki generalize this procedure to any set of values linearly ranked from 'highest' to 'lowest', e.g., {good, medium, bad}, where 'good' ranks above 'medium' which ranks above 'bad'. To socially evaluate an object *x*, first order the individuals such that the first individual evaluates *x* at least as highly as the second; the second at least as highly as the third; etc. The social value of *x* is the value assigned to *x* by the middle individual in that order (assuming an odd number of individuals for simplicity). So, if in a wine evaluation among three judges some wine receives the evaluations (exquisite, exquisite, drinkable), then this wine is socially evaluated as exquisite. This so-called *majority-judgment rule* (one might have called it the *median rule*) has several appealing features (see Balinski and Laraki 2011), but remains controversial within social choice theory with its ordinalist tradition.

The aggregation of evaluations can be regarded as a special case of judgment aggregation, namely over either ordinal ranking propositions of type '*x* ranks over *y*' (for options *x* and *y* in *X*), or absolute value propositions of type '*x* has value *v*' (for options *x* in *X* and values *v* in *V*).

2.2 Making social decisions

The procedures discussed above generate social beliefs of different sorts. We now turn to procedures for making social decisions. To stay within social epistemology (broadly construed), we assume the decisions in question track some truth, i.e., can be 'correct' or 'incorrect' in some procedure-independent sense (see Sections 1.2, 1.3 and 1.5). For example, a group might have to choose a member from a set *K* of social alternatives. *K* could contain just two alternatives, as when a court decides whether to convict or acquit the defendant. *K* could instead contain many alternatives, as when a court decides between different sentences. In so-called *plurality rule*, each individual votes for exactly one alternative in *K*, and society chooses the alternative receiving the highest number of votes (or one such alternative in case of a tie). Plurality rule reduces to simple majority rule in the two-alternative case. In the many-alternative case, plurality rule can lead to problems: Suppose alternative *k** in *K* is the 'correct' alternative. Plurality rule will normally fail to select *k** if *K* contains many alternatives similar to *k** ('clones' of *k**), because *k** will tend to lose votes to its clones. Worse, plurality will usually not even select one of the 'approximately correct' clones of *k**, because these clones will themselves tend to lose votes to similar alternatives. One response is to replace plurality rule by *approval voting*, in

which individuals can vote for any number of alternatives and the most often approved alternative wins; here the correct alternative will not tend to lose votes to its clones because individuals can approve many alternatives. Another response is to base the social choice on how each individual ranks alternatives from 'best' to 'worst'. Here we aggregate individual rankings into a socially winning alternative (or set of winning alternatives, if ties are permitted). This aggregation problem differs structurally from the preference-aggregation problem mentioned above, in that social outputs are winning alternatives rather than rankings of alternatives. Procedures with such outputs are called *social-choice* rules rather than *preference-aggregation* rules. Nonetheless, the logical difficulties surrounding preference aggregation – illustrated by Condorcet's voting paradox and culminating in Arrow's Theorem – reemerge for social choice procedures. For instance, there may not exist any 'Condorcet winner', an alternative that is majority-ranked over each other alternative – a problem closely related to Condorcet's voting paradox.

2.3 Jury theorems: formal truth-tracking arguments

Do aggregation procedures like those just introduced succeed in tracking the truth? Jury theorems provide formal 'wisdom of crowds' arguments, to the effect that appropriate procedures – typically majoritarian procedures – tend to generate true social beliefs, or 'correct' social decisions. Jury theorems can be powerful instruments – but they can also convey a misleading message when applied wrongly, because their optimistic conclusions may rely on misguided assumptions. The simplest and most common jury theorem is Condorcet's (1785) jury theorem. It assumes a majority vote between two social alternatives of which exactly one is 'correct' or 'better', for instance whether or not to convict or acquit a defendant. Jury theorems address the effect of increasing the size of the group. They ''typically conclude that 'crowds are wise' in one or both of the following senses:

- *The growing-reliability thesis*: Larger groups are better truth-trackers. That is, they are more likely to select the correct alternative (by majority) than smaller groups or single individuals.
- *The infallibility thesis*: Huge groups are infallible truth-trackers. That is, the likelihood of a correct (majority) decision tends to full certainty as the group becomes larger and larger.'' (Dietrich and Spiekermann forth.)

