Juho Ritola

Deliberative Democracy, the Deliberating Agent, and Critical Thinking: An ideal picture and some empirical challenges[1]

Abstract

According to some prominent theorists, the conditions of deliberative democracy call for reasoned decisions from mutually justifiable premises. The deliberative ideal places demands on the epistemic quality of the deliberating process and on the epistemic habits and beliefs of the relevant agents. In this essay, I discuss this ideal in light of empirical literature. I examine some empirical literature on human reasoning that paints a bleak picture of human rationality: we fall victim to heuristics and biases, persevere in our beliefs in the face of contrary evidence, and justify our current moral judgments by post hoc-reasoning. In addition, the deliberating groups have specific problems. The groups may, for example, amplify errors or fall victim to information cascades. I argue that, given that we are interested in securing that deliberative process is epistemically valuable, the literature gives further support to the idea that education must foster not only skills but also dispositions for critical thinking. I conclude with a brief defense of epistemic internalism against the argumentation by M. Solomon.[2]

Keywords

Deliberative democracy, critical thinking, reasoning

1. Introduction

One important trend in current political philosophy is the theory of Deliberative Democracy. The guiding thought of this line of thinking is that legitimate lawmaking issues from the public deliberation of citizens.[3] An important division in this literature is the way the value of deliberation is conceived. In his influential article, Jon Elster notes that "[a]ccording to the theorists of participatory democracy from John Stuart Mill to Carole Pateman, the goal of politics is the transformation and education of the participants."[4] Elster argues that

1 This paper was written while working in a project titled "The Sociality of Knowledge", project number 1251076, funded by the Academy of Finland.

2 Miriam Solomon, "Groupthink versus The Wisdom of Crowds: The Social Epistemology of Deliberation and Dissent," *The Southern Journal of Philosophy* 44 (2006), doi: 10.1111/j.2041-6962.2006.tb00028.x.

3 James Bohman and William Rehg, introduction to *Deliberative Democracy: Essays on Reason and Politics*, ed. James Bohman and William Rehg, (Cambridge, MA: MIT Press, 1999), ix.

4 Jon Elster, "The Market and the Forum," in *Deliberative Democracy*, 3.

Juho Ritola, University of Helsinki, Finland
E-mail: juho.ritola@utu.fi

Studier i Pædagogisk Filosofi | www.ojs.statsbiblioteket.dk/index.php/spf | ISSN nr. 22449140

Årgang 4 | Nr. 1 | 2015 | side 29-54

this view is incoherent. While the transformation and education of the participants are good things, ultimately they ought to be viewed as by-products of the democratic process. We should only choose methods of political decision-making that are good for that particular purpose.

It indeed seems counterintuitive to accept that any imaginable deliberation that had no epistemically valuable properties should *fully* legitimate a decision. Suppose citizens convene to deliberate which of the existing computer programs on the market the community should choose for use. The participants discuss and agree that they should pick program A. The only premise for A is that it is the most famous of the programs. If, however, the participants do not know whether program A is within their means, serves the needs of the community, runs on their computers, or even what it is famous for, they have made a bad decision. Examples of bad but democratic decisions are not uncommon.

Unsurprisingly, a common demand in deliberative theorizing is that the process should be reason-based.[5] For example, Robert A. Dahl notes that "…citizens must have adequate and equal opportunities […] for expressing reasons for endorsing one outcome rather than another."[6, 7] Joshua Cohen holds that a decision is legitimate if and only if it could be the object of free and reasoned agreement among individuals."[8] Amy Gutmann and Dennis Thompson state that "[d]eliberative democracy asks citizens and officials to justify public policy by giving reasons that can be accepted by those who are bound by it."[9] David Estlund avers that "[the legitimacy of a decision] derives, partly, from the epistemic value, even though it is imperfect, of the procedure that produced it."[10]

These conditions of legitimate deliberation emphasize justification: the decision must be based on good reasons, and good reasoning must lead from these premises to the conclusion. Otherwise, arguably, the decision could not be the object of free and reasoned agreement. From an epistemic point of view, this is to be applauded. It also seems reasonable to specify that there are at least two factors that are required for reasonable deliberation from an epistemic point of view. First, the procedure itself should be rational. For example, it should give each participant sufficient time to justify their view.[11] If the citizen is

5 It is not the case, however, that the aforementioned deliberation was not reason-based. It was based on a bad reason. But I will ignore this in what follows. I assume that we are interested in bringing about decisions based on good reasons.

6 Robert A. Dahl, "Procedural Democracy," in *Contemporary Political Philosophy: An Anthology*, ed. Robert E. Goodin and Philip Pettit (MA: Blackwell Publishing, 2nd edition, 2006), 109-110.

7 Dahl's discussion is directed at choosing a method of political decision-making: only a method that fulfills this requirement (and some others) should be chosen. Dahl also notes that the words he uses are ambiguous, but it seems that his discussion clearly has epistemic overtones.

8 Joshua Cohen, "Deliberation and Democratic Legitimacy," in *Deliberative Democracy*, 73.

9 Amy Gutmann and Dennis Thompson, *Democracy and Disagreement* (MA: The Belknap Press of Harvard University, 1996), 52.

10 David Estlund, "Beyond Fairness and Deliberation: The Epistemic Dimension of Democratic Authority," in *Deliberative Democracy*, 174.

11 Dahl further sets up the criterion of enlightened understanding: "In order to express his or preferences accurately, each citizen ought to have adequate and equal opportunities for discovering and validating, in the time permitted

not allowed to justify the crucial premises and the cogency of his or her reasoning against relevant challenges, other parties to the deliberation cannot make an informed judgment about that position and, consequently, about the case at hand. Second, the citizen that participates in a democratic deliberation should behave in an epistemically respectable way. For example, the discussants should only make claims that they can justify, and only propound reasoning they think is good. The reasonability of the decision is dependent on these epistemic conditions, but the conditions do not guarantee good results: we are fallible. There also seems to be little hope of designing a procedure of deliberation that could not be short-circuited by resolute individuals, who give no regard to evidential considerations or other reasonable demands of proper dialogue. Knowingly giving misleading information or persuasive but fallacious reasoning can harm the reasonability of the process or, in the very least, give the process an unacceptably arbitrary character.[12]

The ideal of deliberative theorizing thus calls for a certain kind of process and a certain kind of citizen. The citizen naturally should have certain inferential and dialectical skills that allow him or her to approximate truth and moral value in deliberation, or reach justified beliefs. However, there is a large body of empirical literature that casts serious doubt on our ability as cognitive agents to live up to these requirements.[13] This literature seems to show that poor quality judgment does not result only from a lack of knowledge about the subject matter, lack of proper standards of reasoning, motivational issues, or from succumbing to various fallacies[14] identified in the literature. The typical individual is further argued to fall victim to poor quality heuristics and allow various biases influence judgments,[15] persevere in our beliefs in the face of contradicting evidence,[16] and use moral reasoning to justify existing moral judgments post-hoc[17]. And as if these problems were not enough, the deliberating bodies seem vulnerable to problems of their own: the groups may amplify

by the need for a decision, what his or her preferences are on the matter to decided." ("Procedural Democracy," 111).

12 For example, a case where two mistakes by the participants cancel each other out by chance seems arbitrary, even if the participants came to a reasonable conclusion.

13 The classic source for this literature is Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., *Judgment Under Uncertainty: Heuristics and Biases*, (Cambridge: Cambridge University Press, 1982). It contains the seminal article by Tversky and Kahneman from 1974, bearing the same name as the article collection. The seminal article concentrates on probability reasoning, but the flourishing research program soon identified various other problems of similar magnitude in other reasoning tasks. See below for further references to the ongoing discussion.

14 By this I refer to something like the "traditional gang of eighteen," an expression coined by John Woods (see John Woods, "Who Cares about Fallacies?," in *Argumentation Illuminated*, ed. Frans van Eemeren et. al. (Amsterdam: SicSat, 1988)) to refer to the fallacies that the fallacy literature typically talks about. These include fallacies like ad hominen and other 'ad' -fallacies, begging the question, and the like.

15 Kahneman, Slovic, and Tversky, *Judgment Under Uncertainty*, 1982.

16 C. Lord, Lee Ross, and M. R. Lepper "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence," *Journal of Personality and Social Psychology* 37 (1979).

