Assessing the epistemic quality of democratic decision-making in terms of adequate support for conclusions

Published in *Social Epistemology* 31:3, 2017, 251–265.

doi: https://doi.org/10.1080/02691728.2017.1317866

Henrik Friberg-Fernros
Department of Political Science, University of Gothenburg
Box 711, 405 30 Gothenburg, Sweden
Phone: +46-31-7861231 E-mail: henrik.friberg-fernros@pol.gu.se

Johan Karlsson Schaffer (corresponding author)
Norwegian Centre for Human Rights, University of Oslo
School of Global Studies, University of Gothenburg
Box 700, 405 30 Gothenburg, Sweden
Phone: +46-31-7865931 E-mail: j.k.schaffer@gu.se

Abstract

How can we assess the epistemic quality of democratic decision-making? Sceptics doubt such assessments are possible, as they must rely on controversial substantive standards of truth and rightness. Challenging that scepticism, this paper suggests a procedure-independent standard for assessing the epistemic quality of democratic decision-making by evaluating whether it is adequately supported by reasons. Adequate support for conclusion is a necessary (but insufficient on its own) aspect of epistemic quality for any epistemic justification of democracy, though particularly relevant to theories that emphasize public deliberation. Finding existing methods for measuring the quality of public discourse to fall short, we draw on an epistemological theory of argumentation to provide a more sophisticated approach to evaluating the epistemic quality of

democratic decision-making, illustrate how the approach can be used, and demonstrate its relevance for the epistemic turn in democratic theory.

Keywords

Epistemic democracy, deliberative democracy, epistemological theory of argumentation, argument schemes, support adequacy, discourse quality.

Acknowledgments

For their comments on earlier versions of this paper, we thank Svenja Ahlhaus, Sveinung Arnesen, Simone Chambers, Felix Gerlsbeck, Peter Hallberg, Andrew Knops, Helene Landemore, Olof Leffler, Peter Niesen, Valeria Ottonelli, Markus Patberg, Sean Sinclair and Daniele Santoro. We have also benefited greatly from discussing some of these ideas with Christoph Lumer and Douglas Walton.

Funding and grant-awarding bodies

This research was financially supported by an Independent Project Grant from the Research Council of Norway ('The Consensus Paradox', project no. 213547).

Bio notes

Henrik Friberg-Fernros is an associate professor at the Department of Political Science, University of Gothenburg. **Johan Karlsson Schaffer** is an associate professor at the School of Global Studies, University of Gothenburg, and a senior researcher at the Norwegian Centre for Human Rights, University of Oslo. Correspondence to: Johan Karlsson Schaffer, School of Global Studies, University of Gothenburg Box 700, 405 30 Gothenburg, Sweden. E-mail: j.k.schaffer@gu.se

1. Epistemic quality – impossible to measure?

How can we assess the epistemic quality of democratic decision-making? Recent years have seen increased academic interest in the epistemic dimensions of democracy, as political theorists have debated whether and how democracy can be justified in terms of producing better decisions than feasible alternatives (e.g. Estlund 2008; Landemore 2013; List and Goodin 2001), while empirical researchers have examined the epistemic effects of various decision-making procedures (e.g. Caluwaerts and Ugarriza 2012; Niemeyer 2011).

Yet this growing literature on epistemic democracy still lacks viable methods for assessing the epistemic quality of democratic decision-making. For instance, while the theory of deliberative democracy has long claimed that producing more robust, well-informed decisions is a key advantage of deliberation over alternative procedure, recent overviews of current methods for analysing group deliberation report no approaches that aim to measure epistemic quality (Black et al. 2011; Beste 2013). Some even claim that measuring the epistemic quality of democratic decision-making would be virtually impossible. For one thing, any such measure would be inherently controversial, involving the use of 'external … standards of fairness or rightness, on which even heated philosophical debates proved endless' (Caluwaerts and Ugarriza 2012, 5; cf. Beste 2013, 28; Cooke 2000, 955). And even if we could agree on the relevant standards of correctness, we will probably disagree 'over how particular decisions measure up against those standards' (Ingham 2013, 139).

In this paper, we will accept the challenge of this scepticism. We shall propose a procedure-independent approach for assessing the epistemic quality of democratic decision-making that is neither dependent on substantive external standards of correctness nor determined by specific measurements based on such standards. Expanding on previous suggestions to employ argumentation theory to assess the rationality of the product of democratic decision-making

(Knops 2006; cf. 2017 [this issue]; Friberg-Fernros and Schaffer 2014), we propose that one can measure an important aspect of epistemic quality by assessing whether democratic decision-making is adequately supported on reasons. Our approach draws on established practices in formal and informal logic to evaluate whether the inferences that a collective decision comprises are logically sound or internally valid: Does the conclusion follow from the premises? In practice, this approach entails, first, that one reconstructs inferences, either from speeches made in the deliberative phase preceding a decision or from the decision itself. Second, one evaluates whether the inference meets the relevant epistemic criteria for the argument scheme being used. An inference that fails to pass this test is inadequately supported by reasons, and, we submit, should be considered as having less epistemic value than one that is adequately supported by reasons, ceteris paribus.

Consequently, the epistemic quality of democratic decision-making can be assessed in terms of how well the inferences made through the process or in the decision meet these criteria. This, we shall argue, suggests a necessary but – importantly – insufficient element of epistemic quality for any epistemic justification of democracy, though it is particularly crucial to deliberative theories of democracy, as they emphasize that the public reason-giving aspects of democratic decision-making promote epistemic quality.

Hence, the paper makes two important contributions: First, meeting the sceptics' challenge, we make a theoretical contribution by demonstrating that a procedure-independent standard for assessing the epistemic quality of decision-making need not rely on external standards of moral or factual correctness, which may be both inherently controversial and difficult to employ in empirical research. Second, drawing on an epistemological theory of argumentation, we offer a framework for assessing the epistemic quality of democratic decision-making by detecting inadequate support for conclusions, and demonstrate the potential of the approach.

The rest of this paper is structured in five parts. In section two, we engage with the sceptics' challenge by arguing that the epistemic quality of decision-making can be detected without assessing whether the decision in question is morally or factually correct. We also show how the approach matters for different theories about what mechanisms make democratic decision-making epistemically superior. In section three, we survey methods for measuring the epistemic quality of decision-making in existing research, and show why even the much-used Discourse Quality Index (Steiner et al. 2004) falls short of assessing whether reasons adequately support conclusions. In section four, we present our approach in greater detail: Inspired by argumentation theory, we suggest three types of argument schemes and their corresponding criteria of inadequate support, and outline how those criteria can be used to assess the epistemic quality of democratic decision-making. In section five, we respond to three important challenges to our framework, before we finally conclude the paper by elaborating what our argument implies for empirical and theoretical research on the epistemic aspects of democracy.

2. Political cognitivism and support adequacy

Now, our first task is to meet the sceptics challenge by showing that it is possible to evaluate the epistemic quality of democratic decision-making by the procedure-independent standard (adequate support for conclusion) we shall present in subsequent sections. In doing so, we subscribe to the view Hélène Landemore (2013, 210) calls political cognitivism: the position that 'at least some political questions admit of something like a "better" or "worse" answer.' In other words, there is a standard independent of the decision-making procedure that one can use to assess the result of that procedure. We endorse political cognitivism, but we also acknowledge certain flaws in existing justifications of it – flaws that our proposed method can address, we shall argue.