Different jury theorems differ in which of the two conclusions they reach and which premises they rest on. For instance, Condorcet's theorem reaches both conclusions, based on two simple premises: an *independence*

assumption whereby voters have independent probabilities of voting for the correct alternative, and a *competence* assumption whereby these correct-voting probabilities exceed ½ and are the same for each voter. The theorem's infallibility conclusion follows easily from the law of large numbers. The theorem's growing-reliability conclusion follows from a more sophisticated combinatorial argument. The infallibility conclusion is overly optimistic and has left many with the (correct) impression that 'something' must be wrong with jury theorems that reach that conclusion, although there is confusion about the source of the problem. Some blame Condorcet's competence assumption, arguing that real groups contain some incompetent individuals; while the competence assumption is indeed untenable as such, it can be weakened without reversing the infallibility conclusion, namely to the more plausible assumption that *average* competence in the group exceeds ½ (Owen et al. 1989; Dietrich 2008). The real problem is the independence assumption. Although this assumption can too be weakened without reversing the infallibility conclusion (e.g., Pivato 2017 and the literature reviewed in Dietrich and Spiekermann forth.), *plausible* weakenings of independence make the infallibility conclusion collapse (Dietrich and List 2004, Dietrich 2008). An important source of independence failure comes from *common causes* affecting voters: common evidence, common theoretical paradigms, or even 'non-evidential' common causes such as room temperature (Dietrich and Spiekermann 2013). The limited nature of available evidence places objective limits on the reliability of majority judgments, which cannot be miraculously overcome by including more and more voters exposed to the same limited evidence. But it would be hasty to dismiss jury theorems. Although the infallibility conclusion is too optimistic, the growing-reliability conclusion *can* be saved: that conclusion is reached by a jury theorem, which revises Condorcet's independence and competence assumptions in defensible ways (Dietrich and Spiekermann 2013). In other work, Condorcet's jury theorem has been extended beyond binary social decisions, such as choices between many alternatives via plurality voting (List and Goodin 2001) and premise-based judgment aggregation (List 2005). In fact, almost any standard voting rule has a corresponding jury theorem.[4] The question is thus not whether jury theorems 'exist' to defend a given aggregation rule, but whether those theorems make acceptable premises. While standard jury theorems address the epistemic performance of a *given* procedure (usually majority voting), one may alternatively search for the epistemically optimal procedure

[4] See Pivato (2017) for an abstract jury theorem that simultaneously covers most standard decision problems and voting methods, including majority voting, plurality voting, approval voting, Borda rule and other scoring rules, Kemeny rule and other distance-based rules, the median rules, etc.

which typically departs from majority voting, notably because of differential competence across individuals (Ben-Yashar and Nitzan 1997, Dietrich 2006).

3 Institutions beyond mere aggregation

Besides voting procedures, there are many other institutional and social processes for promoting social knowledge and correct social decisions. In this section we give three important examples: prediction markets, deliberation, and distributed search. Like belief aggregation procedures and other voting rules, prediction markets have predefined objects on their agenda (propositions, events, alternatives etc.). Deliberation, by contrast, is a more open and malleable process. Distributed search is different again, dividing cognitive labor in a structured way to solve complex problems.

3.1 Prediction markets

Suppose we want to predict whether a future event *E* will take place, such as "Bernie Sanders will be nominated as a candidate for the presidential election in the US in 2020 by the Democratic Party". We could ask a large number of people whether they believe this will happen. Alternatively, we could create a prediction market for winner-takes-all "event futures" that pay out $1 if the event takes place, and $0 if it does not. Starting from some initial allocation of these futures, traders can buy and sell them on a market created for that purpose. The market price for the future indicates the probability of the event, *Pr*(*E*). This is because the value of a future is the *expected* payout, which equals *Pr*(E) x $1. Traders thinking the current market price is below the event probability buy, traders thinking the opposite sell (Wolfers and Zitzewitz 2004).