17 Jonathan Haidt, "Emotional Dog and Rational Tail: A social intuitionist approach to moral judgment," in *Reasoning: Studies of Human Inference and Its Foundations*, ed. Jonathan E. Adler and Lance J. Rips (Cambridge: Cambridge University Press, 2008).

error rather than correct it, fall victim to informational cascades, and give disproportionate evidential value to agreement.[18]

In this essay, I will first discuss theories of deliberative democracy, noting some features that have been deemed essential for proper deliberation. I will then turn to empirical literature and examine some of the challenges it poses to the ideal picture of a rational deliberator. Though part of the literature paints a gloomy picture of human rationality, and hence raises doubts on the prospects of epistemically commendable citizen deliberation, we need not draw the conclusion that we are determined to produce epistemically suboptimal results. The literature does seem to indicate, however, that education needs to pay special attention to fostering both the appropriate skills of and dispositions for critical thinking, a position championed by Harvey Siegel.[19]

At the end of the essay, I will discuss the nature of justification in the deliberative context. From an epistemological perspective, the ideal picture of a citizen and the conditions of a proper deliberation seem to call for two further theses that are disputed by many. The first of these is epistemological internalism.[20] I think the empirical literature in fact gives some support to an internalistic picture of justification. This is diametrically opposed to Solomon, who argues that empirical evidence about group reasoning in fact supports externalism.[21] I support my case with the fact that a crucial factor in attaining objectively justified beliefs is the ability to monitor one's belief-formation and apply critical methods both to its results and to the specific processes that formed them. This applies both to beliefs formed in solo reasoning and group reasoning.

The second thesis is dialectical justification. It states that a necessary condition of being justified in one's belief is that one is able to defend one's belief against (relevant) challenges. This thesis is not popular among epistemologists, but the deliberative context requires this dialectical sense of justification. I think Solomon's criticism is not effective against internalism but is rather directed against this dialectical requirement. However, the dialectical requirement is important for deliberation regardless of the problems of group reasoning.

2. Conditions of deliberative democracy

As was noted, the guiding thought of deliberative democracy is that legitimate lawmaking issues from the public deliberation of citizens. Bohman and Rehg relate that the central

18 Solomon "Groupthink;" Cass Sunstein, "Deliberating Groups versus Prediction Markets (or Hayek's Challenge to Habermas)," *Episteme* 3, (2006).

19 See for example Harvey Siegel, *Educating Reason: Rationality, Critical Thinking, and Education* (New York: Routledge, 1987).

20 While I accept this thesis, most of the discussion in this paper does not hinge on it. If one sees the ideal picture of the deliberating citizen as appealing, empirical evidence about our reasoning performance and its failings is important whether one is an externalist or an internalist. Both can also accept that a critical thinking class should aim to objectively improve the reasoning performance of the student that is or will later be a citizen involved in deliberation.

21 Solomon, "Groupthink."

ideas of deliberative democracy started to take shape in 70's, and the term 'deliberative democracy' was coined only 1980 by Joseph M. Bessette.[22, 23] By now, it has become an important part of the discussion in political philosophy. I will briefly discuss how the condition of basing decisions on reasons is motivated and formulated in deliberative democracy. I do not claim comprehensiveness; the purpose is only to highlight some theorists that emphasize the concept of reason. This will motivate the ideal picture of a citizen as an epistemic agent that we will then contrast with the empirical findings.

The deliberative movement juxtaposes itself with the purely "aggregative" view of democracy based on theories of rational choice, which postulate that individuals have fixed sets of preferences, ordered according to their desirability. The individual is seen rational in the sense that he/she, under the constraints and opportunities of his/her situation, will make choices that best serve those interests.[24] Under the aggregative view, the individual, a political actor, may also be characterized as taking part in a bargaining process: knowing what he/she wants, the individual maneuvers to arrive at an outcome that is as close to his/her preferences as possible.[25] The main form of political participation is voting, and the institutions of government are designed to allow free pursuit of interests.

Bohman and Rehg see as a central tenet of deliberative democracy that political processes cannot be reduced to individual choices in the "market," but instead, the citizens should assume a civic standpoint and orient themselves to the *common good*[26] and discuss issues in the "forum."[27] The individual actors, in this view, do not come to the decision situation with fixed preferences: they are willing and able to change their minds, based on the information and argumentation encountered in the process deliberation. Shawn Rosenberg notes that even though deliberative theorists explicitly reject Rawls'[28] concept of the veil of ignorance, they do draw on his conception of the individual citizen as having much greater cognitive capacities and moral potential than in rational choice theories.[29]

We should now look more specifically at some of the conditions that have been set on a good deliberative process. Joshua Cohen argues that "outcomes can only be legitimate if

22 Joseph M. Bessette, "Deliberative Democracy: The Majority Principle in Republican Government," in *How Democratic is the Constitution?*, ed. Robert Goldwin and William Schambra (Washington, D.C.: American Enterprise Institute, 1980).

23 Bohman and Rehg, introduction, xii.

24 Shawn W. Rosenberg, introduction to *Deliberation, Participation, and Democracy*, ed. Shawn W. Rosenberg (Hampshire: Palgrave Macmillan, 2007), 5.

25 Jürg Steiner, The *Foundations of Deliberative Democracy*: Empirical Research and Normative Implications (Cambridge, UK: Cambridge University Press, 2012), 4,doi: 10.1017/CBO9781139057486.

26 Unfortunately, I have to leave this crucial term undefined. I simply assume it is something that can legitimately be con sidered as an answer to the question "what would be good for all of us?" which is different from, but not necessarily opposed to, "what would be good for me?"

27 Bohman and Rehg, introduction, xiii; the use of the terms "the market" and the "forum" derive from Elster "The Market and the Forum," which was originally published in *The Foundations of Social Choice Theory*, ed. J. Elster and A. Hylland (Cambridge: Cambridge University Press, 1986).

28 In, for example, John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993).

29 Rosenberg, introduction, 6.

and only if they could be the object of free and reasoned agreement among equals".[30] This is spelled out by further conditions, of which we can concentrate on the idea that it must be reasoned. The reasoned condition means:

> that the parties to it are required to state their reasons for advancing proposals, supporting them, or criticizing them. They give reasons with the expectation that those reasons [...] will settle the fate of their proposal. [...] Proposals may be rejected because they are not so defended with acceptable reasons, even if they could be so defended.[31]

So, even though the process is premised on equal participation, the process itself is governed by reason-giving. We should note that the condition of justification is distinctly dialectical: one must be able to defend one's claim publicly. Private musings, no matter how justified, are not sufficient, and the place for giving the reasons is during the process, not afterwards, or in other forums.

Amy Guttmann and Dennis Thompson set three essential conditions to the reasonability of a deliberation. First, the deliberating process must fulfill the principle of reciprocity: you must seek reasons that you and your opponent can accept. Even if they are not to be found, you must keep on searching. This reciprocity is given the role of regulator of reason.[32] The second important principle is publicity, which means that the reasons the officials and the citizens give to justify political actions, and the information to assess those reasons, should be public.[33] The third important principle is that justificational demands apply between everyone: in a deliberative forum, each is accountable to all.[34]

What epistemic characteristics might we require from a decision so that it could be the object of free and reasoned agreement? For the purposes of this discussion, we can draw a rough, intuitive, picture. We want a decision that is premised on justified, preferably true premises, and on cogent reasoning from those premises. The decision should also be based on appropriate reflection of moral and practical concerns. In all its complexity, this would entail, among other things, that the relevant evidence was gathered well, conclusions from the evidence were drawn with reliable methods, all relevant evidence was considered, etc., but we can do with this sketch of ideal rational justification here.

An important issue here is that of finding common premises.[35] This problem is premised on the fact that the political life, and democratic deliberation, must start from reasonable

30 Cohen, "Deliberative Democracy," 73.
31 Cohen, "Deliberative Democracy," 74.
32 Gutmann and Thompson, *Democracy and Disagreement*, ch. 2.
33 Gutmann and Thompson, *Democracy and Disagreement*, ch. 3.
34 Gutmann and Thompson, *Democracy and Disagreement*, ch. 4.
35 I address some of the related issues in my "Reasonable pluralism and the dialectical conditions of knowledge," forthcoming.

pluralism[36]: citizens possess different, yet reasonable, comprehensive doctrines (religious, philosophical, and moral) on which they cannot hope to agree fully. Because of this, there is no guarantee that the deliberators share enough common premises to reach a decision. But in order to argue together, the citizens must reflect on what premises they can use and find common premises that could be accepted by all parties.

The nature of appropriate premises is a question that has occupied theorists of political philosophy. The intuitive starting point is the shared premises requirement.[37] Rawls develops the idea of an overlapping consensus.[38] Another suggestion is that the decision could be based on an incompletely theorized agreement.[39] This debate reflects the fact that the rules of deliberation are not solely epistemic but also political and moral. It is true that the requirement of common premises (or the obligation to keep on searching for common premises) is justifiable from the viewpoint of argument as *rational persuasion*: to bring about rational persuasion, it is necessary to start from premises that the opponent can rationally accept. But we can easily imagine situations, where the speaker has no way of rationally persuading the hearer, even though the claim in question is epistemically justified to the arguer herself. Yet the obligation to keep on searching for common premises to justify the decision obtains, because the context is that of a *political* decision. This is the import of the central idea of modern political philosophy that political order, realized through political decisions, can only be justified if it is justified to those who are bound by the order in question.[40] One may have to, for example, keep on searching for new corroborating evidence that can be used in public deliberation, though one knows full well that the standards of reasonable evidence gathering have been met.