Political cognitivism comes in two versions: weak and strong (Landemore 2013, 211). Weak political cognitivism builds the epistemic case by arguing that democratic decision-making helps us avoid bad decisions. It refrains from positing a substantive idea about good decisions; rather, it remains agnostic about what a good outcome is, beyond the avoidance of major harm. It helps us evade bad decisions, rather than reach good ones. Weak political cognitivism may hold different ideas about the procedure-independent standard. For instance, David Estlund (2008, chap. IX) suggests a list of 'primary bads', such as war, famine or genocide, that we all agree are undesirable, while James Bohman (2007) suggests that multiperspectival deliberation is epistemically valuable in terms of reducing error, rather than maximizing the truth. By contrast, strong political cognitivism more assertively holds that democratic decision-making helps us reach good decisions by some substantive standard, such as respect for basic rights (Arneson 2003), the maximization of aggregate utility (Mill 1991), the general will (Rousseau 1762) or, more generally, tracking the truth, whatever we take it to be (Landemore 2017 [this issue]).

Now, whatever the merits of these two approaches, they seem to invite precisely the sort of worries sceptics raised, especially if we are looking for a way to turn them into a measure that can be used to assess real-world democratic decision-making outcomes. First, whether we endorse weak or strong political cognitivism, the procedure-independent standard as such seems to involve substantive assumptions about moral or factual correctness, which seem inevitably to stir controversy (Estlund 2008, 99). True, we might sometimes manage to resolve such disagreements about facts or values – Landemore (2013, 213ff) suggests how this could be achieved – but this might seem to jeopardize the instrumentalist epistemic justification of democracy: Before we can assess whether democracy promotes the procedure-independent standard, we have to resolve, by some second-order decision-making procedure, our differences about what that standard is (error avoidance, truth, rights, utility?). Second, there is the 'just off on timing' prob-

lem (Landemore 2013, 212): If the procedure-independent standard is some desirable state of the world that political decisions may or may not produce, it would seem possible to evaluate them only in hindsight. But exactly when should we draw the deadline? How should we weigh short-term against long-term consequences? Given that democracy is an open, on-going process (Saffon and Urbinati 2013, 448), there may be no point in time from which to test the correctness of decisions retrospectively.

The approach we propose here can avoid these problems. While we share the political cognitivist argument against scepticism, we believe existing theories fall short in suggesting viable standards and hands-on methods for testing the epistemic quality of democratic decision-making. The method, which we shall present in further detail in the fourth section, is based on the idea that a necessary-but-insufficient condition for epistemic quality (on any version of political cognitivism) is that concluding decisions are adequately supported by reasons. 'Adequate support for conclusion' provides a procedure-independent standard that can withstand some of the problems other such standards face. First, unlike the primary bads version of weak political cognitivism, our method can assess the epistemic quality of decision-making of a more mundane nature, where any impact on primary bads is so indirect and piecemeal so as to make it practically indistinguishable. Our method also avoids the problem that there may be reasonable disagreement about the causes and injustice of primary bads (Gaus 2011, 293ff). Second, our method is also unaffected by the timing problem: The adequacy with which reasons support the conclusion can be assessed as soon as a decision has been made, without having to take into account possible effects in some near or distant future. Third, the standard we suggest is compatible with a range of ideas about good or bad outcomes, with a range of theories of the epistemic mechanisms at work in democracy (Anderson 2006) and with a range of epistemological theories (Misak 2008). Finally, our approach suggests a hands-on method for empirical research.

While the method needs further specification, it integrates well with existing empirical methods in research on democratic decision-making.

Thus far we have (a) sided with political cognitivism by claiming that it is possible to assess outcomes of democratic decision-making by some procedure-independent standard of epistemic quality and (b) hinted that 'adequate support for conclusion' suggests one such standard that can withstand some of the problems facing more substantive standards of moral or factual correctness or error avoidance. Now, we need to show in greater detail the nature of support adequacy as a procedure-independent standard and what traction it provides on the epistemic aspects of democracy. We shall argue that support adequacy is a necessary but insufficient condition for the epistemic quality of any democratic decision-making procedure.

How does support adequacy matter for political cognitivism? Consider a situation in which democratic decision-makers reach a decision, yet fail to provide any reason whatsoever for it. The decision may still be correct, according to some procedure-independent standard of correctness (or it may have avoided some egregious errors or primary bads). Hence, even if decision-makers provide no support for the conclusion, we cannot rule out that the decision is correct (or, alternatively, non-erroneous). However, in that case, we have no reason to believe the decision to be correct (or non-erroneous), nor to believe that the decision-making procedure is producing correct decisions (or avoiding error). Consequently, if the aim is to make correct decisions (or avoid errors) in a reliable way, providing adequate support for conclusions is essential. Therefore, we shall assume that 'truth-tracking is a stronger aim than just making a factually right decision. It means to make the right decision for the right reasons' (Hartmann and Sprenger 2012, 213). Hence, truth-tracking is a reliable process through which to discover the truth (or the rightness) (List 2004, 526, n.7). You might arrive at the truth by mere chance (say, by flipping a coin), but the process would not be truth-tracking, but rather truth-picking. Alter-

natively, on weak political cognitivism, a strong case for democracy's epistemic superiority should show that it avoids error for the right reasons.

The upshot of this argument is that one can evaluate the truth-tracking or error-avoiding potential of democratic decision-making procedures by examining whether the decisions consist of valid inferences, i.e., whether the conclusions are supported by reason. Hence, support adequacy, we submit, is a necessary condition for truth-trackingness and robust error-avoidance. While it is not entirely uncontroversial, most theorists accept support adequacy as a necessary condition for the rational credibility of a claim or belief (Pinto 2001:37). However, it is not a sufficient condition, since a perfectly valid conclusion can still be false if the premises are false. The approach we suggest here does not involve testing whether the premises are also correct or plausible, as any such tests would lead us back to the problem of disagreements over external standards of moral or factual correctness and, in particular, specific measurements of such standards, but in the final section, we shall suggest some strategies for how to deal with this limitation of the support adequacy approach. Let us conclude, however, by noting that any presumptive method for assessing decision-making according to some procedure-independent standard of moral or factual correctness (or incorrectness) would also have to involve an assessment of whether decisions are based on valid inferences, at least if it is to assess whether democratic decision-making track the truth for the right reasons, rather than by mere chance.

Support adequacy is also an insufficient criterion in another sense: It is only a measure of epistemic quality, not an exhaustive standard by which to assess political decision-making. For example, there might be circumstances where managing to find a compromise is more important than to provide reasons for decisions (e.g., a so-called 'under-theorized agreement'). Sometimes decisions that would score low on support adequacy may perform well on other standards. Support adequacy does not suggest a notion of general desirability, but by capturing one important, epistemic aspect of decision-making, it may help researchers explore correlations

between different standards of decision-making quality (for a similar view, see Landemore 2017 [this issue]).

