The (Severe) Limits of Deliberative Democracy as the Basis for Political Choice*

Gerald F. Gaus

1. A Puzzle: The Majoritarianism of Deliberative Democracy

As Joshua Cohen observes, “[t]he notion of a deliberative democracy is rooted in the intuitive idea of a democratic association in which the justification of the terms and conditions of association proceeds through public argument and reasoning among citizens” (1989: 21). Deliberative democrats insist that this deliberation must be public in a radical sense: only reasons that can be embraced by all of us are truly public, and hence justificatory. As Gerald Postema has put it, a public reason must be a shared reason (1995b: 70). Postema, and most deliberative democrats, insist that this idea of public, shared reasoning, is not simply a theoretical commitment, but a regulative aim of practical politics: “[a]greement among members of the community is set as the open-ended task or project of ...[the] exercise of practical reason and judgment” (Postema, 1995a: 356). Thus “[t]he aim of the regulative idea is agreement of conviction on the basis of public reasons uttered as assessed in public discourse...” (Postema, 1995a: 356).

Of course advocates of deliberative democracy are well aware that complete actual consensus is not a reasonable aim “even under ideal conditions” (Postema, 1995a: 356). The deliberations of citizens may not yield a consensus, and so we may have to cut the discussion off by taking a vote (Postema, 1995a: 356). Habermas’ view is more complex. His “two-track” conception of deliberative politics conceives of an interplay between democratic procedures and deliberative collective will formation: the procedures do not

*I am especially grateful for the insightful comments and criticisms of Jon Anomaly and Tom Christiano.
simply follow deliberation (“cut it off”), but also affect the formation of a deliberative rational agreement (1996: 304ff.). Still, Habermas is quite clear that “[p]olitical deliberations...must be concluded by majority decision in view of the pressures to decide” (1996: 306).

It is not a puzzle that, because of the practical requirements of politics, we cannot wait for consensus to emerge and so must vote. What is, I think, puzzling, is that deliberative democrats, seeing that actual consensus is an impossible goal, suppose that the majority rule procedures are the best way to cope with the demands of practical politics. If the regulative ideal is agreement among N citizens, why should it be the case that the preferred institutional approximation is that measures are adopted if and only if they are supported by more than N/2 voters?

In this essay I argue that majority rule is most unlikely to be the procedure that would be justified through an appeal to public reason: there is a strong case that the optimal deliberative democratic voting rules require extraordinary majorities. And because of that, I shall argue, deliberative democracy is severely limited as a political theory about how actual disputes are to be resolved under modern conditions of deep disagreement. Sections 2 and 3 consider important preliminary matters concerning how voting and public reason are best conceptualized. Although we often refer to “deliberative democracy” as if it is a single view, there are many accounts of deliberative democracy (Dryzek, 2004). We need to be careful about just what form of deliberative democracy we have in mind. Section 2 examines three competing views of public reason and deliberation while Section 3 distinguishes two ways that voting can seen as expressing public reason. Sections 4 and 5 then advance the main argument of the paper: drawing
on the classic work of James Buchanan and Gordon Tullock (1962), I argue that, given
the preferred understandings of public reason and voting, the best voting rule for
deliberative democracy is a supermajority rule. In Section 6 I argue that this result shows
that deliberative democracy is often ill-equipped to provide guidance as to how actual
political disputes are to be legitimately resolved. As a political theory of actual political
choice, I shall argue, deliberative democracy will often be indeterminate: it provides no
guidance about what we should do, just when we most need it.

2. What is the Relation between Public Deliberation
and a Publicly Justified Outcome?

2.1 The aim of deliberative democrats is to reason together so as to arrive at publicly
justified shared moral norms, political laws, and public policies.¹ Let us call the outcome
of appropriate public deliberation on issue I the publicly justified position, or the publicly
correct position, on I. Theories of deliberative democracy differ in how they relate
appropriate public deliberation to the publicly correct answer. We can distinguish at least
three different accounts; let us begin with

Collaborative Reasoning:² A deliberation D is collaborative only if (i) D is about some
issue I for which there is a deliberation-independent criterion of a correctness (e.g.,
justified true belief regarding I); (ii) participants to D about I are more likely to arrive
at this correct outcome than are individuals who do not participate in D; and (iii)
that a large proportion of the participants disagree about the publicly correct answer
regarding I is a strong reason to question anyone’s claim to have reached the publicly
correct answer regarding I.
In collaborative reasoning we reason together about some matter in a way that makes it more likely that we arrive at the publicly correct view. Not all sound reasoning is collaborative. While it might reasonably be argued that good reasoning qua presenting arguments and replying to objections is inherently collaborative in the sense of (i) and (ii), claim (iii) is an additional and stronger claim. According to (iii), if after engaging in deliberation on \( I \), the participants still disagree, then every participant has strong reason to doubt his or her own view. Clause (iii) points to a conviction that, on this matter, reasoners really ought to come to the same conclusion, and something has gone wrong when they do not. I take it that Rawls denies that reasoning about conceptions of the good is collaborative in this sense; competing credible or reasonable conclusions about conceptions of the good are “the natural outcome of the activities of human reason under free enduring institutions” (1996: xxvi). In contrast, Rawls does think that deliberation about matters of justice ought to tend towards consensus, at least among reasonable people. This is important. Any account of collaborative reasoning will have to identify a class of competent participants in the deliberation, whose dissent is relevant to (iii).