The picture of ideal deliberation places considerable requirements on the citizen: making epistemically and morally sound decisions under reasonable pluralism is no easy task. If one holds that the deliberative decisions must be epistemically and morally commendable, one should ask how we as citizens fare in this task. We have belief-forming habits that make many of the beliefs we hold epistemically suspect. It is now time to turn to some empirical evidence about human reasoning performance.

3. Challenges to the deliberative ideal

In this section, I will discuss some of the epistemically problematic habits, methods, and procedures people resort to in forming beliefs. I will first discuss issues that pertain mainly to empirical (contingent) beliefs and assess their meaning based on the literature. Second,

36 Term coined by Rawls; see for example *Political Liberalism*, 100.

37 Discussed by, for example, Bruce Ackerman, "Why Dialogue?," *Journal of Philosophy* 86 (1989), doi: 10.2307/2027173.

38 Rawls, *Political Liberalism*, ch. 2.3.

39 Cass R. Sunstein, "Incompletely Theorized Agreements," *Harvard Law Review* 108 (1995).

40 A good introduction to this idea is Kevin Vallier and Fred D'Agostino, "Public Justification," *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), ed. Edward N. Zalta. Accessed May 30, 2015. URL = <http://plato.stanford.edu/archives/spr2014/entries/justification-public/>.

I will discuss issues in the justification of moral beliefs. Third, I will discuss problems that relate to beliefs formed in a group.

3.1. The reasoning agent

A group of citizens deliberating on what to do is typically faced with complex problems that require rational intake, analysis, and evaluation of information. The decisions, it has been argued, ought to be reasonable in light of the available information. But the evidence on human performance in reasoning tasks is not reassuring. One research tradition that has significantly influenced our picture of human reasoning is *Heuristics and Biases* (HB).[41] It has produced results that cast a serious shadow on the reasonability of the individual citizen's performance on a variety of reasoning tasks. Although the reasonability of the group decision is not a direct function of the reasonability of the individuals in the group, the reasonability of the individuals is hardly irrelevant either. If all the individuals are very unreliable in processing information, it seems unlikely that the group as a whole is very reliable.

The crucial finding of HB, documented in a wide array of studies, is that the human agent seriously underuses the normatively appropriate reasoning strategies and overuses more primitive intuitive strategies. Arguably, the poor performance is based on the same implements that are the basis of successful reasoning. Our reasoning performance is based on two interrelated systems that are deeply embedded in our cognitive system. First, there are *knowledge structures* that allow the individual to define and interpret the data of physical and social life. These structures are something we need to deal with the abundance of information with which we are faced. Without them, life would be a "buzzing confusion." Second, there are *judgmental heuristics* that reduce complex inferential tasks to simple judgmental operations. Though these strategies work well for the individual in many instances, there is a price to be paid for the mental economy.[42]

The main judgmental heuristics identified by HB are *availability* heuristic and *representativeness* heuristic. Let us start with the availability. When we judge the relative frequency of particular objects or the likelihood of particular events, we are often influenced by their relevant availability, that is, their accessibility in the processes of perception, memory, or construction from imagination. When this availability is paired with objective frequency, it can be a useful heuristic. However, there are many factors uncorrelated with frequency that can influence an event's perceptual salience, vividness or completeness with which it is recalled, or the ease with which it is imagined. These factors can make the availability misleading. For example, people consistently err in judging the relative frequency of two kinds of English words. When asked to estimate the number of words beginning with a specific letter (for example, 'r' or 'k') in relation to words where the same letter appears third,

41 Cf. note 13. My description of this research program here is mainly based on another important book of this tradition by Richard E. Nisbett and Lee Ross, *Human Inference: Strategies and Shortcomings* (Englewood Cliffs, New Jersey: Prentice Hall, 1980).

42 Nisbett and Ross, *Human Inference*, 3-7.

they falsely think there are more words beginning with those letters. Apparently, the ease at which one can generate words beginning with a certain letter in relation to generating words where the same letter is third, leads one to think that the first are more numerous. However, the ease of generation does not correlate with the objective frequency. People commit similar mistakes in causal reasoning. For example, having had to explain a certain event (which makes a specific causal scenario very available to the person) increases the subject's belief in the likelihood of that event.[43]

The representativeness heuristic involves the application of relatively simple resemblance criteria to categorization. When we make a categorization judgment, we assess the degree to which the salient features of the object are representative of the features presumed to be characteristic of that category. For example, subjects are asked to evaluate the relative likelihood of three sequences of births of girls (G) and boys (B): 1) BBBBBB, ii) GGGBBB, iii) GBBGGB. If they go by representativeness heuristic, they tend to answer iii), based on what they know about the population of babies and about the randomness of the process, meaning that each birth is a "random" event in which the probability of "boy" and "girl" are nearly equal. However, the likelihood of each of these series is nearly equal. The explanation is that under representativeness heuristics, ii) does not seem representative of the randomness, and i) seems representative neither of randomness nor the population from which the sample was taken. The representativeness heuristics is a very strong heuristic and many studies[44] have documented its force in various settings.

It bears repeating to note that these heuristics are vital to many inferential tasks; induction and generalization are not possible without the subjects being able to decide what class or category one is observing, and such judgments essentially hinge upon judgments of representativeness. So "…the problem is clearly one of overapplication […] the generating process is generally valid. It leads people to recognize that an all-male or an all-white jury is more likely to reflect a biased selection procedure than will a jury with a more proportionate representation of the population."[45]

Another important part of our making sense of the world is the use of various knowledge structures or *schemas*. These structures pertain to for example events (termed *scripts*) or persons (termed *personae*). They are fairly loose structures, but once an object fulfills enough conditions to be placed under a schema (say, 'a dog'), the user of that concept readily assigns various other characteristics to that object (for example 'capable of loyalty,' 'likely to chase cats' etc.). For the management of mental life, schemas are even more impor-

43 Nisbett and Ross, *Human Inference*, 18-22.

44 The most famous of them probably being Amos Tversky and Daniel Kahneman, "Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment," *Psychological Review* 90 (1983), doi: 10.1037/0033-295x.90.4.293. In the study, subjects overwhelmingly violated the conjunction rule of probability, which states that any conjunction of two properties is less probable than the conjuncts separately. When subjects were given a description they thought representative of a certain type of person, they selected a conjunction containing that property rather than singular options not containing the conjunct deemed representative.

45 Nisbett and Ross, *Human Inference*, 29.

tant than heuristics, but they are also often overused and are liable to mislead, especially when combined with the availability and representativeness heuristics. I will not delve into the literature on this phenomenon but note only one representative case mentioned by Nisbett and Ross[46], from E.R. May's book *"Lessons" of the Past*[47]. The example highlights the kind of problems to which the combination of poor knowledge structure and availability can lead. It has been noted that President Harry Truman developed a strong trust to his wartime ally Joseph Stalin. According to May, Truman's correspondence reveals that he developed this trust, because Stalin evoked the persona of Tom Pendergast, Truman's long time benefactor. Though Pendergast was known to be a ruthless and corrupt kingmaker, he was always completely trustworthy in his relations to Truman. The thought runs that as Stalin evoked certain characteristics of Pendergast, Truman assumed that other characteristics could also be assumed. Nisbett and Ross note that typically the case is that a given schema is overused, but sometimes the conceptual category is so lacking in foundation and predictive value that it almost invariably serves its user badly. Many racial and ethnic stereotypes fit this description.[48]

Another phenomenon that seems problematic for rational deliberation is *belief perseverance*. Originating in the work of Lord, Ross, and Lepper[49] this term refers to the problems people exhibit in dealing with new knowledge. This shows mainly three ways. First, when a subject has a theory about a given topic, exposure to new evidence (whether it supports, opposes, or is mixed), will tend to result in more belief in the correctness of the original theory than the normative dictates allow. Second, when a belief is formed based on the evidence, subsequent evidence will tend to be disregarded, counter to obvious dictates about the totality of evidence. Third, once a theory is formed, it survives the total discrediting of the original evidence.[50] To caricaturize, whatever gets in first, regardless of how it got there, will stay there, and becomes more entrenched, come what may.

