10.1 Social Epistemology

Franz Dietrich and Kai Spiekermann

Summary

Social epistemology studies knowledge in social contexts. We distinguish two types, depending on the holder of knowledge. Epistemology *in* groups studies individuals learning from others, while epistemology *of* groups studies groups as knowers, literally or metaphorically. 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 or actions aim at "correctness"—in virtue of facts that are empirical or normative, real or constructed, objective or only intersubjective, universal or relative, and so on. Procedures and institutions are evaluated by epistemic performance: Are they truth-conducive? Do groups become "wiser" than their members? This chapter reviews several procedures and institutions, discussing epistemic successes and failures. Jury theorems provide formal arguments for epistemic success, but only some have defensible premises. It will be argued that larger groups can outperform smaller groups, yet without becoming infallible.

1. Scope and Problems of Social Epistemology

Social epistemology is the branch of epistemology that studies knowledge in social contexts. This review first sets the stage by introducing clarifications, distinctions, and applications (section 1). It then discusses aggregation procedures (section 2) and institutional arrangements (section 3) that generate collective outcomes that "track the truth" or lead to epistemic failures (section 4).

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 as a species of individual epistemology, although its focus lies in how individual knowledge depends on social inputs; typical questions are rational responses to testimony and peer disagreement and "irrational" or subpersonal responses such as belief contagion through social media. Epistemology *of* groups ascribes knowledge to the group; let us talk here of *social* knowledge or belief (List & Pettit, 2011; see also Goldman & O'Connor, 2019; 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 qualifies as an *agent* and for what it *means* for the group to believe or know something, ranging from literal to purely metaphorical meanings. Arguably, 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 metaphorically. List (2014) distinguishes between three types of social beliefs: *aggregate* beliefs are mere summaries of 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" 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 is studied using aggregation models. In this review, we focus on epistemology *of* groups, but for other aspects of social epistemology, see Goldman and O'Connor (2019), Goldman and Whitcomb (2011), Goldman (1999), and Schmitt (1999).[1]

1.2 Narrow versus Broad Social Epistemology

Narrowly construed, social epistemology addresses only social knowledge or belief: 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; and so on. Broadly construed, however, social epistemology addresses social decisions in a wide sense, covering not only beliefs but also, for instance, actions: courts passing sentences, scientists choosing experiments, parliaments passing laws, governments setting goals, commissions working out ethical standards, and so forth. For such a social decision to be an epistemic matter, it must have the property of being *correct or incorrect* in virtue of some fact or truth. A court sentence may be an epistemic matter, its being correct or incorrect depending on whether the crime has been committed and merits the sentence. Selecting Oscar winners *might* be an epistemic matter. One reason for adopting this broad notion of social epistemology could be that correct decisions often are, or can be rationalized as, the result of beliefs. This chapter uses "(social) decision" in the widest sense, to include, for instance, social beliefs and social actions.[2] To be an object of social epistemology, broadly construed, the decision must, however, be correctness-apt.[3]

1.3 The Truth to Be Tracked

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

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

Social epistemology with normative facts is controversial, since the existence and the nature of normative facts are debatable.[4] But epistemology about normative facts should not be dismissed, because truth-tracking is meaningful regardless of whether normative facts are real or constructed (e.g., socially constructed), objective or at least intersubjective (within the group), universal or relativistic (as in cultural relativism), and natural or non-natural. The fact—normative or other—should, however, be suitably stable and procedure-independent. If the fact is of a constructed kind, then its existence and nature should not emerge from the social procedure but precede that procedure; otherwise, there is no truth to be tracked, but one to be constructed.

1.4 Truth-Tracking and Rationality

Rationality, understood here as internal coherence (i.e., formal rather than substantive rationality), is necessary but not sufficient for correctness. Beliefs can be perfectly consistent yet false, betterness judgments can be transitive yet misguided, and so on.[5] Rationality is thus only one step toward correctness. But this step is already very difficult to take when the holder of the beliefs (judgments, evaluations, and so on) is a group rather than an individual. Importantly, rationality of the group members does not guarantee rationality of the group, as we know from many aggregation paradoxes, such as Condorcet's voting paradox and the discursive dilemma, and from general impossibility results such as Arrow's Theorem (see section 2.1 and also chapter 10.2 by Schmid, this handbook). The sheer difficulty of achieving social rationality might explain why much of social choice theory focuses on social rationality, setting aside correctness. This limited focus does not place social choice theory outside the social-epistemological endeavor, however, as rationality is necessary for correctness. The interaction between individual rationality, social rationality, and truth-tracking is complex. For example, irrational individual belief revision can hinder individual truth-tracking while leading to a fruitful epistemic diversity that promotes social truth-tracking (see section 3.2).