The mechanisms behind successful prediction markets are aggregation and incentivization (Mann 2016). We have already seen that aggregation through voting can track the truth. The clever addition of prediction markets is the incentive created by gains when the future is under- or overprized. The competitive setting motivates participants to try to outwit others by finding more information. The prediction market is ''successful'' when it predicts – more precisely, attaches high price (probability) to – events that eventually happen. Prediction markets have often successfully predicted political events (see, e.g., the Iowa Electronic Markets[5]), but also business events, sports events, and even scientific events (Arrow et al 2008; Dreber et al. 2015). However, they are not infallible, mostly because of inefficient markets, systematic mistakes of traders, and insufficient or misleading

[5] Available at https://iemweb.biz.uiowa.edu/ [last accessed 17 April 2019].

information. For example, the Iowa Prediction Markets wrongly predicted a ‘’no’’ in the Brexit referendum and a Clinton victory over Trump in the 2016 US presidential election.[6] Prediction markets tend to perform best when the payout conditions are well-defined, the event happens in the near future, there are many market participants with access to independent evidence, and sources of systematic bias are limited.

3.2 Deliberation

The procedures discussed so far emphasize information merging, be it via aggregation rules or prediction markets. Deliberation, by contrast, emphasizes the exchange of reasons, prior to or instead of an aggregation. Is deliberation truth-conducive? Several epistemic effects are conceivable: (i) Deliberation can make private information or reasons public and allow participants to incorporate information and reasons of others; (ii) critical exchanges can eliminate bad reasons or inconsistencies and make good reasons and consistent viewpoints stand out more; (iii) deliberation might eliminate biases and reduce the influence of opinion leaders or other common causes; (iv) while deliberating, the decision problem can evolve, as new options come on the agenda and existing issues are reframed (Goodin and Spiekermann 2018, ch. 9); (v) deliberation can induce a meta-consensus about the structure of the decision space (Dryzek and List 2003; cf. List 2003, Dietrich and List 2010).

A good overview of the deliberation literature is Bächtiger et al. (2018); Landemore (2017) is helpful for the epistemic aspect. Extensive empirical research on “deliberative polls” suggests many benefits of deliberation (e.g., Fishkin 2018 and much of his other work). It is particularly striking that opportunities to deliberate with decision makers increase political knowledge (Esterling et al. 2011), though it is hard to disentangle whether this is an effect of the deliberative process or opportunities for influence. However, there are also warnings that deliberation can lead to increased polarization (Sunstein 2002).

3.3 Distributed Search

Some problems are best tackled by dividing epistemic labor. Scientific research, for example, progresses not because all scientists work on the same problems, with the same frameworks, theories, evidence, etc., but because of a competitive diversity of approaches, with incentive structures

[6] The low probabilities attached to Trump or Brexit victories were, strictly speaking, not mistakes. They merely showed that prediction markets, like many pollsters, underestimated support for Trump and the Brexit.

rewarding originality (Kitcher 1993, Strevens 2003). In a similar vein, some have argued that federalism is an epistemically advantageous political system because different political sub-units can search for and try out different solutions to a problem, an advantage that Judge Brandeis (1932) referred to as the "laboratory" of federalism. Distributed search is particularly suitable when (i) the search domain is large, e.g., many rival theories could be tested; and (ii) identifying correct choices requires effort, e.g., testing a theory (to an extent sufficient for acceptance or rejection) is an expensive and lengthy process. Such epistemic problems can be modelled in different ways (e.g., Zollman 2010; Weisberg and Muldoon 2009, Hong and Page 2012). The theoretical results suggest that division of epistemic labor is typically most productive if a balance is struck between diversity and learning. A group that eliminates diversity too quickly by converging on one candidate solution may miss better options. By contrast, a group that does not learn from the results that its members feed back will not benefit from the evidence the distributed search unearthed.