To complete this quick overview of HB, I will mention one final case, the significance of which has been debated in the literature intensely. In the famous *selection task*, designed by P.C. Wason (1960),[51] people were shown four cards that either had a number or a letter on it. One card had a vowel pictured on it, second card a consonant, third an even number, and the fourth an odd number. Subjects were then asked to test the following rule: "If a card has vowel on one side, then it has an even number on the other side." The normatively correct way to test the rule is to turn two cards: i) the card featuring a vowel,

46 Nisbett and Ross, *Human Inference*, 39.

47 E.R. May, *"Lessons" of the Past* (New York: Oxford University Press, 1973).

48 Nisbett and Ross, *Human Inference*, 40.

49 C. Lord, Lee Ross, and M.R. Lepper, "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence," *Journal of Personality and Social Psychology* 37 (1979), doi: 10.1037/0022-3514.37.11.2098.

50 Nisbett and Ross, *Human Inference*, 169.

51 P.C. Wason "On the failure to Eliminate Hypotheses in a Conceptual Task." Quarterly Journal of Experimental Pscychology 12 (1960), doi: 10.1080/17470216008416717. Reprinted in *New Horizons in Psychology, ed.* B.M. Foss (Harmondsworth: Penguin, 1966).

and ii) the card featuring an odd number. The explanation runs through the semantics of the conditional in the rule. The rule is an implication by its logical form, and implication is false only if the antecedent (the sentence followed by 'if') is true and the consequent (the sentence followed by 'then') false. In testing a rule, you try to find counterexamples to it. By the semantics of the conditional, this could only be the case where either, the antecedent is true, so by turning the card that has vowel on it and finding the consequent false (an odd number), you would have refuted the rule (i.e. case i)) or, the consequent is false with respect to the rule (an odd number), turning the card and finding a true antecedent, you would have refuted the rule (i.e. case ii)).

The surprising result of the test was that in its basic form, only 4 percent of the respondents provided the normatively correct response (Wason and Shapiro 1971). Though the result has been shown to differ somewhat depending on various factors,[52] for example the semantic content of the rule (Johnson-Laird, Legrenzi, and Sonino Legrenzi 1972), this startling result paved the way for a host of new research programs trying to explain why humans, clearly capable of great intellectual achievements, can also fail so miserably.

The evidence on poor human performance is plentiful, and it is not limited to the kind of examples discussed here. Even the basic forms of deductive reasoning do not seem to be clear to people. Jonathan Evans, Stephen Newstead, and Ruth Byrne collected evidence from different studies on human performance in simple reasoning tasks involving modus ponens (If p, then q. p. Therefore, q.), modus tollens (If p, then q. Not q. Therefore, not p), and the corresponding fallacies of affirming the consequent (If p, then q. q. Therefore, p.) and denying the antecedent (If p then q. Not p. Therefore, not q).[53] The results are surprising. While modus ponens was widely recognized as valid (from 91% to 100%), modus tollens was not. It was accepted as a reasonable inference from 41% to 81%.[54] The two fallacies, affirming the consequent and denying the antecedent, received acceptance rates from 23% to 75% and 17% to 73%, respectively![55]

3.2 Assessing the evidence

The HB-tradition thus paints a bleak picture of the average human reasoning capabilities. Given that one accepts that a deliberative decision ought to be epistemically valuable, one might take this literature to show that deliberative democracy is not a good idea. However, this conclusion can be resisted with good reason. I will now highlight some important

52 But see below.

53 Jonathan St. B. T. Evans and Stephen E. Newstead, *Human Reasoning: The Psychology of Deduction* (Hove, UK: Lawrence Erlbaum Associates 1972): 26-36, doi: 10.1111/j.2044-8295.1972.tb01287.x.

54 If negation was used to make a valid inference more complex (e.g. If not p, then q. Not q. Therefore, p), the acceptance rates dropped radically, ranging from 12% to 34%.

55 Two quite good popular books discussing these issues are Thomas Gilovich, *How We Know What Isn't So* (New York: Free Press, 1991) and Stuart Sutherland, *Irrationality: The Enemy within* (London: Constable, 1992). For an in-depth discussion of human reasoning, I recommend *Reasoning*, ed. Adler and Rips.

arguments and results from the literature that give us a better understanding of the HB-literature. These insights also bear on how critical thinking should be taught.

First, L. Jonathan Cohen has argued that there is no straightforward inference from these results to the overall individual irrationality.[56] Though Cohen did not aim to deny the significance of these results, he argued that those who tend to overestimate human reasoning powers tend to concentrate on the human competence, while the underestimation usually derives from exaggerated concentration on performance. We are clearly capable of normatively good reasoning. If there is evidence of systematic failure of our reasoning performance, we need to explain how that comes about and find ways to combat it. Furthermore, it is inherently problematic for a human being to say that there are normatively appropriate reasoning systems, but human beings are not able to use them. Our intuition serves as a crucial tool in identifying the normatively correct inference systems. But if intuitions serve at the ultimate basis of any normative theory of reasoning, one must assume that humans, though fallible, are basically rational. If not, we have no promise of us being able formulate rational theories about inference.[57]

At this stage, a different type of objection might come to mind. If humans indeed are basically rational but fail miserably in tests of performance, maybe we have applied the wrong standards. Maybe it is the standards of deductive, inductive, and abductive reasoning that ought to yield. According to some researchers, there is indeed a grain of truth in this: we do not always and automatically apply these standards of reasoning, yet can manage our life quite well. But this is not a reason for abandoning the standards, although it is probably part of the explanation why the performance in these tests is so poor. Keith Stenning and Michael Lambalgen[58] argue that the study of reasoning in laboratory settings that produced the gloomy results assumed that all the subjects were trying to do same thing.[59] But if we take seriously the role of *interpretation*, we will come to see that the subjects were trying to reason to an interpretation of what they were asked to do. Since the laboratory experiments happen in a "vacuum," it is natural to expect that the answers of subjects varied greatly. Without having the normal background for reasoning, they reason about the parameters of the situation and come to different solutions. Basically, subjects apply their knowledge of everyday language, which does not function according to canons

56 L. Jonathan Cohen, "Can Human Irrationality Be Experimentally Demonstrated?", *Behavioral and Brain Studies* 4 (1981), doi: 10.1017/S0140525X00009092.

57 Cohen, "Human Irrationality," 318-323. Cohen's argumentation has been challenged. See for example David Shier, "Can Human Rationality Be Defended A Priori?," *Behavior and Philosophy* 28 (2000).

58 Keith Stenning and Michiel van Lambalgen, "Interpretation, Representation, and Deductive Reasoning," in *Reasoning*, ed. Adler and Rips.

59 The discussion by Stenning and Lambalgen is situated in the debate that assesses rule-based theories of reasoning (see for example Lance J. Rips, "Logical Approaches to Human Deductive Reasoning," in *Reasoning*, ed. Adler and Rips, 187-205) and model-based theories of reasoning (see for example Philip N. Johnson-Laird, "Mental Models and Deductive Reasoning," in *Reasoning*, ed. Adler and Rips). These theories tried to unearth the real reasoning processes that individuals use. This would naturally make us better positioned to say exactly why the subjects do so miserably in these tests.

of classical logic. Since the subjects are trying to do different things in the tests, it does not make sense to take their responses to indicate a position on some other dimension, such as intelligence.[60]

Stenning and Lambalgen draw on dual processing theory by Jonathan Evans[61] (2003) that separates between System 1, an automatic nonconscious reasoning processes and System 2, explicit controlled reasoning. The System 1 mechanism, shared between humans and animals, is logical but non-monotonic. The System 2 is conceived as a repair mechanism when problems are found in level 1, and requires a change of machinery to a monotonic logic. This machinery is some version of classical logic, (but only when reasoning *from* an interpretation, not *to* an interpretation.) The System 2 processing, (so-called skeptical stance) is essential for several contexts, or better, basically for any context where disagreement is widespread, such as political deliberation. But both systems are important in view of deliberation. In order to reach enough common ground for arguing, that is, to understand what other deliberators are trying to say, the subjects need to employ some machinery of the defeasible logics, i.e. System 1 processing. However, in order to do the kind of examination of claims and arguments that is essential to reasonable deliberation, the subjects must be taught the methods of classical logic, which also form the basis of conceptual learning. Stenning and Lambalgen note that:

> […] classical reasoning is important not because an implementation of it is the "universal deductive reasoning mechanism," but, rather, because classical reasoning is important for aligning and repairing mutual interpretation across some gulf of understanding or agreement, and that learning more explicit control over System 2 process, and their relation to System 1 processes, can have a large impact on many student's interpretation, learning and reasoning process. The skills of skeptical reasoning are then one extremely important set of concepts for learning to learn.[62]

The startling results of HB do not show, then, that people are, in general, logically inept, or that the principles of deductive, inductive, and abductive reasoning are not important to reasoning in the wild, that is, outside the classroom. We need both systems. Yet, it is neither the case that the gloomy results are not important nor that they can be fully explained away. Many aspects of relevant norms, the set-up of the studies, interpretation, and pragmatics complicate the picture, but it remains the case that we are prone to commit many kinds of inferential errors, and without taking a pessimistic or a condescending[63] tone, we can probably agree that this shows in many ways in all walks of life. The teaching of good

60 Stenning and Lambalgen, "Interpretation," 223-227.

61 Jonathan St. B. T. Evans, "In Two Minds: Dual-Process Accounts of Reasoning," *Trends in Cognitive Science* 7 (2003), doi: 10.1016/j.tics.2003.08.012.