1.5 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 "Gettier cases," in which a subject (for us: a group) forms justified true beliefs by sheer coincidence (Gettier, 1963). It is also controversial whether justification to hold a belief comes from sufficient evidence or from a reliable procedure (see Goldman & Beddor, 2016, on evidentialism versus reliabilism). Despite their familiarity, such "knowledge problems" can take a distinctive form for *social* knowledge, notably because groups use different procedures than individuals to acquire beliefs (e.g., voting procedures). Nonetheless, this chapter shall set aside the question of what exactly social knowledge means and instead zoom in on how social beliefs and decisions more broadly can be formed and track the truth.

1.6 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 core claim of *epistemic* justifications of democracy is normative: institutions and procedures *should*

track the truth to be legitimate (Estlund, 2008). By contrast, *proceduralist* justifications of democracy maintain that institutions or procedures are not legitimate because they track any independent truth but because they are 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 the oracle's recommendation for society is epistemically good (it tracks the truth) but procedurally bad (it is totally undemocratic). The debate between epistemic and procedural theories of democracy is ongoing (for a hybrid position, see Peter, 2007). For epistemic theories of democracy, social epistemology is a highly relevant enterprise.

2. Aggregation Procedures to Track the Truth

This section turns to prominent procedures generating social beliefs (section 2.1) or other truth-tracking decisions (section 2.2) and then sketches how *jury theorems* help establish the truth-conduciveness of such procedures (section 2.3).

2.1 Forming Social Beliefs or Judgments

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. Social abstentions are allowed: society can say "yes," or say "no," or abstain (but for simplicity, individuals cannot abstain). Here are some examples. *Majority rule* makes society believe what the majority believes (and abstain in case of a tie). *Asymmetric supermajority rules* make social belief in 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* 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., Dietrich, 2007; List & Pettit, 2002). For instance, a court may need collective beliefs about three propositions: the defendant has committed a certain act (p), such an act breaks the law (q), and 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 the 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 & Sager, 1986) or the *discursive dilemma* (Pettit, 2001). The discursive dilemma generalizes far beyond propositionwise majority voting. General impossibility theorems establish that, as soon as propositions are sufficiently interconnected, there exist *no* propositionwise judgment-aggregation procedures that are well behaved in certain senses; see in particular List and Pettit's (2002) theorem and the Arrow-type theorem for judgment aggregation (Dietrich & List, 2007; Dokow & Holzman, 2010; both building on results of Klaus Nehring and Clemens Puppe later published in Nehring & 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., Dietrich & Mongin, 2010; Pettit, 2001). 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, thereafter 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 that we can prioritize certain (premise) propositions over other (conclusion) propositions. Rival

holistic procedures are distance-based rules (e.g., Miller & Osherson, 2008), sequential rules (e.g., Dietrich, 2015; List, 2004), relevance-based rules (Dietrich, 2015), and "approximate majoritarian" rules (Nehring, Pivato, & 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 Dietrich & List, 2016; Genest & Zidek, 1986). Once beliefs are probabilistic rather than binary, coherence of beliefs consists in respecting probabilistic principles like additivity, rather than logical principles like 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"), as explained in chapter 4.1 by Hájek and Staffel (this handbook). 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* averaging repairs this flaw, that is, produces dynamically rational social beliefs. Here, social probabilities of worlds are obtained by a possibly weighted *geometric* average of individual probabilities (followed by a simple normalization ensuring that the total probability of worlds is 1). Geometric opinion pooling is in fact the *only* way to guarantee "Bayesian" group beliefs of a suitably well-behaved sort, as has only recently been shown (Dietrich, 2019; Russell, Hawthorne, & Buchak, 2015). Yet geometric pooling has a different flaw, which linear pooling avoids: social probabilities are not robust to refining or coarsening the algebra of propositions considered, making social beliefs description-sensitive.