2.1 Support adequacy and competing accounts of epistemic democracy

Support adequacy may provide useful resources for evaluating the epistemic performance of different democratic decision-making procedures. To illustrate this utility, consider two prominent, competing arguments for the epistemic value of democracy. On the aggregative approach, which follows the logic of the Condorcet Jury Theorem, a group best tracks the truth by aggregating individual judgements (List and Goodin 2001). Thus, the procedure must secure that individual voters reach their judgements independently of each other. Some even claim that public interaction may actually lead to worse decisions than would otherwise have occurred (Schkade, Sunstein, and Hastie 2010). By contrast, the deliberative approach assumes better decisions to follow from a process of public deliberation among individuals who must be ready to be swayed by the 'forceless force of the better argument' (Habermas 1971, 131).

With its focus on public reason-giving as the key causal mechanism producing epistemically better decisions, deliberative democratic theory already entails an expectation that deliberation increases the adequacy of support for conclusion – after all, this is what it means for the better argument to triumph. Support adequacy may seem to matter less, however, for the aggregative approach: On this approach, the truth is tracked by aggregating individual, independent judgments on the conclusion, but the mechanism by which democratic decision-making may outperform alternative procedures does not depend as directly on argumentative soundness as does the deliberative approach. Moreover, given that individuals might endorse the conclusion for quite different reasons, the level of support reasons offer for the conclusion is irrelevant for assessing whether the decision-making procedure tracks the truth. In practice, when decisions are made by aggregating individual judgements, participants need not issue a joint statement of

their decision and its supporting reasons (though this problem also affects deliberative polls – see Thompson 2008, 503). Consequently, the support adequacy standard may seem less applicable to an aggregative procedure of democratic decision-making.

While this concern points to a potential limitation of the framework we propose here, we believe support adequacy is relevant for the aggregative procedure too, for three reasons. First, in actual democratic practice, decisions tend to be choices between reasoned policy alternatives. Even if a decision-making procedure does not engage individuals in actual public deliberation prior to the decision (which they normally do), the choice they face is typically not just between different policy options, but also between different justifications for and against those options. Second, support adequacy is relevant for the aggregative approach because, as we have already argued, in order to be truth-tracking rather than just truth-picking, the procedure must arrive at the better decision for the right reasons. If the aggregation of judgements failed to produce decisions that are supported by reasons, it would meet serious objections in terms of democratic legitimacy. Third, while the aggregative approach may just axiomatically assume that citizens are more likely than chance to arrive at the truth, any explanation of that assumption must point to some capacity for reasoning or 'deliberation within' (Goodin 2000; Goodin and Niemeyer 2003; Landemore and Mercier 2012). Indeed, precisely this innate human capacity for rational reflection may be distorted through public debate, some critics of the deliberative approach argue. Hence, both approaches may assume an individual capacity for reasoning, and disagree about whether it is best put to collective, epistemic use through aggregation of independent judgements or through public deliberation. Therefore, our support adequacy framework applies, in theory, to both approaches; in practice, how well it applies depends on the research design. Given that the deliberative approach has been more preoccupied with reason-giving aspects of decision-making, this is also where we find the most developed ideas about how to measure the quality of public discourse, to which we now turn.

3. Existing methods for measuring the quality of deliberative outcomes

As deliberative democracy has matured as an empirical research programme, the methods for testing its key claims have become increasingly diverse and sophisticated (for overviews, see Thompson 2008; Black et al. 2011; Beste 2013; Neblo 2007). Yet for all the talk about reason-giving, argumentation and justification, empirical research on deliberative democracy has mostly focused on other aspects of public deliberation than whether reasons given adequately support the conclusions when groups debate and decide on matters of public policy. This neglect is unfortunate, because it means that the deliberative research programme lacks empirical methods for assessing a key claim in deliberative democratic theory: that public deliberation improves the practical rationality of democratic outcomes (Cooke 2000, 952).

To explore this lacuna in existing research, let us consider the Discourse Quality Index (DQI) (Steiner et al. 2004), a sophisticated coding scheme widely employed in empirical research on deliberation in parliaments and other arenas (e.g. Lord and Tamvaki 2013; Lord 2013; Himmelroos and Christensen 2014; Monnoyer–Smith and Wojcik 2012; Roger and Schaal 2013). On the DQI, one parameter, the 'level of justification', addresses the question of reasons offered for conclusions, which is evidently central to the truth-tracking function of deliberation (Steenbergen et al. 2003). The level of justification 'refers to the nature of the justification of demands'; it aims to 'judge to what extent speech gives complete justification of demands', and, similar to our approach, focuses on inferences, as the 'completeness of justification is judged in terms of the inferences that are made' (Steenbergen et al. 2003, 28).

The DQI proposes four levels of justification:

0. *No justification*: A speaker only says that X should or should not be done, but no reason is given.
1. *Inferior justification*: Here a reason Y is given as to why X should or should not be done, but no linkage is made between X and Y — the inference is incomplete. This code also applies if a conclusion is merely supported with illustrations.
2. *Qualified justification*: A linkage is made as to why one should expect that X contributes to or detracts from Y. A single such complete inference already qualifies for code 2.
3. *Sophisticated justification*: Here, at least two complete justifications are given, either two complete justifications for the same demand or complete justifications for two different demands.

Now, as a measure of whether conclusions are adequately supported by reasons (in order to determine whether viable conditions for truth-tracking exist), the DQI parameter 'level of justification' falls short, and admittedly, it was designed for other purposes. To see why this approach is insufficient for measuring the adequacy of support, let us consider the DQI's illustration of what should be classified as a complete justification (code 2):

> 'Does the hon. Lady [Caroline Spelman, Conservatives, Meriden] agree that there is a further point on the separate taxation of men and women? Women who are abused in the household sometimes find it difficult to get away from the home. Separate taxation helps women to have the courage to move out of an abusive household' (Steenbergen et al. 2003, 32).

The authors interpret this example as providing a complete justification (code 2) for the demand that women and men be taxed separately: 'such policy would give women the courage to escape from an abusive household' (Steenbergen et al. 2003, 32). From our point of view, the im-

portant question is whether the conclusion (that there should be a separate taxation of women and men) is sufficiently supported by reasons. To answer that question we need to reconstruct this inference:

1. Women who are abused in the household sometimes find it difficult to move out.
2. Separate taxation encourages women to move out of an abusive household.
3. Therefore, there should be separate taxation of men and women.

Formally, this inference obviously lacks a normative principle (p) about the desirability of promoting women to 'move out of an abusive household', but we can assume that such an uncontroversial principle implicitly exists as a premise. However, even if we include some such normative principle (p) as a premise in the inference, the conclusion lacks support. Even if we grant that the premise (2) is true and adheres to the uncontroversial implicit normative premise (p), the conclusion (3), that there should be separate taxation of men and women, does not follow. The content of the premises certainly points in that direction, but the support is quite weak. For example, let us assume that such a law would indeed help women move out of abusive households, but also that the effects were quite small, and let us further assume that the law would also have other effects that would threaten other desirable values. Then the conclusion would not be supported, all things considered. Consequently, as long as premises are lacking which state that other effects would not be bad enough to undermine the desirability of separate taxation, the conclusion to enact such a law is only weakly supported (cf. Walton 2011).