62 Stenning and Lambalgen, "Interpretation," 245.

63 Nisbett and Ross (*Human Inference*, 14) write that "[…]this book was written in a spirit of genuine humility. We have found that our richest source of data demonstrating human inferential failings comes not from the

reasoning is therefore crucial: we need to be able to elicit System 2 processing, which also leads to normatively better responses. Let us elaborate on this and the role intelligence by looking into another empirical study.

Keith Stanovich notes that the ability to provide normatively correct responses in reasoning tasks has been shown to correlate with high scores in IQ-tests.[64] But Stanovich argues that important differences pertaining to rational thought are ignored if we only concentrate on the intelligence-related variance. Such a focus could be particularly harmful in the context of assessing the reasonability of democratic deliberation. To understand this, we need to introduce some terminology.

Cognitive theorizing recognizes different levels in the object to be studied. At the base is the biological level that physically realizes the operations. It is inaccessible to cognitive theorizing. The second level is the algorithmic level that is concerned with the computational processes. The cognitive scientist typically operates on this level, positing certain information-processing mechanisms that would explain how a given task gets executed. Third level is the intentional level, which involves the system's goals, beliefs, and choices of action. The constructs of this level consist of control states that regulate behavior, epistemic dispositions that can indirectly alter information pickup tendencies, and regulatory systems that may bring about for example consistency checks. Whereas cognitive psychologists have largely focused on the algorithmic level, personality psychologists have focused on thinking dispositions. These dispositions are intentional-level constructs that pertain to, for example, attitudes toward forming and changing beliefs.[65]

The execution of rationality encompasses two things: the thinking dispositions and the algorithmic level. Intelligence tests concentrate on differences at the algorithmic level and measure performance under optimal situation, which are construed to be situations where the task interpretation is determined externally, the participant is told to maximize performance, and is told how to do so. These situations are called constrained: an attempt is made to specify the task demand so explicitly that variation in intentional level thinking dispositions are minimally influential. In contrast, critical thinking tests are not constrained at the intentional level, and they allow high-level personal goals and their regulation to be implicated in performance. The tendency to change beliefs in the face of contrary evidence or the tendency to not think through different possibilities a problem might involve are examples of behavior that can affected by these goals. A typical test of intelligence strips away features that might bring about belief biases (by for example, using letters in place of sentences or using unfamiliar content). This allows the test to concentrate on the algorith-

undergraduates in our experiments or classrooms but from ourselves and our friends, most of whom are trained social scientists.

64 Keith E. Stanovich, "Individual Differences in Reasoning and the Algorithmic/Intentional Level in Cognitive Science," in *Reasoning* ed. Adler and Rips, 414. See the text for references to these tests. Stanovich's text also contains many references to other texts that use the terms introduced here somewhat differently. These matters are subject to fundamental debates in cognitive science. I will follow Stanovich's classifications and descriptions.

65 Stanovich, "Individual Differences," 414-416.

mic level processing.[66] It is worth pointing out that the reasoning challenges endemic to democratic deliberation are not constrained.

According to Stanovich, the current research on intelligence is starting to converge on the idea that a common feature behind good performance on algorithmic level processing is the concept of decoupling of mental representations. Decoupling supports hypothetical reasoning, which is one of our most important mental functions. Hypothetical reasoning involves representing possible states of the world rather than actual states and is involved in great many reasoning tasks from deductive reasoning to decision making. The ability to decouple makes humans able to distance themselves from the issue, form metarepresentations about oneself and assess learning, and, above all, form and test hypotheses. It appears that a key to intelligence is the ability to maintain decoupling while carrying out mental simulation. Together with a high capacity working memory[67], decoupling is a prime indicator of intelligence as the ability to process information.[68]

Given this picture, it is no surprise that intelligence correlates with the ability to perform well in constrained reasoning tasks. But a more interesting problem, especially from the viewpoint of critical thinking and deliberation, is the fact that intelligence does not always correlate with good performance in unconstrained reasoning tasks. Stanovich's explanation of this "on/off" -nature of correlations runs through dual-processing theory. Typical reasoning tasks in HB-literature pit System 1 processes against System 2 responses. To reach a normatively correct response, the System 2 processing needs to override System 1 response. This override determines whether performance in unconstrained reasoning tasks correlates with intelligence. Stanovich argues that it just such situations that create correlations between intelligence and task performance. He argues that some subjects are more likely to operate entirely on System 1 processing, and these subjects are more likely to be low IQ subjects. Subjects with high IQ are more likely to resolve the conflict in favor of a System 2 response. Yet, successful override of the System 1 happens only through coordinated action of the intentional-level operations recognizing the need to override the System 1 response, and the algorithmic level carrying out the cognitive decoupling operations necessary to cancel the System 1 response. Because intelligence tests are fixed on the individual differences on the algorithmic level, there is room for differences in the intentional-level thinking dispositions to predict difference in unconstrained reasoning tasks.[69]

The empirical evidence seems to corroborate this idea. One essential factor in critical thinking tests is the ability to evaluate the force of arguments independently of prior belief. The evidence has it that a good performance on this score is predicted by thinking dispositions even after general cognitive ability has been partialled out. Other studies[70] have

66 Stanovich, "Individual Differences," 416-418.

67 Working memory is not mainly about memory as such, but about the ability to maintain or suppress (irrelevant or distracting) information.

68 Stanovich, "Individual Differences," 418-420.

69 Stanovich, "Individual Differences," 420-423.

70 See Stanovich, "Individual Differences," 423, for references.

found that the degree to which subjects criticized belief-inconsistent evidence was unrelated to cognitive ability. If the individual is able to, based on intentional-level constructs such as epistemic dispositions, to decouple prior belief from argument evaluation, their performance will be higher on critical thinking scores, regardless of the underlying cognitive capacity. The fact that correlations between high IQ and performance in reasoning tasks in HB-literature are inconsistent seems to be explainable from these intentional-level constructs.[71]

So, although the results that HB has produced are important, there is no need for pessimism. We are prone to err in certain ways, but we are able to learn the proper standards of reasoning. However, in order to perform according to these standards, we must also develop and foster appropriate attitudes for overcoming these errors. The development of such attitudes, and the consequent improvement in reasoning tasks is not tied to sheer cognitive computing ability.

The kinds of skills and dispositions required for epistemically appropriate reasoning bear resemblance to the virtues emphasized in the literature on deliberative democracy. As deliberators we must be able to decouple prior belief from argument evaluation. Though this is not exactly the same as 'assuming a civic standpoint,' proper evaluation does seem to call for the ability to detach ourselves from our own perspective and to evaluate the reasoning of others through proper standards. A decision of a deliberation can be 'the object of free and reasoned agreement among equals,' only if it fulfills the standards or proper reasoning. We can bring about rational persuasion, and be rationally persuaded by others, only if all the parties to the deliberation are able to take in new evidence and evaluate it on its own merits, and further evaluate its implications for our prior beliefs. Rational persuasion is also not possible, if we are not willing to change our mind based on the argumentation presented in the deliberation.

This section has so far concentrated on theoretical reasoning on the individual level: on judgments and reasoning about what is the case. But public deliberation, though based on what is the case, is also about what ought to be the case. Let us now turn to examine a view that challenges some of our rationalistic preconceptions about moral judgment.

3.3 The reasoning agent and the irrelevance of moral reasoning

Deliberative democracy holds that the citizens that come together to deliberate ought to assume a civic standpoint; that they put their own interests aside and orient themselves towards the common good. This picture has been complicated recently, as prominent theorists of deliberative democracy have argued forcefully that self-interest, if suitably constrained, does have a place in deliberation.[72] Though qualified, the moral perspective still forms an essential part of democratic deliberation. But this moral perspective is not static:

71 Stanovich, "Individual Differences," 423-425.

72 See Jane Mansbridge et. al., "The Place of Self-Interest and the Role of Power in Deliberative Democracy," *Journal of Political Philosophy* 18 (2010).

the process of giving reasons that others can comprehend and accept is central to deliberative democracy. This picture seems to entail that in order for deliberation to really be worth our while, the subjects ought to be able to reach the relevant decisions based on moral reasoning given in a deliberation, i.e. based on that specific reason-giving process.