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 some evaluations (e.g., of size) and more controversial for others (e.g., of well-being). 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, or ranks weakly above, y (with respect to the relevant criterion). *Absolute* evaluations are captured by a function assigning to each object in X a value or rank from a set V of possible values or ranks, for example, 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 the *social* evaluations of objects be, given the evaluations by the group 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. This impossibility can be escaped by using some "holistic" rather than pairwise aggregation rule, for instance, the Borda rule or a distance-based rule. A different escape is to replace ordinal by absolute evaluations.
- 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 "nonapproved"—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 nonapproved otherwise (Brams & Fishburn, 2007). Balinski and Laraki generalize this procedure to any set of values linearly ranked from "highest"

to "lowest," for example, {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, and so on. 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 & Laraki, 2011) but remains controversial within social choice theory with its ordinalist tradition.

Aggregating evaluations can be regarded as a special case of judgment aggregation by using propositions of type "x ranks over y" (for options x and y in X), in the ordinal case, or "x has value or rank v" (for options x in X and values v in V), in the cardinal case.

2.2 Making Social Decisions

The procedures discussed above generate social beliefs of different sorts. This section turns to procedures for making other social decisions.[6] To stay within social epistemology (broadly construed), the decision in question is assumed to track some truth, in particular, can be "correct" or "incorrect" in some procedure-independent sense (see sections 1.2, 1.3, and 1.6). 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. When using *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 with *approval voting* (Brams & Fishburn, 2007), 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 (1963) 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 correct social beliefs or other 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 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 "groups are wise" in the sense that one or both of two controversial hypotheses holds:

- *The growing-reliability hypothesis*: Larger groups are better truth-trackers. That is, they are more likely to reach the correct outcome (in most jury theorems: by majority) than smaller groups or single individuals.
- *The infallibility hypothesis*: Huge groups are infallible truth-trackers. That is, the likelihood of a correct outcome tends to full certainty as the group becomes larger and larger (Dietrich & Spiekermann, 2020).

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 infallibility hypothesis then follows easily using the law of large numbers and the growing-reliability hypothesis follows from a more sophisticated combinatorial argument. The infallibility hypothesis, however, is overly optimistic and has left many with the (correct) impression that "something" must be wrong with those jury theorems that reach that conclusion, although there is confusion about the source of the problem. Some blame Condorcet's unrealistic competence assumption. However, the implausible infallibility conclusion remains if competence can vary across individuals as long as it exceeds ½ *on average* (Dietrich, 2008; Owen, Grofman, & Feld, 1989). The real problem is the independence assumption. Although this assumption can also be weakened *without* losing infallibility (e.g., Pivato, 2017, and the literature reviewed in Dietrich & Spiekermann, 2020), *plausible* weakenings of independence make infallibility collapse (Dietrich, 2008; Dietrich & List, 2004).

An important source of independence failure are *common causes* affecting voters: common evidence, common paradigms or perspectives, and even noninformational common causes such as room temperature (Dietrich & Spiekermann, 2013). The limited nature of available evidence places objective bounds 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 infallibility is unrealistic, the growing-reliability conclusion *can* be saved: that conclusion is reached by a jury theorem that revises Condorcet's independence and competence assumptions in defensible ways (Dietrich & Spiekermann, 2013, 2020). In other work, Condorcet's jury theorem has been extended beyond binary social decisions, such as choices between many alternatives via plurality voting (List & Goodin, 2001). In fact, there is a jury theorem for almost any standard voting rule, including approval voting, distance-based rules, the Borda rule, and other scoring rules (Pivato, 2017). The question is thus not whether jury theorems "exist" to defend a given aggregation rule but whether those theorems use acceptable premises.[7]

3. Institutions beyond Mere Aggregation

Besides voting procedures, there are many other institutional and social processes for promoting social knowledge and correct social decisions. This section gives three important examples: deliberation, distributed search, and prediction markets.

3.1 Deliberation

The procedures discussed so far emphasize information aggregation. Deliberation, by contrast, emphasizes the exchange of reasons, prior to or instead of aggregation. Is deliberation truth-conducive? Several epistemic effects are at work: (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 and other common causes; (iv) while deliberating, the decision problem can evolve, as new options come on the agenda and existing issues are reframed (Goodin & Spiekermann, 2018, chapter 9); and (v) deliberation can induce a meta-consensus about the structure of the decision space (Dryzek & List, 2003; cf. Dietrich & List, 2010; List, 2002).