4 Some Sources of Social Epistemic Failure

We have already encountered several reasons why groups can fail to form true beliefs or to make correct decisions: its members can fall prey to biases, be influenced by opinion leaders or follow misleading evidence. In this section, we briefly review some further sources of social epistemic failure. Most epistemic failures begin as individual failures, pertaining to epistemology *in* groups. But since individual beliefs shape group beliefs, individual failures become social failures and may even get amplified through deliberation or aggregation. So they also pertain to epistemology *of* groups, our main focus. An exception is strategic voting, which involves no false individual beliefs, but may lead to social epistemic failures.

Strategic Voting. Even if voters care for nothing but true social beliefs or correct social decisions, they can have a strategic incentive to misrepresent their own view – which is surprising given the absence of conflicting goals. Consider a majority decision between options *x* and *y*. Voter Ann personally attaches high probability to *x* being the correct option. Instead of voting for *x*, she reasons as follows. Her vote makes a difference (is 'pivotal') only if the other voters are equally split between *x* and *y* (let the total number of voters be odd). Thinking strategically, Ann ignores all cases where she is not pivotal and assumes that *x* and *y* receive equal support from the others. But (so she reasons) if half of the others believe *y* to be correct, half of the others must have recognized a quality in *y* that she is unaware of. She infers that *y* is probably the correct option, and therefore votes for *y*, against her

individual judgment (Austen-Smith and Banks 1996).[7] Is such strategic voting an epistemic failure or enhancement? One might suspect the latter, given that individuals strategize in order to improve social beliefs. However, strategic voting equilibria can be epistemically worse than non-strategic voting because strategic voters discard private information. Fortunately, voters often have rational grounds for voting non-strategically. For example, they may care about expressing their own view instead of (or in addition to) caring about correctness of the voting outcome (Dietrich and Spiekermann forth.).

Social pressure. A potent cause of epistemic failure is the desire to conform with perceived expectations or norms, to avoid conflict or cognitive dissonance, to avoid mistakes in public, or to please peers. For example, group deliberation is likely to over-emphasize information that was already widely accepted prior to deliberation and under-emphasize information that is sparsely spread and yet crucial for a correct decision. Such an under-emphasis can happen not just because fewer people can mention the information, but also because those who can mention it stay silent, predicting, perhaps correctly, that the information will be controversial (Stasser and Titus 2011; Gigone and Hastie 1993). Related effects of social pressure are the amplification of individual errors in deliberation, and the possible enforcement of extreme positions – two deliberation failures (Sunstein and Hastie 2014a,b).

Motivated Cognition. The desire to fit into an identity group can bias the processing of factual information even at a cognitive level (e.g., Spiekermann and Weiss 2016). For example, Kahan et al. (2012) describe how different political-cultural backgrounds create different pressures to seek or avoid information about climate change and to process that information in biased ways. Kahan et al.'s left-liberal subjects believe that climate change is a problem, right-libertarians much less so. Scientific or numeracy skills further increase this divergence, suggesting that better understanding leads to even more biased processing.

Epistemic Injustice. Someone suffers epistemic injustice if she is "wronged specifically in her capacity as a knower" (Fricker 2007, 20). Fricker distinguishes between testimonial and hermeneutic injustice. Testimonial injustice is experienced if testimony is discounted because of an identity

[7] Note that Ann implicitly assumes that the others do not vote strategically, hence that their votes are evidence for what option is correct. Game-theoretically, the example shows that it may not be a (Bayes-Nash) equilibrium that everyone votes truthfully. Interestingly, there are many 'bad' equilibria, such as everyone voting for option y regardless of private information; here y wins even when all private information supports x.

prejudice. The evidence of a person of minority background might be dismissed because of his minority background. Hermeneutic injustice arises if oppressed individuals or groups do not have adequate concepts to conceive of, make sense of, or communicate their experiences. For example, Fricker explains that victims of sexual harassment found it difficult to understand their experience and explain it to others before the term 'sexual harassment' was coined. Testimonial and hermeneutic injustice undermine the ability of individual and ultimately the group to make correct decisions or form true beliefs.