Yet, there are strong challenges to the rationalistic idea that people's moral judgments are the end-product of moral reasoning. It is not possible to fully enter this debate here. Instead, I take a different tack. Suppose that you accept the idea that moral judgment is based on quick moral intuition, and is then followed by slow *ex post facto* moral reasoning. A model based on this view, proposed Jonathan Haidt, is called the social intuitionist approach to moral judgment (SIAM).[73] This model is based on findings in many fields, and these findings cast strong doubt on the idea that moral reasoning is causally effective[74] in moral judgment. But accepting something along these lines does not imply that moral reasoning could not be effective in moral judgment. Haidt notes that his model is antirationalist in a very limited sense.[75] The claim is, specifically, "that moral reasoning is *rarely the direct cause* of moral judgment."[76] It should be noted, as is stressed by Haidt himself, that his model is a descriptive model.[77] It is not a normative model of how moral judgment ought to be made. One could hold that moral reasoning ought to be the sole ruler of moral judgment. I take no issue with that position. But since we are interested in epistemically valuable democratic deliberation that normal human beings can do, we should be aware of the empirical reality.

Let us now view some of the theses that orient SIAM and some features of the model itself. We can be brief.[78] First, moral reasoning is often motivated. Haidt likens it to a lawyer building a case, premised on the idea of defending the client. It is less like a scientist or judge seeking truth.[79] Second, people construct justifications for their moral positions quite easily

73 Haidt, "Emotional Dog and Its Rational Tail: A social intuitionist approach to moral judgment," Psychological Review 108 (1995/2008), doi: 10.1037/0033-295X.108.4.814.

74 It is not commonly accepted in epistemology that reasons-relation can be construed as a causal relation. An originally influential example of a case where reasons, though causally ineffective, justify an epistemic judgment, has been put forth by Keith Lehrer, "How Reasons Give us Knowledge, or the Case of a Gypsy Lawyer," *Journal of Philosophy* 68 (1971). It may be added that not everyone now finds that example convincing. A causal-doxastic theory of epistemic basing has been put forth by Keith A. Korcz, "The Causal-Doxastic Theory of the Basing Relation," *Canadian Journal of Philosophy* 30 (2000), and criticized by Daniel M. Mittag, "On the Causal-Doxastic Theory of the Basing Relation," *Canadian Journal of Philosophy* 32 (2002). Though that debate is ongoing, the idea that moral reasoning does not causally bring about moral judgment seems troublesome from the viewpoint of deliberative democracy. See Keith A. Korcz, "The Epistemic Basing Relation," *The Stanford Encyclopedia of Philosophy* (Spring 2010 Edition), ed. Edward N. Zalta, URL = <http://plato.stanford.edu/archives/spr2010/entries/basing-epistemic/> for discussion and further references.

75 Haidt, "Emotional Dog and Its Rational Tail," 1025.

76 Emphasis added.

77 Ibid.

78 See Haidt, "Emotional Dog and Its Rational Tail" for references to the empirical literature.

79 Haidt, "Emotional Dog and Its Rational Tail," 1033-1035. Deanna Kuhn, *The Skills of Argument*, (Cambridge: Cambridge University Press, 1991), doi: 10.1017/CBO9780511571350, has found that very few groups, philosophers among them, have been found to reason well. What specific implications this has in another matter. For example, Anthony S. Laden "The Justice of Justification," in *Habermas and Rawls: Disputing the Political*, ed. James G.

and are able to achieve a sense of objectivity, feeling that the position is the outcome of the reasoning.[80] Third, moral action correlates more strongly with moral emotion than with moral reasoning.[81] We now turn to the features of the model.

According to SIAM, a given situation elicits an intuition, based on automatic processing. This intuition then turns into judgment, followed by reasoning. Sometimes this reasoning works back to affect the judgment and intuition. The model is a social one in that one individual's judgment and reasoning both in turn affect the fellow human's intuition about the eliciting case (this link is called social persuasion), which further affect judgment and reasoning, which both in turn influence the individual with whom the process started (another link of social persuasion).[82]

What specific implications can we draw from this in respect to deliberation? Haidt (2008: 1045) draws on other researchers, who through arduous processes, have managed to tune up intuitions about justice, rights, and fairness, leading to good moral talk. Haidt argues that if the main obstacle to good moral reasoning is the biased search for evidence, people should take advantage of the social persuasion link. By seeking out discourse partners who are respected for their wisdom and open-mindedness, and by talking about evidence and justification, people can reach judgments that are likely to be more nuanced and reasonable.[83] So it is not the case that we cannot improve our moral judgment, but we do have to make a specific effort for it.

What emerges from this is the importance of certain attitudes and dispositions to the quality of one's moral judgment. One needs to be able to decouple one's prior moral judgment from moral arguments, and assess the relevant arguments for their worth, and be willing to change one's mind based on them. Haidt surmises that by creating a surrounding that promotes good moral talk, i.e. the use of social persuasion link, the researchers might have also improved the quality of personal reflection.[84] But the literature also reminds us of the difficulty: creating stable dispositions for this activity is much more difficult than teaching the relevant skills.[85] In addition, the use of the social persuasion link is not without its problems, and we must now turn to issues in group reasoning.

Finlayson and Fabian Freyenhagen (New York: Routledge, 2011) has argued, basing his view on the work of Rawls and Habermas, that the political sphere is not to be defined by philosophical experts.

80 Haidt, "Emotional Dog and Its Rational Tail," 1035-1036.

81 Haidt, "Emotional Dog and Its Rational Tail," 1036-1038.

82 Haidt, "Emotional Dog and Its Rational Tail," figure 2.

83 Haidt, "Emotional Dog and Its Rational Tail," 1045.

84 Haidt, "Emotional Dog and Its Rational Tail," ibid.

85 Cf. Raymond S. Nickerson, "The Teaching of Thinking and Problem Solving," in *Thinking and Problem Solving*, ed. Robert J. Sternberg (San Diego, CA: Academic Press, 1994), that discusses the problem of transfer from classroom to outside of it.

3.4 The reasoning group

According to Miriam Solomon, the traditional picture[86] of group reasoning has it that a group of individuals deliberating is more likely to reach objectively better results than the individuals of the group deliberating alone. The thought is that the interaction of the group provides for a critical testing of arguments, their explicit premises, and implicit assumptions. But a quick review of the literature shows that we should not take this idea for granted. Cass Sunstein has noted that there are pressures in public deliberation that may lead the group to converge on falsehood rather than truth, and individual errors, instead of being corrected, can be amplified[87]. In the same vein, Miriam Solomon has argued that group deliberation may lead to suppression of relevant data.[88] Yet, neither writer views these problems as unavoidable, but rather as instructive for the design of group deliberation. We should examine these problems briefly.

Both writers partially base their worry on an epistemologically troublesome phenomenon called *groupthink*, identified by I. Janis.[89] Sunstein, drawing on Janis, observes that at times deliberation leads to irrational results, because of two pressures. The first is that group members fail to disclose information out of deference to the information publicly announced by others. The second is that social pressures lead people to silence themselves in order not to face reputational sanctions, such as disapproval by others. These mechanisms can lead to poor results through amplification of individual error. Sunstein further shows that the information that is shared by the participants may affect the decisions disproportionally, and the information may cascade when shared sequentially: the opinions already voiced tend to lead to withholding of private knowledge to the detriment of the quality of the overall decision. Further, the phenomenon of group polarization may affect the reasonability of the group choice: members of the deliberating body may end up adopting a more extreme version of the position they had before deliberation began.[90]

It seems inevitable that the social persuasion link can be used in good and bad ways. As Sunstein notes, the process of group polarization is not entirely unreasonable. If several individual reasoners, based on different premises, arrive at the same conclusion, this does affect the rational believability of the given claim.[91] The process is troublesome only if the

86 Solomon, "Groupthink," 28, identifies the traditional picture with the views of Plato, Mill, Popper, and Longino.

87 Sunstein, "Deliberating Groups." Part of Sunstein's argumentation is based on the heuristics and biases –literature and summarized as "garbage in – garbage out," or worse, in cases of amplification as "some garbage in – more garbage out."

88 Solomon, "Groupthink".

89 Irving Janis, *Groupthink*. (Boston: Houghton Mifflin, 1982). The force of this effect has also been questioned; see Won-Woo Park, "A Review of Research on Groupthink," *Journal of Behavioral Decision Making* 3 (2011). Nevertheless, groupthink seems like a real worry. Sunstein (2006) provides references to many relevant studies.

90 Sunstein, "Deliberating Groups," 197-205.

91 Sunstein, "Deliberating Groups," 204-205. Also, one of the major positions in the epistemology of disagreement is the *conciliatory position*, which holds that when one disagrees with an epistemic peer, one ought to scale down one's belief accordingly. See for example Richard Feldman and Ted A. Warfield, eds. *Disagreement*, (Oxford: Oxford University Press, 2010), for discussion.

agreement affects the judgment disproportionately. If the original basis for the individual judgments is poor, essentially the same, or involves biases, the result may be irrational.