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 for citizens to deliberate with decision makers increase political knowledge (Esterling, Neblo, & Lazer, 2011), although it is hard to disentangle whether this is an effect of the deliberative process or of opportunities for influence. However, there are also warnings that deliberation can lead to increased polarization (Sunstein, 2002). A good overview of the deliberation literature is Bächtiger, Dryzek, Mansbridge, and Warren (2018); Landemore (2017) is helpful for epistemic aspects.

3.2 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, and so on, but because of a competitive diversity of approaches, with incentive structures rewarding originality while ensuring that resources are not too concentrated. In a similar vein, some have argued that federalism is an epistemically advantageous political system because different political subunits can search for and try out different solutions to a problem, an advantage that Judge Brandeis referred to as the "laboratory" of federalism (*New State Ice Co. v. Liebmann*, 1932). Distributed search is particularly suitable when (i) the search domain is large (e.g., many rival theories could be tested); (ii) identifying correct choices requires effort, for example, testing a theory (to an extent

sufficient for acceptance or rejection) is an expensive and lengthy process; and (iii) search in different locations causes useful by-products (e.g., new discoveries).

In distributed search, individual rationality need not be conducive to social rationality. In scientific research, for example, a bad incentive system would render it individually rational for each scientist to pursue research in novel areas receiving most attention and funding. However, that will likely lead to inefficient allocation of resources while valuable subfields are left unexplored, hampering overall scientific progress. A socially rational division of scientific labor would promote more diversity, more replication, and work on important projects with less imminent rewards. The influential work by Kitcher (1993) and Strevens (2003) shows how a mix of incentive schemes helps modern science manage this tension, often with success. The importance of diversity and the tension between individual and social rationality have been modeled in different ways (e.g., Hong & Page, 2012; Weisberg & Muldoon, 2009; Zollman, 2010; see also chapter 14.1 by Andersen & Andersen, this handbook). The theoretical results suggest that a socially irrational division of epistemic labor can emerge from individual rationality.

3.3 Prediction Markets

Prediction markets are financial markets from which predictions of contingent events, such as election outcomes, emerge. The securities traded on these markets are futures on the occurrence of these events. The mechanisms behind prediction markets are aggregation and incentivization (Mann, 2016; Wolfers & Zitzewitz, 2004). 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 overpriced. 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[8]), 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 information. For example, the Iowa Prediction Markets wrongly predicted a "no" in the 2016 UK "Brexit" referendum and a Clinton victory over Trump in the 2016 US presidential election.[9]

Prediction markets track the truth in an interestingly different way: while standard aggregation procedures transform explicit individual judgments into an explicit group judgment or other decision, prediction markets transform entirely behavioral inputs (buying and selling) into market prices from which implicit, but precise, group judgments can be read off. Although individual judgments are revealed only indirectly through acts of buying or selling, they are elicited reliably: while voting gives incentives to misrepresent, prediction markets do not normally provide such opportunities because counter-judgmental trading leads to losses.

4. Some Sources of Social Epistemic Failure

This chapter has already shown several reasons why groups can fail to form correct beliefs or other correct decisions: its members can fall prey to biases, be influenced by opinion leaders, or follow misleading evidence. This section briefly reviews 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, the main focus of this chapter. An exception is strategic voting, which involves no false individual beliefs but may lead to social epistemic failures.

4.1 Strategic Voting

Even if voters care for nothing but correctness of social beliefs or decisions more broadly, they can have strategic incentives to misrepresent their views—which is surprising given the absence of conflicting goals. For instance, a juror may vote for convicting the defendant, *although* her private information suggests innocence, as she reasons strategically that this makes the jury's (majority or supermajority) decision more likely correct, despite making her own vote less likely correct (see Austen-Smith & Banks, 1996; for details of the strategic rationale, see Dietrich & Spiekermann, 2020). So individual and social rationality can conflict even when, unlike in a prisoner's dilemma or tragedy of the commons, individual and social interests are perfectly aligned.

It is, however, important to put things into perspective. First, it is debatable whether such strategizing is socially harmful: while strategic voters systematically ignore private information (leading to underinformed social outcomes), they do so in order to improve social outcomes. The harmfulness of strategizing depends on the models and assumptions used. Second, voters often have perfectly rational grounds for voting truthfully, as they may care about expressing their own view instead

of (or in addition to) caring about correctness of the voting outcome, a phenomenon usually ignored in the strategic voting literature with its consequentialist orientation (Dietrich & Spiekermann, 2020). Paradoxically, voters stop strategizing because they *stop* (not *begin*) sharing the same goal of correct outcomes.