Epistemic Skepticism and Nihilism. Another threat to epistemic success is the refusal to take a truth-seeking attitude. Epistemic skepticism is the belief that there is no truth to be tracked or that access to the truth is impossible. It can be distinguished from epistemic nihilism, the view that truth does not matter, regardless of whether it exists or not. While skepticism may be sometimes be justified, nihilism is not. Harry Frankfurt (2005) has suggested that the lack of care for the truth is one aspect of "bullshitting", a lack of commitment to the truth and epistemic norms of truthfulness (cf. Cassam 2019, especially ch. 4). The so-called "fake news" phenomenon seems fueled by epistemic nihilism, though the debate about this is still in its infancy (e.g., Mukerji 2018, Goodin and Spiekermann 2018, ch. 21). The skepticism or nihilism of individuals can scale up to the group level, undermining the truth-tracking ability of media organizations, public forums, or the political system as a whole.

References

Arrow, K. J. (1963). *Social choice and individual values* (2nd. rev.). New York/London: Wiley.

Arrow, K. J., Forsythe, R., Gorham, M., Hahn, R., Hanson, R., Ledyard, J. O., … Zitzewitz, E. (2008). The Promise of Prediction Markets. *Science*, *320*(5878), 877–878.

Austen-Smith, D., & Banks, J. S. (1996). Information aggregation, rationality, and the Condorcet jury theorem. *American Political Science Review*, *90*(1), 34–45.

Bächtiger, A., Dryzek, J. S., Mansbridge, J., & Warren, M. E. (Eds.). (2018). *The Oxford Handbook of Deliberative Democracy*. Oxford, New York: Oxford University Press.

Balinski, M. L., & Laraki, R. (2011). *Majority judgment: measuring, ranking, and electing*. Cambridge, Mass.; London: MIT Press.

Ben-Yashar, R. C., & Nitzan, S. I. (1997). The Optimal Decision Rule for Fixed-Size Committees in Dichotomous Choice Situations: The General Result. *International Economic Review*, *38*(1), 175.

Brams, S., & Fishburn, P. C. (2007). *Approval Voting* (2nd ed.). New York: Springer Science & Business Media.

Brandeis, L. *New State Ice Co. v. Liebmann, 285 U.S. 262 (Dissenting Opinion)*. , (1932).

Cassam, Q. (2018). *Vices of the Mind: From the Intellectual to the Political*. Oxford: Oxford University Press.

Condorcet De, M. (1785). *Essai sur l'application de l'analyse á la probabilité des décisions rendues á la pluralité des voix*.

Dietrich, F. (forthcoming). A Theory of Bayesian Groups. *Noûs*.

Dietrich, F. (2006). General Representation of Epistemically Optimal Procedures. *Social Choice and Welfare*, *26*(2), 263–283.

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

Dietrich, F. (2008). The premises of Condorcet's jury theorem are not simultaneously justified. *Episteme*, *5*(1), 56–73.

Dietrich, F. (2015). Aggregation theory and the relevance of some issues to others. *Journal of Economic Theory*, *160*, 463–493.

Dietrich, F., & List, C. (2004). A model of jury decisions where all jurors have the same evidence. *Synthese*, *142*, 175–202.

Dietrich, F., & List, C. (2007). Arrow's theorem in judgment aggregation. *Social Choice and Welfare*, *29*(1), 19–33.

Dietrich, F., & List, C. (2010). Majority voting on restricted domains. *Journal of Economic Theory*, *145*(2), 512–543.

Dietrich, F., & List, C. (2016). Probabilistic Opinion Pooling. In C. Hitchcock & A. Hajek (Eds.), *The Oxford Handbook of Probability and Philosophy* (pp. 519–543).

Dietrich, F., & List, C. (2017). Probabilistic opinion pooling generalized. Part one: general agendas. *Social Choice and Welfare*, *48*(4), 747–786.