Since the DQI considers a conclusion that is only weakly supported by the premises to be complete, it cannot tell us whether inferences are maximally truth-tracking in the sense of having premises which support the conclusion. Neither does the third level – the 'sophisticated justification' – add any stronger criterion for when to consider the conclusion to be supported. The difference between the second and the third level is only numerical: at the second level, the speaker offers just one constellation of premises to justify the conclusion, while s/he offers at

least two constellations of premises at the third level. Hence, the level of support for conclusions can be quite low even in 'sophisticated justifications'.

Besides the DQI, other noteworthy micro-analytic approaches for measuring the quality of the content of deliberation (Bächtiger et al. 2009) come some way in mapping arguments (e.g. Nanz and Steffek 2005; Dutwin 2003), but likewise lack tools for measuring whether they are sound and valid. For instance, the 'argument repertoire' (Cappella, Price, and Nir 2002), includes both the relevant reasons one has for a position and the reasons one knows of for the opposite position, while Stromer-Galley's (2007) coding scheme for content analysis of deliberation has one component ('reasoned opinion expression') that seeks to capture whether participants elaborate their views, by providing, for instance, a justification, a reason, a definition or a fact, etc. However, as an operationalization of the epistemic quality of deliberative outcomes, these approaches fail for similar reasons as the DQI: The numerical measures cannot distinguish whether the reasons sufficiently support the conclusion. Hence, the most advanced approaches for measuring deliberative reasoning seem insufficient for assessing a key aspect of epistemic quality. In order to close this gap in existing methods for assessing discourse quality, it seems fruitful to consult scholarly literatures in argumentation theory and informal logic (for a similar point, see Knops 2006), which specifically address the question of how to determine whether a conclusion is adequately supported by reasons.

4. Assessing the adequacy of support for conclusions

Now, we propose that one could capture an essential aspect of epistemic quality in democratic decision-making by assessing whether reasons adequately support the conclusion. Hence, we wish to complement the numerical approaches counting the number of justifications speakers offer for a demand with an approach that also tests whether their justifications are in fact good ones. In this section, we draw on an epistemological theory of argumentation in order to suggest

an analytical framework for assessing premise support in three generic types of arguments: deductive, inductive and practical argument schemes.

The epistemological approach to argumentation (Feldman 1993; Lumer 2011; Lumer 2005; Biro and Siegel 2006; Aikin 2008) offers a compelling theoretical explanation of what makes an argument valid, and it can helpfully be distinguished from rival accounts of argumentation by the way it views the function of argumentation: Rhetorical theories assume that argumentation aims at persuading an addressee, while consensus theories, e.g., pragma-dialectics (Eemeren and Grootendorst 2004), assume that argumentation serves to reach shared beliefs through a discursive process. Epistemological theories of argumentation, by contrast, assume that the purpose of argumentation is to rationally convince – to produce rationally justified belief, i.e., knowledge.

An advantage of the epistemological approach, not least for our purposes, is that it offers criteria for assessing an argument's epistemic value. The epistemological approach suggests that in order to convince rationally, arguments must rely on established principles of epistemology, which both help explain what makes a belief rationally justified and provide criteria for assessing it (Lumer 2000; Lumer 2011). These epistemological principles, along with acceptability criteria developed in logic, probability theory and prudential decision or desirability theory, also suggest an exhaustive typology of generic classes of argument (Lumer 2011): Deductive, inductive and practical argument schemes. For each type of argument scheme, inadequate premise support for the conclusion can take different forms, but we can distinguish two distinct, generic ways in which the premises of an inference can fail to support the conclusion: Either the argument scheme as such fails to convince because it rests on an inadequate epistemological principle, or it fails because the arguer applies the scheme in a distorted way. Now, let us look closer at how these failures may appear in deductive, inductive and practical argument schemes.

Deductive argument schemes follow the rules of formal logic. According to formal logic, the premises support a conclusion if the conclusion follows necessarily from the premises (Johnson 1999; Johnson and Blair 1994). Hence, the inference is truth-preserving since the conclusion is true if the premises are true. Applied correctly, a deductive argument scheme is therefore per se effective, as anyone who accepts the premises must also accept the conclusion. The first type of failure is therefore an empty category since truth-preserving inferences are always epistemologically adequate. Needless to say, though, deductive argument schemes are not always sound since a deductively valid inference may contain false premises. Moreover, applied in a distorted way, a deductive argument scheme leads to a *non-sequitur*, i.e. a formal fallacy where the premises fail to support the conclusion, such as 'affirming the consequent' or 'denying the antecedent'.

Inductive argument schemes rest on a probabilistic principle in which the conclusion is inferred on the basis of statistics. Unlike deductive reasoning, inductive reasoning is defeasible: The conclusion can certainly be wrong even though the premises are true, which makes inductive validity less strict as a necessary truth-tracking condition. An example of an inductive argument schemes that apparently mobilizes a probability in support of the conclusion is 'argument from cause to effect', which predicts that one type of event tends to cause another (Walton, Reed, and Macagno 2008, 168). Many other kinds of argument schemes are also warranted by a probabilistic principle. For instance, in 'argument from expert opinion', the validity of the conclusion depends on the probability that the cited expert is correct and is telling the truth (Lumer 2011, 6). For inductive argument schemes, inadequate support for conclusion may result either if the argument scheme is applied incorrectly (i.e., it fails to meet the probabilistic warranting principle) or if the scheme is ineffective per se. A common example of the former error would be to make generalizations on the basis of an unrepresentative sample. An example of the latter type of failure would be 'argumentum ad consequentiam', according to which an event will not

happen because it has negative consequences (Lumer 2000, 414), and its positive counterpart, 'wishful thinking', which both rest on an inadequate epistemological principle.

Practical argument schemes, finally, are arguments intended to justify and guide action. In contrast to deductive and inductive argument schemes, practical argument schemes do not concern factual but normative matters: What should I, or we, do? They are supported on a normative logic according to which the desirability of different actions is determined. As with inductive schemes, practical argument schemes can fail to provide premise support in two ways: because the argument scheme is ineffective per se and/or because the arguer uses an otherwise acceptable scheme in a distorted way. Exemplifying the first type of failure, 'argument from waste', i.e., an argument that a course of action should not be abandoned because of sunk costs, rests on a doubtful practical principle, at least in contexts where a pure cost–benefit rationality is relevant (Walton 2002). As for distorted applications of practical argument schemes, consider again the separate taxation example from the DQI mentioned above. The inference is a case of the practical argument scheme 'argument from consequences' and in order for its premises to offer sufficient support for the conclusion, they should consider other consequences of bringing about some state of affairs, i.e., possible negative effects of separate taxation relevant to the conclusion. Lacking such considerations, the use of the argument scheme is applied in a distorted way and the inference would not be truth-tracking.