What, then, can be done to combat these effects? Sunstein argues that we might take lessons from the so-called *prediction markets*.[92] This is a fairly new innovation. The central idea of a prediction market is to take advantage of the information-aggregating features of a market. Various such markets have been created and they have done well in predicting future events, and in many domains they have outdone deliberating groups. They impose incentives on diverse people to disclose information without facing reputational sanctions. Because of their structure, they have been able to eliminate, rather than amplify individual errors. They have, in fact, been very accurate, failing only in cases where there has not been much information to disperse. Sunstein also notes that they have not been vulnerable to manipulation. It is not the case that such markets could be used in everything; they do not seem very adequate for dealing with normative matters and the idea of sharing all the relevant information is not always feasible.

Solomon emphasizes that she does not advocate dissent as thought in the traditional picture, which sees dissent as valuable because it leads to criticism, which supposedly leads to the sharpening of views, and ultimately to consensus. The group deliberation may after all lead to amplification of errors too. What is crucial in her view is that dissent is valuable even if there is no discussion, because it keeps the information available.[93]

What we should learn from this, however, is that there should be strong incentives for sharing information: epistemic diversity in the deliberating body ought to be promoted, and dissent valued. Sunstein argues that deliberating bodies should be made aware of the epistemic problems of deliberation. The answer to combating the effect boils down to educating deliberating citizens about the ways that we tend to go wrong on the group level and how information sharing can combat that. Again, we are not determined to do badly in groups, but we need to have appropriate dispositions for reasoning well in groups, and the group reasoning may have to be structured so that irrational effects are marginalized.[94]

4. Discussion

We should now take stock of the empirical material, and its relation to the deliberative ideal described in section 2. After that, I will close by discussing the nature of justification in deliberation.

The literature review above gave reason to believe that although human beings are capable of reasoning well, there are many specifics to our belief-forming methods and strategies that considerably affect its relative reliability. It also emerged that the disposition to apply normatively appropriate methods is a good predictor of normatively appropriate

92 See Sunstein "Deliberating Groups," 205-209. See the text for further references on this fairly new innovation.
93 Solomon "Groupthink," 38-39.
94 Sunstein, "Deliberating Groups," 209.

responses to test questions of different types, after general cognitive ability has been partialled out. The relevant kind of proper examination of claims that ideally results from these dispositions applies both to moral and empirical claims.

In some measure, this picture also lends support to Harvey Siegel's position that a *critical thinker* is the appropriate educational ideal.[95] This ideal contains essentially two elements. First, "[t]o be a critical thinker is to be appropriately moved by reasons."[96] Siegel argues that there is a deep conceptual connection between rationality and critical thinker: "[t]o be a rational person is to believe and act on the basis of reasons."[97] Further, "a critical thinker is one who appreciates and accepts the importance, and convicting force, of reasons."[98] In order to do this, one must understand both the general principles of reasoning and the field-specific criteria of good reasoning, and understand how the field-specific reasons justify through their relation to general principles of good reasoning.[99]

The second, and no less important, part of critical thinking consists of the critical spirit. The critical spirit is realized through certain attitudes and dispositions, habits of mind, and character traits that are required to actually apply the methods of critical thinking. Mere ability to assess the quality of reasoning is not sufficient; one must have the motivation to apply these skills. The ideal critical thinker wants to evaluate the evidential force of reasons, because she values intellectual honesty, justice to evidence, sympathetic and impartial consideration of interests and objectivity.[100]

The notion of critical thinker should be appealing to those deliberative theorists that worry about the epistemic quality of deliberation.[101] If the citizens that enter the deliberative arena are able to approximate this ideal, we would seem to have a better chance of an epistemically justifying deliberative process, and thereby a better chance of reaching a decision that could be the object of free and reasoned agreement among individuals.[102] Such citizens would also seem more likely to fare well in publically justifying public policies to those who are bound by it, and in seeking reasons that you and your opponent can accept.[103] Having such citizens in deliberation would also seem to increase the likelihood that the decision reached is legitimate, because part of the legitimacy of a decision derives from the epistemic value of the procedure.[104] Further, an epistemically commendable process also requires that we give the relevant preferences and reasons of each participant

95 Siegel, "*Educating Reason*."
96 Siegel, "*Educating Reason*," 32.
97 Siegel, "*Educating Reason*," ibid.
98 Siegel, "*Educating Reason*," 33.
99 Siegel, "*Educating Reason*," 37
100 Siegel, "*Educating Reason*," 39.
101 Though even if one did not accept any epistemic commitments in one's justification of deliberation as a political method, one could still be moved by the educational ideal.
102 Cohen, "Deliberation and Democratic Legitimacy," 73.
103 Gutmann and Thompson, *Democracy and Disagreement*, ch. 2.
104 Estlund, "Beyond Fairness and Deliberation."

sufficient time to influence the deliberation.[105] Without proper time, the participants are not able to form adequate conception of the overall justification of dissenters' position and take advantage of the epistemic value of dissent or the social persuasion link. A critical thinker would know this.

There are then empirical, epistemological, moral, and political reasons for teaching not only the skills and norms of proper reasoning but also the appropriate dispositions and attitudes that support the use of those skills and norms. This seems to have strong educational implications on teaching critical thinking. However, I will not discuss that issue further here. Instead, I will close by examining Solomon's argumentation that the problems of group rationality show that epistemological internalism is not acceptable.[106]

Solomon argues that the facts of group deliberation support epistemological externalism. The traditional picture of deliberation is internalist, she argues, because it is premised on the idea that rational dialogue between two individuals improves the reasoning of those individuals: it corrects errors in reasoning, exposes presuppositions, and transmits new evidence. Each individual is able to improve, because the evidence is weighed and errors are pointed out to each in the dialogue. Such a dialogue makes all the individuals reflect on all the relevant evidence in the manner that internalism seems to entail. But as has been observed, the process does not always correct: it may preserve or even amplify errors. Therefore, deliberation, meaning (here) the internalist common reflection of reasons, does not necessarily produce epistemically good results. Therefore, internalism is false. Externalism is the denial of internalism. Therefore, it is true.

The idea that dialogue and mutual testing of claims is epistemically beneficial is indeed traditional, but it also includes many further theses that internalism in itself does not entail.[107] Solomon[108] cites the definition given by a representative internalist, Roderick Chisholm[109]: "If a person is internally justified in believing a certain thing, then this is something he can know by reflecting upon his own state of mind." It is indeed the case that this is the central premise of internalism and denying it makes one an externalist. However, as James Pryor has pointed out, the *simple internalist position* is that justification supervenes on facts that one is in a position to know by reflection alone.[110] But only a separate and stronger claim, known as *access internalism*, demands that one always has special access to one's justificatory states, and deliberation without access to one's reasons seems difficult. Access internalism is stronger, because it is possible for the justification to supervene on

105 Dahl, "Procedural Democracy," 109-110.

106 Solomon, "Groupthink."

107 Admittedly, if one finds the ideal of a critical thinker, or some essential parts of the critical thinking, appealing, one might also find some form of epistemological internalism appealing. The kind of analysis and evaluation of evidence and different arguments emphasized by the ideal seems very congenial to internalism. But we must clearly separate the deliberative ideal and the critical thinking ideal from the epistemological thesis of internalism.

108 Solomon, "Groupthink," 28-29.

109 Roderick Chisholm, *Theory of Knowledge* (Englewood Cliffs, New Jersey: Prentice Hall, 1989), 7.

110 James Pryor, "Highlights of Recent Work in Epistemology: A Survey Article," *British Journal for the Philosophy of Science* 52 (2001): 104-106, doi: 10.1093/bjps/52.1.95.

internal things, which one is in a *position* to know, without actually being able to access those reasons. This could be, for example, because one is epistemically untrained, or, as the case might often be, one has wrong beliefs about what it takes to be justified. So, one could accept simple internalism without requiring that one is able to justify one's beliefs in a debate.

But even accepting access internalism does not imply any commitments on the traditional picture of deliberation or its epistemic qualities. Nor should we claim that a typical epistemological internalist holds the so-called dialectical condition of justification: that a necessary condition of justification is that one is able to defend one's position in a public debate.[111] A dialectical condition is not popular among epistemologists.[112] The standard counterexample is the case of a diffident schoolboy asked about the year on which the Battle of Hastings was fought. Being shy, he is unable to produce the correct answer, though he believes that it was 1066, and remembers well the justifying source on which he formed this belief. Dialectical success is not sufficient for justification either, because one might succeed in defending one's claim with biased evidence, fallacious arguments, or because of informational cascading.