Dietrich, F., & Mongin, P. (2010). The premiss-based approach to judgment aggregation. *Journal of Economic Theory*, *145*(2), 562–582.

Dietrich, F., & Spiekermann, K. (forthcoming). Jury Theorems. In J. W. Miranda Fricker Peter J. Graham, David Henderson, Nikolaj Pedersen (Ed.), *The Routledge Companion to Social Epistemology*. Retrieved from http://www.kaispiekermann.net/s/DietrichSpiekermann-JuryTheorems.pdf

Dietrich, F., & Spiekermann, K. (2013). Epistemic Democracy with Defensible Premises. *Economics and Philosophy*, *29*(1), 87–120.

Dokow, E., & Holzman, R. (2010). Aggregation of binary evaluations. *Journal of Economic Theory*, *145*(2), 495–511.

Dreber, A., Pfeiffer, T., Almenberg, J., Isaksson, S., Wilson, B., Chen, Y., … Johannesson, M. (2015). Using prediction markets to estimate the reproducibility of scientific research. *Proceedings of the National Academy of Sciences*, *112*(50), 15343–15347.

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

Esterling, K. M., Neblo, M. A., & Lazer, D. M. J. (2011). Means, Motive, and Opportunity in Becoming Informed about Politics: A Deliberative Field Experiment with Members of Congress and Their Constituents. *Public Opinion Quarterly*, *75*(3), 483–503.

Estlund, D. M. (2008). *Democratic Authority: A Philosophical Framework*. Princeton: Princeton University Press.

Fishkin, J. S. (2018). *Democracy When the People Are Thinking: Revitalizing Our Politics Through Public Deliberation*. Oxford, New York: Oxford University Press.

Frankfurt, H. G. (2005). *On Bullshit* (1 edition). Princeton, NJ: Princeton University Press.

Fricker, M. (2007). *Epistemic Injustice: Power & the Ethics of Knowing*. Oxford: Oxford University Press.

Genest, C., & Zidek, J. V. (1986). Combining Probability Distributions: A Critique and an Annotated Bibliography. *Statistical Science*, *1*(1), 114–135.

Gettier, E. L. (1963). Is Justified True Belief Knowledge? *Analysis*, *23*(6), 121–123.

Gigone, D., & Hastie, R. (1993). The common knowledge effect: Information sharing and group judgment. *Journal of Personality and Social Psychology*, *65*(5), 959–974.

Goldman, A., & Beddor, B. (2016). Reliabilist Epistemology. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016). Retrieved from https://plato.stanford.edu/archives/win2016/entries/reliabilism/

Goldman, A., & Blanchard, T. (2018). Social Epistemology. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2018). Retrieved from https://plato.stanford.edu/archives/sum2018/entries/epistemology-social/

Goldman, A. I. (1999). *Knowledge in a Social World*. Oxford: Clarendon.

Goodin, R. E., & Spiekermann, K. (2018). *An epistemic theory of democracy*. Oxford: Oxford University Press.

Hong, L., & Page, S. E. (2012). Some Microfoundations of Collective Wisdom. In H. Landemore & J. Elster (Eds.), *Collective Wisdom: Principles and Mechanisms*. Cambridge: Cambridge University Press.

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, *2*(10), 732–735.

Kitcher, P. (1995). *The advancement of science: science without legend, objectivity without illusions*. Oxford: Oxford University Press.

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

Landemore, H. (2017). Beyond the Fact of Disagreement? The Epistemic Turn in Deliberative Democracy. *Social Epistemology*, *31*(3), 277–295.

List, C. (2004). A Model of Path-Dependence in Decisions over Multiple Propositions. *American Political Science Review*, *98*(3), 495–513.

List, C. (2005). The probability of inconsistencies in complex collective decisions. *Social Choice and Welfare*, *24*(1), 3–32.

List, C. (2014). Three Kinds of Collective Attitudes. *Erkenntnis*, *79*(9), 1601–1622.

List, C., & Goodin, R. E. (2001). Epistemic Democracy: Generalizing the Condorcet Jury Theorem. *Journal of Political Philosophy*, *9*(3), 277–306.