Offering a theoretically grounded explanation of what makes a good argument, this approach explicates two necessary but insufficient conditions for making truth-tracking inferences: It should (a) employ an argument scheme that rests on an epistemologically adequate principle in (b) a correct, undistorted manner. Assuming that this typology can be operationalized into a concrete empirical method, it may complement the DQI with more accurate and less arbitrary criteria for determining the support adequacy of arguments in political discourse. Moreover, by focusing on arguments – rather than on arguers, audiences, and their interaction – the episte-

mological approach to argumentation also allows us to offer a procedure-independent method for assessing the epistemic quality of a variety of decision-making procedures and their output. The unit of analysis could be either individual statements presented in a more or less deliberative context preceding a decision (e.g. speeches in parliament, as per the DQI, or bills and motions), or statements resulting from a collective a decision-making process (e.g. court judgments). By comparison, attempts at applying pragma-dialectics to assess the quality of democratic decision-making (Knops 2006; Curato 2012) presuppose a sequential, dialogical situation that may be absent in some decision-making settings.

A method for assessing support adequacy raises important issues of interpretation. As is standard in argumentation analysis, the approach we suggest here would proceed in two steps: First, one seeks to reconstruct the inferences in a way that corresponds as closely as possible to the way the arguer *de facto* made the inference, identifying explicitly stated conclusions and premises. However, since the aim is to present the inference as the arguer *de facto* made it, one must also consider that some premises may be left unstated, yet not make the argument better than the arguer intended. Here, the distinction between 'used' and 'needed' premises offers some help (Ennis 1982; Walton and Reed 2005, 341): 'Used' premises are those that the arguer meant to include in the argument, but left unstated, while 'needed' premises are propositions that make the argument rational or sound, regardless of whether the arguer intended to include them or not. In the reconstruction phase, only 'used' premises should be added, while 'needed' premises are relevant in the evaluation phase.

4.2 Operationalizing support adequacy and argumentative errors

Thus far, we have argued that it is possible in theory to hold democratic decision-making to a procedure-independent standard of epistemic quality, and that support adequacy forms an essential part of that standard. However, in order to be useful for empirical research, the standard should preferably produce a quantifiable measure. Let us now consider how one could opera-

tionalize support inadequacy for the three types of argument schemes and the two kinds of argumentative errors we described in the previous section.

For deductive argument schemes, there is just one kind of error: distorted application, which results in a formal fallacy. This is a dichotomous phenomenon: Either the argument is deductively valid or it is not. For inductive and practical argument schemes, argumentative errors can result either because the argument scheme rests on a flawed epistemological principle, or because it applies an otherwise acceptable argument scheme in a distorted way. Both of these kinds of error are matters of degree, rather than dichotomous. For both the inductive and practical type, various argument schemes have varying epistemic value depending on the epistemological principle at work. Although it seems impossible to provide an exhaustive, transitive ranking of individual argument schemes in terms of their epistemological warrant, we may still classify them according to their epistemic strength as, say, strong, medium or weak schemes.

For instance, as an example of a strong scheme, the inductive argument scheme 'argument from cause to effect' is directly warranted by a probabilistic principle. Other inductive argument schemes may be considered ceteris paribus somewhat weaker because they are more indirect or mediated, such as 'argument from expert opinion': In order to assess the argument, we must evaluate two probabilities: Is the cited expert likely to be correct and to truthfully reveal their opinion? Yet other argument schemes rest on epistemological principles that generally have low epistemic value, such as 'argumentum ad hominem', where an argument is rejected on the basis of the bad character of the arguer: While it might sometimes be relevant to consider the character of the arguer in deciding the epistemic value of his/her argument, usually it is not (Walton, Reed, and Macagno 2008, 142).

Likewise, practical argument schemes, which aim to guide action, can be classified according to the strength of the epistemological principle warranting the inference. For instance, 'argument from consequences' seems to have a stronger epistemological warrant than 'argument

from threat', since the former, unlike the latter, considers not only potential negative outcomes, but also potential positive effects. Finally, 'argument from waste' may represent an epistemologically even weaker argument scheme, which in most situations commits a 'sunk costs fallacy'. This indicates a rough classification of argument schemes according to their epistemic quality that could be useful when turning the framework into a concrete method. The weaker the argument scheme used, the lower the epistemic potential of the argument.

The second type of error, i.e., where an argument scheme is applied in a distorted way is also a matter of degree: The greater the distortion, the lower the epistemic value of the argument. For instance, the more the argument leaves out premises necessary to adequately support the conclusion, the lower its epistemic value. Empirical research in informal logic has identified some 60 different argument schemes, along with criteria that premises should meet in order to provide support for the conclusion (Walton, Reed, and Macagno 2008). Hence, running through the set of critical questions specific to each argument scheme may help to determine the degree of distortion in the way the argument scheme is being used.

Of course, to classify arguments in terms of the strength of the used scheme's epistemological warrant and how well the scheme is applied requires qualitative evaluation, and it is hard to say, for example, exactly how much better 'argument from consequences' is compared to 'argument from threat'. There is no objective, natural numerical scale that corresponds to these qualitative differences. Still, as long as researchers quantify in a systematic, unitary way, they may be able to generate comparative measures of whether reasons adequately support conclusions based on this approach, and, in turn, to capture an important aspect of epistemic quality of democratic decision-making.

5. Challenges and objections

In this section, we respond to three challenges for our methodological framework: that it is meaningless because of its indifferentism to the plausibility of premises; that the method is too labour-intensive to be useful for empirical research; and that the approach fails as an evaluation of the epistemic performance of democracy writ large.

5.1 Can you test truth-tracking without evaluating the plausibility of premises?

First, can an approach for assessing whether decision-making tracks the truth do without evaluating truth-claims? If it focuses solely on validity, the approach would fail to discriminate between valid inferences with more and less plausible premises. First, consider, the following inference:

1. All men are mortal.
2. Socrates is a man.
3. Therefore, Socrates is mortal.

This inference is deductively valid and the premises are very plausible; thus, it is highly truth-tracking. In contrast, consider now this (in)famous inference:

1. All men are immortal.
2. Socrates is a man.
3. Therefore, Socrates is immortal.

This inference is also deductively valid, but, needless to say, it contains a highly implausible premise. However, as our theoretical framework implies indifferentism with regard to the truth of premises, it would not differentiate between these inferences with regard to their truth-trackingness since they are equally valid. Importantly, our approach would not accept false posi-

tives (cf. Estlund 2008, 113), since it would not claim that a valid inference with implausible premises is truth-tracking; as we already stated above, the method does not allow for such affirmative judgments. Nevertheless, it might be considered problematic enough if the framework cannot differentiate between these two inferences with regard to their truth-trackingness. To put this objection frankly: Is a method that cannot differentiate between an obviously true inference and a patently absurd one, with regard to their truth-trackingness, relevant at all?

The indiscrimination objection is an important challenge for our approach. Yet, we believe that the problems are mostly of a practical nature, not a principal error. There is nothing inherently wrong with the method's indifferentism to truth-claims; it does not suggest that a valid inference with implausible premises is truth-tracking, but only that the truth-trackingness of such an inference is not invalidated by lack of support for premises. Consequently, the indiscrimination objection actually concerns the method's practical utility. In most actual cases, the approach would be able to handle the objection, as the severity of the problem largely depends on how one employs the method in a research design. Empirical testing of democracy's epistemic potential will likely involve studies of deliberation in either small group experiments or real-world decision-making arenas (Friberg-Fernros and Schaffer 2014), and so we shall suggest how each of these research designs could tackle the indiscrimination problem.