4.2 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 overemphasize information that was already widely accepted prior to deliberation and underemphasize information that is sparsely spread and yet crucial for a correct decision. Such an underemphasis can happen not just because there are fewer people who hold the information in the first place but also because those who hold it stay silent, predicting, perhaps correctly, that the information will be controversial (Gigone & Hastie, 1993; Stasser & Titus, 2003). Related effects of social pressure are the amplification of individual errors in deliberation and the possible enforcement of extreme positions—two deliberation failures (Sunstein & Hastie, 2015a, 2015b).

4.3 Motivated Cognition

The desire to fit into an identity group or comply with social norms can bias the acquisition and cognitive processing of factual information in the first place (e.g., Spiekermann & 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 such 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.

4.4 Epistemic Injustice

Someone suffers epistemic injustice if she is "wronged specifically in her capacity as a knower" (Fricker, 2007, p. 20). Fricker distinguishes between testimonial and hermeneutic injustice. *Testimonial* injustice is experienced if testimony is discounted because of an identity prejudice; for example, the evidence of a person of minority background is dismissed *because* of their background. *Hermeneutic* injustice can arise 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 individuals and ultimately the group to generate correct beliefs or other correct decisions.

4.5 Epistemic Skepticism and Nihilism

A more fundamental 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 or not it exists and is accessible. Whereas skepticism is a respectable epistemological position, nihilism is an epistemologically and morally dubious attitude. However, genuine nihilists will by definition not care about whether their attitude is defensible, which makes nihilism particularly immune to arguments and hard to fight against. 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 chapter 4). The so-called fake news phenomenon seems fueled by epistemic nihilism, although the debate about this is still in its infancy (e.g., Goodin & Spiekermann, 2018, chapter 21; Mukerji, 2018; cf. Rini, 2017). 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.

Acknowledgments

We thank Wolfgang Spohn and two anonymous referees for very helpful suggestions. We also acknowledge financial support by the French National Research Agency (ANR) through three grants (ANR-17-CE26-0003, ANR-16-FRAL-0010, and ANR-17-EURE-0001).

Notes

1. This volume touches upon other issues of formal epistemology in chapter 2.6 by van Benthem, Liu, and Smets; chapter 5.2 by Rott; chapter 5.3 by Kern-Isberner, Skovgaard-Olsen, and Spohn; chapter 5.5 by Hahn and Collins; and chapter 5.6 by Woods.

2. By using "social decision" broadly to encompass social beliefs we do not assume that group beliefs (let alone individual beliefs) always are under direct voluntary control. We set aside whether groups have such control, and what it would mean.

3. If groups are construed as intentional agents, they also hold non-belief 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; Scanlon, 1998, p. 20). The correctness of holding, say, an intention might depend on whether its content is morally permissible. For instance, it might be correct for a firm to intend to maximize shareholder value if maximizing it is morally permissible. Like social beliefs and correctness-apt actions, correctness-apt social attitudes might be regarded as a matter of social epistemology (broadly construed). Whether or not one adopts this inclusive view, some of the technical machinery of social epistemology—like jury theorems—can be used to study whether group attitudes become correct.

4. A hybrid position is to deny normative facts but to grant that normative judgments can be correct or incorrect as acts (on other grounds than the truth or falseness of their contents). This is still compatible with social epistemology broadly construed (cf. section 1.2).

5. What (formal) rationality means precisely depends on the type of belief (or other attitude or decision) considered and can be fleshed out differently. Binary yes–no beliefs are often required to be consistent and deductively closed. Probabilistic beliefs are often required to satisfy Kolmogorov's axioms. Ordinal betterness judgments over some choice options are often required to be transitive and perhaps complete.

6. The social beliefs or judgments discussed in section 2.1 often already generate social actions: society can choose what it believes to be best (ordinally) or what it approves or rates highest (absolutely) or what is recommended by its judgments, and so on. The current section, however, investigates procedures that generate social actions *directly*, rather than generating beliefs or judgments *supporting* actions.

7. While standard jury theorems address the epistemic performance of a given procedure (usually majority voting), one may alternatively search for the epistemically optimal procedure, thereby accounting for differential competence across individuals (Ben-Yashar & Nitzan, 1997; Dietrich, 2006).

8. Available at https://iemweb.biz.uiowa.edu (last accessed April 17, 2019).

9. 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 Brexit.