List, C., & Pettit, P. (2002). Aggregating Sets of Judgments: An Impossibility Result. *Economics and Philosophy*, *18*(1), 89–110.

List, C., & Pettit, P. (2011). *Group agency: The possibility, design, and status of corporate agents*. Oxford: Oxford University Press.

Mann, A. (2016). The power of prediction markets. *Nature News*, *538*(7625), 308.

McConway, K. J. (1981). Marginalization and Linear Opinion Pools. *Journal of the American Statistical Association*, *76*(374), 410–414.

Miller, M. K., & Osherson, D. (2008). Methods for distance-based judgment aggregation. *Social Choice and Welfare*, *32*(4), 575.

Mongin, P. (1995). Consistent Bayesian Aggregation. *Journal of Economic Theory*, *66*(2), 313–351.

Mukerji, N. (2018). What is Fake News? *Ergo: An Open Access Journal of Philosophy*, *5*, 923–946.

Nehring, K., Pivato, M., & Puppe, C. (2014). The Condorcet set: Majority voting over interconnected propositions. *Journal of Economic Theory*, *151*, 268–303.

Nehring, K., & Puppe, C. (2010). Abstract Arrowian aggregation. *Journal of Economic Theory*, *145*(2), 467–494.

Owen, G., Grofman, B., & Feld, S. L. (1989). Proving a distribution-free generalization of the condorcet jury theorem. *Mathematical Social Sciences*, *17*, 1–16.

Peter, F. (2007). Democratic legitimacy and proceduralist social epistemology. *Politics, Philosophy & Economics*, *6*(3), 329–353.

Pettit, P. (2001). Deliberative Democracy and the Discursive Dilemma. *Philosophical Issues*, *11*(1), 268–299.

Pivato, M. (2017). Epistemic democracy with correlated voters. *Journal of Mathematical Economics*, *72*, 51–69.

Russell, J. S., Hawthorne, J., & Buchak, L. (2015). Groupthink. *Philosophical Studies*, *172*(5), 1287–1309.

Scanlon, T. (1998). *What we owe to each other*. Cambridge, Mass., London: Belknap Press of Harvard University Press.

Schmitt, F. (1999). Social Epistemology. In J. Greco & E. Sosa (Eds.), *The Blackwell Guide to Epistemology* (pp. 354–382). Malden, MA: Blackwell.

Spiekermann, K., & Weiss, A. (2016). Objective and Subjective Compliance: A Norm-Based Explanation of `Moral Wiggle Room'. *Games & Economic Behavior*, *96*, 170–183.

Stasser, G., & Titus, W. (2003). Hidden Profiles: A Brief History. *Psychological Inquiry*, *14*(3–4), 304–313.

Strevens, M. (2003). The Role of the Priority Rule in Science. *The Journal of Philosophy*, *100*(2), 55–79.

Sunstein, C. R. (2002). The Law of Group Polarization. *Journal of Political Philosophy*, *10*(2), 175–195.

Sunstein, C. R., & Hastie, R. (2014a). Garbage in, garbage out? Some micro sources of macro errors. *Journal of Institutional Economics*, 1–23.

Sunstein, C. R., & Hastie, R. (2014b). *Wiser: Getting Beyond Groupthink to Make Groups Smarter* (First Edition edition). Boston, Massachusetts: Harvard Business Review Press.

Wagner, C. (1982). Allocation, Lehrer models, and the consensus of probabilities. *Theory and Decision*, *14*(2), 207–220.

Weisberg, M., & Muldoon, R. (2009). Epistemic Landscapes and the Division of Cognitive Labor. *Philosophy of Science*, *76*(2), 225–252.

Wolfers, J., & Zitzewitz, E. (2004). Prediction Markets. *Journal of Economic Perspectives*, *18*(2), 107–126.

Zollman, K. J. S. (2010). The epistemic benefit of transient diversity. *Erkenntnis*, *72*(1), 17–35.