First, in small-group experiments, one can simply avoid questions about the plausibility of the premises altogether by instructing the participants in the experiment to assess the premise support and to withhold their judgments about the plausibility of the premises. Therefore, an experimental study can be designed in a way that renders the indiscrimination objection irrelevant. Second, in studies of real-world deliberation the indiscrimination problem becomes more urgent because the researchers can control neither the inferences that participants use nor the considerations they address. One way to deal with the problem is simply to ignore the truth-trackingness of inferences with regard to the plausibility of premises, and make reservations that

the method does not rule out the risk that some inferences that pass the test may contain implausible premises.

A more assertive strategy for mitigating the indiscrimination problem in studies of real-world deliberation would be to extend the analysis by not just testing inferences one by one, but also to compare the content of different kinds of inferences. Cross-inference consistency analysis could reveal inconsistencies between inferences, which in turn may allow for assessing the plausibility of premises without actually testing their truth-claims externally (Friberg-Fernros and Brommesson 2013). For instance, if an actor in one inference rests on the premise that Socrates is immortal, and in another inference on the premise that Socrates is mortal, then the actor would be guilty of inconsistency and obviously not truth-tracking. Furthermore, the more implausible the premises, the more likely they are to conflict with premises in other inferences, and so this strategy may help to identify inferences with implausible premises.

The cross-consistency strategy assumes that inferences should be substantially consistent with each other, an assumption which seems reasonable in many situations (Ottonelli 2010); for instance, we expect the law to be coherent in this sense (Dworkin 1986). However, in some situations there are no parallel inferences across which to examine consistency, as for example in the DQI illustration cited above. In such cases, one way to mitigate the indiscrimination problem is to trace the inferences behind the premises. Inferences are rarely presented in isolation; rather, they are embedded in a wider web of argumentation, which, for example, aims to defend the specific content of the inferences. In other words, inferences are being made in support of other inferences. Here, it might be possible to test the support for conclusions in inferences that aim to support the premises in the main conclusion. By assessing these underlying inferences, one could differentiate between valid inferences with, respectively, plausible and implausible premises, as implausible premises are likely to be harder to defend than plausible ones. In theory, there might be cases where there is only one isolated inference to analyse, but in real-world

political deliberation such cases are too rare to invalidate the relevance of our methodological framework.

Finally, the framework's indifferentism to truth claims also offers certain advantages, as it makes the method compatible with a wider range of epistemic theories of democracy, including correctness theories (Estlund 2008, 103ff), and a variety of related epistemological theories (Misak 2008). Any attempt at evaluating the plausibility of premises would lead back to substantive standards of correctness, and limit the method's applicability correspondingly.

5.2 Isn't this approach too labour-intensive to be practically useful?

A second challenge for our methodological framework concerns its feasibility as a method for testing the epistemic quality of democratic outcomes on a mass scale. Even if one accepts that the methodology, in principle, would make it possible to reconstruct inferences and evaluate whether premises adequately support conclusions, the method would be too cumbersome to be practically useful for testing the truth-trackingness of democratic procedures and decisions, or for any other empirical purpose.

Since we claim to present a theoretical framework that could provide a foundation for hands-on methods useful for empirical research, this objection is important. However, the problem is not unique for our approach; all forms of manual content analysis face similar problems and trade-offs between accuracy and feasibility (Black et al. 2011). As for the reconstruction phase, recent advances in research on argumentation may provide a useful path forward. Argumentation schemes have been used as an analytical tool in empirical studies of argumentation in diverse contexts, e.g. artificial intelligence (Bench-Capon and Dunne 2007), science education (Jiménez-Aleixandre and Erduran 2007), and legal reasoning (Verheij 2003). The process of reconstructing naturally occurring arguments can also be facilitated by various tools for argumentation mining (Mochales and Moens 2011). And while the evaluation phase, i.e. the actual assessment of premise support adequacy, must always involve possibly time-consuming interpre-

tation and contextual sensitivity (Pinto 2001, 27), the method we have suggested here provides guidelines that will make the evaluation phase both easier and more transparent.

5.3 How does this indicator help us assess the epistemic performance of democracy as a system?

Finally, critics may worry that the approach we have presented here might perhaps be useful in assessing an essential aspect of epistemic quality of arguments made in democratic decision-making procedures. It may fall short, however, if we wish to assess how well democracy performs as a complex, dispersed system with emergent properties, since the epistemic quality of democracy at the systems level is more than the mere sum of sound arguments in particular instances (Mansbridge et al. 2012, 5f). This is a reasonable concern, especially given recent interest in systemic approaches to democracy, which seek to move beyond the narrow theorising of discrete deliberative initiatives to re-focus, instead, on how the public sphere relates to the state (Chambers 2009; 2017 [this issue]).

However, we make a valuable contribution per se by providing a way to assess the epistemic merits of democratic decision-making in specific institutions, even if epistemic quality at the systems level cannot be measured by simply aggregating from such cases. For instance, there is a dynamic research programme on democratic deliberation in various settings (e.g., parliaments or micro-publics) that has thus far lacked methods for testing a key epistemic claim in deliberative democratic theory – a gap our approach attempts to bridge.

Yet even for research that focuses more on democracy as a complex system, it is crucial to assess how specific, discrete initiatives or institutions contribute to the system's epistemic function, which consists just in 'producing preferences, opinions and decisions that are appropriately informed by facts and logic and are the outcome of substantive and meaningful consideration of relevant reasons' (Mansbridge et al. 2012, 10). Thus, the epistemic functions performed by the components, which include delivering decisions adequately supported on reasons, are central for understanding how a complex deliberative system works. Hence, the approach we have suggest-

ed here offers resources for assessing the epistemic performance of both discrete instances of deliberation and the systems they comprise.

6. Conclusion

The epistemic turn in democratic theory yearns for ways to evaluate the epistemic performance of different democratic decision-making procedures, but such methods have been lacking, presumably because many believe that assessing epistemic quality implies that one must employ some inherently controversial procedure-independent substantive standard of moral or factual correctness.

In this paper, we have presented and defended a procedure-independent way to assess the epistemic quality of democratic decisions that steers clear of substantive standards of correctness, by assessing, instead, whether conclusions are adequately supported by reasons. Adequate support for conclusion is a necessary feature of epistemic quality, whether one understands epistemic quality to consist only in the avoidance of erroneous decisions or in some more substantive notion of desirable features. To the extent that democratic decision-making procedures produce decisions that are inadequately supported by reasons, we can fairly say that they are also lacking an important aspect of epistemic quality.

This paper has demonstrated the fruitfulness of connecting theoretical literatures on epistemic democracy and empirical methods for assessing discourse quality with the toolbox of argumentation theory. Now, while this paper has aimed to offer proof of concept, the ultimate proof is in the pudding: The framework we suggest here needs to be elaborated into an operational method and tested through empirical studies.

Acknowledgements

For their comments on earlier versions of this paper, we thank Svenja Ahlhaus, Sveinung Arnesen, Simone Chambers, Felix Gerlsbeck, Peter Hallberg, Andrew Knops, Helene Landemore, Olof Leffler, Peter Niesen, Valeria Ottonelli, Markus Patberg, Sean Sinclair and Daniele Santoro. We have also benefited greatly from discussing some of these ideas with Christoph Lumer and Douglas Walton.