References

Arrow, K. J. (1963). *Social choice and individual values* (2nd rev. ed.). New York, NY: 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, England: Oxford University Press.

Balinski, M., & Laraki, R. (2011). *Majority judgment: Measuring, ranking, and electing*. Cambridge, MA: 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), 175–186.

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

Cassam, Q. (2019). *Vices of the mind: From the intellectual to the political*. Oxford, England: Oxford University Press.

Condorcet, A., Marquis de. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* [An essay on the application of probability theory to plurality decision-making]. Paris, France: Imprimerie Royale.

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. (2019). A theory of Bayesian groups. *Noûs*, *53*, 708–736.

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 A. Hájek & C. Hitchcock (Eds.), *The Oxford handbook of probability and philosophy* (pp. 519–543). Oxford, England: Oxford University Press.

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. (2013). Epistemic democracy with defensible premises. *Economics & Philosophy*, *29*(1), 87–120.

Dietrich, F., & Spiekermann, K. (2020). Jury theorems. In M. Fricker, P. J. Graham, D. Henderson, & N. J. L. L. Pedersen (Eds.), *The Routledge handbook of social epistemology* (pp. 386–396). New York, NY: Routledge.

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, NJ: Princeton University Press.

Fishkin, J. S. (2018). *Democracy when the people are thinking: Revitalizing our politics through public deliberation*. Oxford, England: Oxford University Press.

Frankfurt, H. G. (2005). *On bullshit*. Princeton, NJ: Princeton University Press.

Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford, England: 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. I. (1999). *Knowledge in a social world*. Oxford, England: Clarendon Press.

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

Goldman, A., & O'Connor, C. (2019). Social epistemology. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from https://plato.stanford.edu/archives/fall2019/entries/epistemology-social/

Goldman, A. I., & Whitcomb, D. (Eds.). (2011). *Social epistemology: Essential readings*. Oxford, England: Oxford University Press.

Goodin, R. E., & Spiekermann, K. (2018). *An epistemic theory of democracy*. Oxford, England: 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* (pp. 56–71). Cambridge, England: 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. (1993). *The advancement of science: Science without legend, objectivity without illusions*. Oxford, England: 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. (2002). Two concepts of agreement. *The Good Society*, *11*(1), 72–79.

List, C. (2004). A model of path-dependence in decisions over multiple propositions. *American Political Science Review*, *98*(3), 495–513.

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 & Philosophy*, *18*(1), 89–110.

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

Mann, A. (2016). The power of prediction markets. *Nature News*, *538*(7625), 308–310.

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), 575–601.

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.

New State Ice Co. v. Liebmann, 285 U.S. 262 (1932).

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.

Rini, R. (2017). Fake news and partisan epistemology. *Kennedy Institute of Ethics Journal*, *27*(2), E-43–E-64.

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

Scanlon, T. M. (1998). *What we owe to each other*. Cambridge, MA: 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. *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. (2015a). Garbage in, garbage out? Some micro sources of macro errors. *Journal of Institutional Economics*, *11*(3), 561–583.

Sunstein, C. R., & Hastie, R. (2015b). *Wiser: Getting beyond groupthink to make groups smarter*. Boston, MA: 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.

This is a section of doi:10.7551/mitpress/11252.001.0001

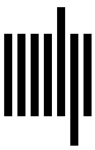
The Handbook of Rationality

Edited by: Markus Knauff, Wolfgang Spohn

Citation:

The Handbook of Rationality
Edited by: Markus Knauff, Wolfgang Spohn
DOI: 10.7551/mitpress/11252.001.0001
ISBN (electronic): 9780262366175
Publisher: The MIT Press
Published: 2021

Funding for the open access edition was provided by the MIT Libraries Open Monograph Fund.



The MIT Press

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Knauff, Markus, editor. | Spohn, Wolfgang, editor.
Title: The handbook of rationality / edited by Markus Knauff and Wolfgang Spohn.
Description: Cambridge : The MIT Press, 2021. | Includes bibliographical references and index.
Identifiers: LCCN 2020048455 | ISBN 9780262045070 (hardcover)
Subjects: LCSH: Reasoning (Psychology) | Reason. | Cognitive psychology. | Logic. | Philosophy of mind.
Classification: LCC BF442 .H36 2021 | DDC 153.4/3—dc23
LC record available at https://lccn.loc.gov/2020048455