So the proper target of Solomon's criticism is not internalism, it is rather the assumption that any deliberation improves reasoning, or that joint deliberation always produces epistemically the best results. One might read this as another reason for not supporting the dialectical requirement of justification, given that we are justified in identifying the dialectical condition as the appropriate target of her criticism. The fact that a matter is debated does not always improve the quality of the relevant beliefs. For the deliberative theorist the dialectical criterion is essential, but he or she need not take a stand on the internalism/externalism –debate. Yet, the deliberative ideal introduced in section two certainly seems congenial to internalist aspirations. The empirical results suggested that the ability to distance oneself from one's own beliefs, to assess the evidence for them, and to monitor one's belief-forming practices is a key factor in improving the objective quality of one's beliefs. This seems to support an internalist approach to justified beliefs.

Solomon's point that a group deliberation is not necessarily epistemically more efficient than a process where the same individuals reflect and give their answer independently seems undeniable. But internalism does not force one to choose one method over another. Instead, it holds that the participants are only justified in their beliefs, and their consequent answers, if they can know this by reflection alone. They would know it by reflection

111 For the view that justification requires that one is, under normal conditions, able to do so, see Adam Leite, "On Justifying and Being Justified," *Philosophical Issues* 14 (2004), doi: 10.1111/j.1533-6077.2004.00029.x. Regardless, we must differentiate two different ideas. It is one thing to hold that a normal human in normal circumstances ought to be able to justify a claim publicly, because one's justification is typically such that one can transfer it into a feasible public case. It is another thing altogether to hold that justification is constituted by this ability.

112 But see for example Edmund Craig, *Knowledge and the State of Nature* (Oxford: Clarendon Press, 1990). See for example Markus Lammenranta, "Disagreement, Skepticism, and the Dialectical Conception of Justification," *International Journal for the Study of Skepticism* 1 (2011), for further discussion.

alone if they were, for example, able to access all their reasons for their answer. We are justified in believing the aggregate result of the prediction market only if we can know by reflection alone whether the result is justified. We check the result, compare it with reality, repeat the test often enough times, and finally reach a state of mind based on which we can say that we are justified in believing that a prediction market produces results worth believing.[113] Internalists and externalists place different demands on the chosen method of belief-formation, whatever the method is. Personally, I would be quite surprised if the individuals responsible for the aggregate result typically had no reasons for their responses. But if they typically did not and yet produced the correct result, this would cast a doubt on the importance of internalist requirement. However, what is important here is that internalist need not require that there is always a deliberation. If some questions are best treated first separately and then aggregated, so be it. But there would still be a condition on being justified in believing the individual result and the aggregate result.

In any case, the internalist criterion of justification is not sufficient for proper deliberation. It is not enough to be justified in believing or know that one's proposal is good; one must be able to defend it sufficiently clearly in relevant arena from mutually acceptable premises.[114] But as a deliberative requirement, the dialectical condition is just what is needed and the deliberative theorist should be aware of the problems that groupthink can produce. There are many suggestions on how to improve group reasoning. Solomon notes that some structure must be followed to avoid the effect, but more research on the topic is needed.[115] But to put the result into perspective, we might observe that it has not been shown that a deliberation could not justify the conclusion, or that it goes wrong more often than not, or that in general the truth-approximating qualities of deliberation do not meet some intuitively acceptable threshold. Deliberation has only been shown to be vulnerable to problems, especially under certain conditions. But perception, memory, intuition, and testimony are also known to be subject to problems of their own, and that has not led us to conclude that they cannot justify. We have discussed various reasons why deliberation might produce unjustified results. But we have also seen reason to believe that given that one's case for a given belief is examined by one's epistemic peers, who have been properly informed[116] about the issue at hand, care about the common good, and who are

113 The mere fact that sensory perception, external evidence, must be used to reach a justified belief about the reliability of prediction markets does not support externalism.

114 Supporting the traditional picture does not in any case entail that dialogue is the only, or the most important, test of justification. One might just assume that over an extended period, open dialogue is likely to zero in on the truth, which is again a different claim.

115 Solomon,"Groupthink," 32.

116 Deliberations may be organized so that the deliberators are provided with expert information about the issue both in written and oral form. This was the case, for example, in the study by Maija Setälä, Kimmo Grönlund, and Kaisa Herne, "Citizen Deliberation on Nuclear Power: A comparison between two decision methods," *Political Studies* 58 (2010). This study did not find systematic evidence of social pressure between groups that ended their deliberation with a secret ballot and groups that had to formulate a common statement, thus countering the evidence discussed by Sunstein. However, in addition to this availability of expert knowledge (which Cass Sunstein ("Group Judgements: Statistical Means, Deliberation, and Information Markets," *N.Y.U. Law Review* 80 (2005): 1011)

able to override system 1 judgments with system 2 processing, because they possess the relevant epistemic dispositions, some justification is incurred.

5. Conclusion

This essay started out by formulating an ideal picture of democratic deliberation and the deliberative agent. The ideal of deliberative democracy requires that one is rational in deliberating about a given policy. To participate in a deliberation, one must accept that the process, in which the reasons are given and publicly scrutinized, aims to produce a rational consensus, and thus a decision on the issue, based on which a given policy is implemented. The rationality of this process is dependent on the epistemic behavior, acts, and dispositions of the agents participating in it. This means, among other things, that one ought listen to each participant, to analyze the strength and evidence presented in the discussion carefully, and the reasoning from that evidence to beliefs, and further to policy decisions. It was also noted that the ability to take a skeptical stance and evaluate matters critically was more important than individual computing abilities.

This ideal was then compared with some empirical results about the reasoning capabilities of the typical citizen and groups of citizens. It was argued that although there are serious problems in the quality of our reasoning, there is no evidence that human beings could not deliberate rationally. However, the empirical literature tells us that in many cases a critical thinker would be justified in thinking that the mere fact that one's fellow citizen believes something is not sufficient for the critical thinker to believe that the belief's content is true. But it must be borne in mind that the debate on the significance of the empirical evidence is complex. Adler (2008) argues that it has not been shown that the empirical results of Heuristics and Biases cast no shadow on the typical justificational methods of the average citizen and I agree.[117] Though the implications of these results need to be put in perspective, the justificational problems do not disappear altogether.

So, in objective epistemic terms, it seems that one ought to carefully examine the evidence one is given by one's fellow citizens, and its implications to the deliberative decision at hand. We all fall victim to epistemically inferior practices occasionally. The testing function of public deliberation can be useful to differentiate the beliefs based on bad practices from beliefs based on good practices, but also to prevent the use of new bad epistemic practices and support the use of good practices. Yet, we also saw that unless the procedure of deliberation is controlled, it is liable to produce problems of it own.

---

has acknowledged as a possible factor reducing error amplification), these groups were composed of people from different segments of society (also noted by Sunstein ("The Law of Group Polarization," *The Journal of Political Philosophy* 10 (2002): 186, doi: 10.1111/1467-9760.00148) as a reasonability enhancing factor, because it counters group polarizing), was guided by rules of procedure, and had an impartial moderator present.

117 Jonathan E. Adler, "Presupposition, Attention, and Why-Questions," in *Reasoning*, ed. Adler and Rips.

Both the ideal of deliberation and the ideal of critical thinking could make one hospitable to the epistemological thesis of internalism, but this conceptual matter was not treated extensively here. It was only argued that the traditional picture of epistemically beneficial dialogue contains more commitments than internalism, and that at least parts of the empirical evidence can be interpreted so as to support internalism. Noting both the fallibility of individual and group reasoning and the importance of the political decisions, the kind of reflection required by internalism (and the critical thinking –ideal) seems essential for justification. To reach normatively appropriate results reflection of one's beliefs requires that one justifiedly believes the relevant premises and principles and methods of reasoning. To actually reach justified beliefs with a meaningful frequency requires further that one possess certain dispositions, attitudes, and habits of mind. This means that we also reflect on the ways we form beliefs: every once in a while, we must be able to ask "why do I believe *that*?" A democratic deliberation, a natural habitat for disagreement, is just one context where this should happen. We must be willing and able to ask ourselves why we believe as we do.

Regardless, part of the motivation of a rational code of conduct for the deliberation is not solely epistemic: the basis is partially moral and political. Because a deliberation must end with a decision, a common ground must be found, even when the prospects for establishing it are not good. Whether one is an externalist or internalist, one has a right to demand that decisions are justified with reasons one can accept. One has that right as an epistemic agent but also as a citizen bound by the decision. The epistemologist can use the inability of the diffident schoolboy to establish that showing and knowing are different things. The political philosopher must instead show how we can manage to decide together despite this difference. The empirical evidence that we have examined here does not rule out the possibility of reasonable democratic deliberation, given appropriate education of the deliberating agents.