References

Aikin, Scott F. 2008. "Three Objections to the Epistemic Theory of Argument Rebutted." *Argumentation and Advocacy* 44: 130–42.

Anderson, Elizabeth. 2006. "The Epistemology of Democracy." *Episteme* 3 (1–2): 8–22. doi:10.3366/epi.2006.3.1-2.8.

Arneson, Richard J. 2003. "Defending the Purely Instrumental Account of Democratic Legitimacy." *Journal of Political Philosophy* 11 (1): 122–32.

Bächtiger, André, Susumu Shikano, Seraina Pedrini, and Mirjam Ryser. 2009. "Measuring Deliberation 2.0: Standards, Discourse Types, and Sequentialization." In *5th ECPR General Conference*. Potsdam. http://www.scribd.com/doc/63594625/Measuring-Deliberation-2-0.

Bench-Capon, T.J.M., and Paul E. Dunne. 2007. "Argumentation in Artificial Intelligence." *Artificial Intelligence* 171 (10–15): 619–41. doi:10.1016/j.artint.2007.05.001.

Beste, Simon. 2013. "Contemporary Trends Of Deliberative Research: Synthesizing A New Study Agenda." *Journal of Public Deliberation* 9 (2). http://www.publicdeliberation.net/jpd/vol9/iss2/art1.

Biro, John, and Harvey Siegel. 2006. "In Defense of the Objective Epistemic Approach to Argumentation." *Philosophy Articles and Papers*, January. http://scholarlyrepository.miami.edu/philosophy_articles/13.

Black, Laura W., Stephanie Burkhalter, John Gastil, and Jennifer Stromer-Galley. 2011. "Methods for Analyzing and Measuring Group Deliberation." In *Sourcebook for Political Communication Research: Methods, Measures, and Analytical Techniques*, edited by E. Page Bucy and R. Lance Holbert, 323–45. Communications Series. New York: Routledge.

Bohman, James. 2007. "Political Communication and the Epistemic Value of Diversity: Deliberation and Legitimation in Media Societies." *Communication Theory* 17 (4): 348–55. doi:10.1111/j.1468-2885.2007.00301.x.

Caluwaerts, Didier, and Juan Ugarriza. 2012. "Favorable Conditions to Epistemic Validity in Deliberative Experiments: A Methodological Assessment." *Journal of Public Deliberation* 8 (1). http://www.publicdeliberation.net/jpd/vol8/iss1/art6.

Cappella, Joseph N., Vincent Price, and Lilach Nir. 2002. "Argument Repertoire as a Reliable and Valid Measure of Opinion Quality: Electronic Dialogue During Campaign 2000." *Political Communication* 19 (1): 73–93. doi:10.1080/105846002317246498.

Chambers, Simone. 2009. "Rhetoric and the Public Sphere Has Deliberative Democracy Abandoned Mass Democracy?" *Political Theory* 37 (3): 323–50. doi:10.1177/0090591709332336.

———. 2017. "Balancing Epistemic Quality and Equal Participation in a System Approach to Deliberative Democracy." *Social Epistemology*.

Cooke, Maeve. 2000. "Five Arguments for Deliberative Democracy." *Political Studies* 48 (5): 947–969. doi:10.1111/1467-9248.00289.

Curato, Nicole. 2012. "A Sequential Analysis of Democratic Deliberation." *Acta Politica* 47 (4): 423–42. doi:10.1057/ap.2012.15.

Dutwin, David. 2003. "The Character of Deliberation: Equality, Argument, and the Formation of Public Opinion." *International Journal of Public Opinion Research* 15 (3): 239–64. doi:10.1093/ijpor/15.3.239.

Dworkin, Ronald. 1986. *Law's Empire*. Harvard University Press.

Eemeren, Frans H. van, and Rob Grootendorst. 2004. *A Systematic Theory of Argumentation: The Pragma-Dialectical Approach*. Cambridge University Press.

Ennis, Robert H. 1982. "Identifying Implicit Assumptions." *Synthese* 51 (1): 61–86. doi:10.1007/BF00413849.

Estlund, David. 2008. *Democratic Authority : A Philosophical Framework*. Princeton N.J.: Princeton University Press.

Feldman, Richard. 1993. *Reason and Argument*. Englewood Cliffs N.J.: Prentice Hall.

Friberg-Fernros, Henrik, and Douglas Brommesson. 2013. "The Responsibility to Protect – An Incoherent Doctrine?" *International Politics* 50 (4): 600–622. doi:10.1057/ip.2013.16.

Friberg-Fernros, Henrik, and Johan Karlsson Schaffer. 2014. "The Consensus Paradox: Does Deliberative Agreement Impede Rational Discourse?" *Political Studies* 62 (S1): 99–116. doi:10.1111/1467-9248.12103.

Gaus, Gerald. 2011. "On Seeking the Truth (Whatever That Is) through Democracy: Estlund's Case for the Qualified Epistemic Claim." *Ethics* 121 (2): 270–300. doi:10.1086/658141.

Goodin, Robert E. 2000. "Democratic Deliberation within." *Philosophy & Public Affairs* 29 (1): 81–109.

Goodin, Robert E, and Simon Niemeyer. 2003. "When Does Deliberation Begin? Internal Reflection versus Public Discussion in Deliberative Democracy." *Political Studies* 51 (4): 627–49. doi:10.1111/j.0032-3217.2003.00450.x.

Habermas, Jürgen. 1971. "Vorbereitende Bemerkungen Zu Einer Theorie Der Kommunikativen Kompetenz." In *Theorie Der Gesellschaft Oder Sozialtechnologie: Was Leistet Die Systemforschung?*, by Niklas Luhmann and Jürgen Habermas, 101–41. Frankfurt am Main: Suhrkamp.

Hartmann, Stephan, and Jan Sprenger. 2012. "Judgment Aggregation and the Problem of Tracking the Truth." *Synthese* 187 (1): 209–21. doi:10.1007/s11229-011-0031-5.

Himmelroos, Staffan, and Henrik Serup Christensen. 2014. "Deliberation and Opinion Change: Evidence from a Deliberative Mini-Public in Finland." *Scandinavian Political Studies* 37 (1): 41–60. doi:10.1111/1467-9477.12013.

Ingham, Sean. 2013. "Disagreement and Epistemic Arguments for Democracy." *Politics, Philosophy & Economics* 12 (2): 136–55. doi:10.1177/1470594X12460642.

Jiménez-Aleixandre, María Pilar, and Sibel Erduran. 2007. "Argumentation in Science Education: An Overview." In *Argumentation in Science Education*, edited by Sibel Erduran and María Pilar Jiménez-Aleixandre, 3–27. Science & Technology Education Library 35. Springer Netherlands. http://link.springer.com/chapter/10.1007/978-1-4020-6670-2_1.

Johnson, Ralph. 1999. "The Relation between Formal and Informal Logic." *Argumentation* 13 (3): 265–74. doi:10.1023/A:1007789101256.

Johnson, Ralph, and J. Anthony Blair. 1994. *Logical Self-Defense*. U.S. ed. New York: McGraw-Hill.

Knops, Andrew. 2006. "Delivering Deliberation's Emancipatory Potential." *Political Theory* 34 (5): 594–623. doi:10.1177/0090591706290780.

———. 2017. "Validity and Scope as Criteria for Deliberative Epistemic Quality across Pluralism." *Social Epistemology*.

Landemore, Hélène. 2013. *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton ; Oxford: Princeton University Press.

———. 2017. "Beyond the Fact of Disagreement? The Epistemic Turn in Deliberative Democracy." *Social Epistemology*.

Landemore, Hélène, and Hugo Mercier. 2012. "Talking It out with Others vs. Deliberation within and the Law of Group Polarization: Some Implications of the Argumentative Theory of Reasoning for Deliberative Democracy." *Análise Social* 47 (205): 910–34.

List, Christian. 2004. "On the Significance of the Absolute Margin." *The British Journal for the Philosophy of Science* 55 (3): 521–44. doi:10.1093/bjps/55.3.521.

List, Christian, and Robert E. Goodin. 2001. "Epistemic Democracy: Generalizing the Condorcet Jury Theorem." *Journal of Political Philosophy* 9 (3): 277–306. doi:10.1111/1467-9760.00128.

Lord, Christopher. 2013. "No Representation without Justification? Appraising Standards of Justification in European Parliament Debates." *Journal of European Public Policy* 20 (2): 243–59. doi:10.1080/13501763.2013.746123.

Lord, Christopher, and Dionysia Tamvaki. 2013. "The Politics of Justification? Applying the 'Discourse Quality Index' to the Study of the European Parliament." *European Political Science Review* 5 (1): 27–54. doi:10.1017/S1755773911000300.

Lumer, Christoph. 2000. "Reductionism in Fallacy Theory." *Argumentation* 14 (4): 405–23. doi:10.1023/A:1007809709996.

———. 2005. "The Epistemological Theory of Argument – How and Why?" *Informal Logic* 25 (3). http://ojs.uwindsor.ca/ojs/leddy/index.php/informal_logic/article/view/1135.

———. 2011. "Argument Schemes—an Epistemological Approach." In *Argumentation. Cognition and Community*, edited by Frank Zenker.

Mansbridge, Jane, James Bohman, Simone Chambers, Thomas Christiano, Archon Fung, John Parkinson, Dennis Thompson, and Mark Warren. 2012. "A Systemic Approach to Deliberative Democracy." In *Deliberative Systems : Deliberative Democracy at the Large Scale*, edited by John Parkinson and Jane J Mansbridge, 1–26. Cambridge: Cambridge University Press.

Mill, John Stuart. 1991. "Considerations on Representative Government." In *On Liberty and Other Essays*. Oxford: Oxford University Press.

Misak, Cheryl. 2008. "A Culture of Justification: The Pragmatist's Epistemic Argument for Democracy." *Episteme* 5 (1): 94–105. doi:10.3366/E1742360008000257.

Mochales, Raquel, and Marie-Francine Moens. 2011. "Argumentation Mining." *Artificial Intelligence and Law* 19 (1): 1–22. doi:10.1007/s10506-010-9104-x.

Monnoyer–Smith, Laurence, and Stéphanie Wojcik. 2012. "Technology and the Quality of Public Deliberation: A Comparison between on and Offline Participation." *International Journal of Electronic Governance* 5 (1): 24–49. doi:10.1504/IJEG.2012.047443.

Nanz, Patrizia, and Jens Steffek. 2005. "Assessing the Democratic Quality of Deliberation in International Governance: Criteria and Research Strategies." *Acta Politica* 40 (3): 368–83. doi:10.1057/palgrave.ap.5500118.

Neblo, Michael A. 2007. "Family Disputes: Diversity in Defining and Measuring Deliberation." *Swiss Political Science Review* 13 (4): 527–557. doi:10.1002/j.1662-6370.2007.tb00088.x.

Niemeyer, Simon. 2011. "The Emancipatory Effect of Deliberation: Empirical Lessons from Mini-Publics." *Politics & Society* 39 (1): 103–40. doi:10.1177/0032329210395000.

Ottonelli, Valeria. 2010. "What Does the Discursive Paradox Really Mean for Democracy?" *Political Studies* 58 (4): 666–87. doi:10.1111/j.1467-9248.2010.00821.x.

Pinto, Robert. 2001. *Argument, Inference, and Dialectic : Collected Papers on Informal Logic*. Dordrecht: Kluwer Academic.

Roger, Léa, and Gary S. Schaal. 2013. "The Quality of Deliberation in Two Committees of the European Parliament: The Neglected Influence of the Situational Context and the Policymaking Stage." *Politics and Governance* 1 (2): 151–69.

Rousseau, Jean-Jacques. 1762. "The Social Contract or Principles of Political Right." http://www.constitution.org/jjr/socon.htm.

Saffon, Maria Paula, and Nadia Urbinati. 2013. "Procedural Democracy, the Bulwark of Equal Liberty." *Political Theory* 41 (3): 441–81. doi:10.1177/0090591713476872.

Schkade, David, Cass R. Sunstein, and Reid Hastie. 2010. "When Deliberation Produces Extremism." *Critical Review* 22 (2–3): 227–52. doi:10.1080/08913811.2010.508634.

Steenbergen, Marco R., André Bächtiger, Markus Spörndli, and Jurg Steiner. 2003. "Measuring Political Deliberation: A Discourse Quality Index." *Comparative European Politics* 1 (1): 21–48. doi:10.1057/palgrave.cep.6110002.

Steiner, Jürg, André Bächtiger, Markus Spörndli, and Marco R. Steenbergen. 2004. *Deliberative Politics in Action: Analyzing Parliamentary Discourse*. Cambridge: Cambridge University Press.

Stromer-Galley, Jennifer. 2007. "Measuring Deliberation's Content: A Coding Scheme." *Journal of Public Deliberation* 3 (1). http://www.publicdeliberation.net/jpd/vol3/iss1/art12.

Thompson, Dennis F. 2008. "Deliberative Democratic Theory and Empirical Political Science." *Annual Review of Political Science* 11 (1): 497–520. doi:10.1146/annurev.polisci.11.081306.070555.

Verheij, Bart. 2003. "Dialectical Argumentation with Argumentation Schemes: An Approach to Legal Logic." *Artificial Intelligence and Law* 11 (2–3): 167–95. doi:10.1023/B:ARTI.0000046008.49443.36.

Walton, Douglas. 2002. "The Sunk Costs Fallacy or Argument from Waste." *Argumentation* 16 (4): 473–503. doi:10.1023/A:1021108016075.

———. 2011. "Argument Mining by Applying Argumentation Schemes." *Studies in Logic* 4 (1): 38–64.

Walton, Douglas, and Chris Reed. 2005. "Argumentation Schemes and Enthymemes." *Synthese* 145 (3): 339–70. doi:10.1007/s11229-005-6198-x.

Walton, Douglas, Chris Reed, and Fabrizio Macagno. 2008. *Argumentation Schemes*. Cambridge: Cambridge University Press.