Reasoning in a scientific community is, at least on most accounts, collaborative. Persistent disagreement within a group of competent inquirers about \( I \) is an indication that something is amiss: none of the positions are sufficiently well-justified to claim the truth, insofar as the deliberation about \( I \) has not reached its conclusions. Recall here C. S. Peirce’s notion that truth is that which a diligent community of inquirers would come to agree on (Misak: 2000). That some competent members of the community disagree is an indication that deliberation has not yet reached its successful conclusion. One of the
revolutionary features of Thomas Kuhn’s (1970) analysis of science was precisely to deny that scientific inquiry was collaborative in this sense; only within a shared paradigm, Kuhn maintains, does scientific inquiry have this collaborative character.

2.2 That science is typically understood as an instance of collaborative reasoning suggests that it is not a sufficiently strong sense of public reasoning to ground deliberative democracy. Some, perhaps, conceive of deliberative democracy as akin to scientific inquiry—for example, those who see it as an epistemic procedure seeking to uncover objective truths about morality (e.g., Estlund, 1997)—but more usual, I think, is to conceive of public reason as more intimately tied to what the parties actually believe rather than a search for an external truth. The publicly correct position is constructed out of the reasoning of the citizens. Let us call this

Public Constructivism: Deliberation \( D \) about some matter \( I \) is a case of public constructivism only if (i) there exists some justified function \( f \) that generates from the set of all individual beliefs, values, etc., a publicly justified position on \( I \) and (ii) \( D \) tracks, approximates, or tends to \( f \).

Like accounts of collaborative reasoning, plausible theories of public constructivism in some way idealize the set of individuals from whose beliefs and values public reason is constructed. For example, rather than claiming that a public reason is constructed out of the set of beliefs, values and desires of each and every person as now constituted in his or her present position, a public constructivism may maintain that the relevant set is all citizens considered as reasonable or rational. Rawls takes this route, excluding unreasonable beliefs and judgments from the set of individual judgments on which the
public constructivist function \(f\) operates (1996: 53-54). As Rawls sees it, a slave-owning class which insists that the two principles of justice do not express public reason—because the slave owners reject them in favor of a system allowing property rights in other people—is unreasonable. “We say: ‘Their proposal was perfectly rational given their strong bargaining position, but it was nevertheless highly unreasonable, even outrageous’” (Rawls, 1996: 48). For our purposes, the important point is that such constructivist theories understand public reason as an expression of shared judgments under some idealized condition, e.g., assuming individuals are rational or reasonable, and are addressing considerations to others. Let, then, \(\{b_1, ..., b_n\}\) be the belief systems of suitably idealized reasonable persons 1 through \(N\) on matter \(I\); the basic idea of public constructivism is that there is function \(f\) such that \(f\{b_1, ..., b_n\} \rightarrow \alpha\) means that \(\alpha\) is the correct public position about \(I\). That is, the public constructivism function operates on the set of suitably constrained individual beliefs to yields the publicly correct position in \(I\).

2.3 A public constructivism must somehow take into account every idealized person’s beliefs, values, reasons, etc., and construct public reason out of it. The most obvious candidate is a consensus justification; on a consensus justification, \(\alpha\) is publicly justified if each suitably idealized member of the public holds \(\alpha\). Consensus justification is at the heart of Rawls’s political constructivism, which seeks a free-standing view built on shared “fundamental ideas that are drawn from the public political culture of a democratic society” (Rawls, 1996: 25n). A justification, however, can also be by convergence (D’Agostino, 1996: 30-31). Here, instead of justifying \(\alpha\) by showing it to be shared by all
suitably idealized members of the public (or built up from shared ideas), it is shown that each has her own, often different, reasons to endorse $\alpha$. Both convergence and consensus public constructivism are versions of collaborative reasoning. At least as I have depicted them, both hold (i) that there is a deliberation-independent correct answer (what the public constructivist function would generate given the set of idealized individual beliefs, values and desires), and (ii) insofar as the aim of politics is to uncover the publicly correct answer, that actual political deliberation tends to track this function, i.e., yield the correct public answer. Thus one can adopt public constructivism while also adopting collaborative reasoning: a public constructivist can accept that there is a deliberation-independent criterion of correctness (say that $f$ yields $\alpha$), that participants to the discourse ($D$) are more apt to arrive at $\alpha$, and that continued disagreement among the participants of the discourse is a significant reason to question one's opinion that $\alpha$ is publicly justified. It is important to keep in mind here that even if $f$ is depicted as a sort of deliberation, such as discussion in the original position, this does not undermine the claim that the criterion of correctness is independent of $D$—our actual political deliberation about matter $I$.

2.4 A form of public constructivism that apparently is not also a form of collaborative reasoning is suggested by the Habermas's (1987) discourse theory. On this view the justified public construction is closely linked to actual public discourse and the way in which people's beliefs may change as they are tested in actual exchange. This sort of public constructivism characterizes public reason as a vector of $\{b_1...b_N\}$ about $I$ which either does not (in the rare case) depart from $\{b_1...b_N\}$, or for which there is a series of
points which “dialogically” lead to publicly correct answer to \( I \). On the face of it, there is no deliberation-independent criterion of correctness; the publicly justified outcome is that to which deliberation leads, given the initial set of individual beliefs, reasons and desires. On further reflection this is not perhaps quite so obvious. The deliberation that yields public justification is still a suitably idealized deliberation, call it \( D^* \). Our actual deliberation, \( D \), is not identical to \( D^* \), since \( D^* \) is, say, a deliberation conducted in complete conformity to certain norms of rational discourse — norms to which we never quite conform. But now we once again have a deliberation-independent criterion of the success of \( D \), i.e., what \( D^* \) would yield. This looks just like Rawls’s theory, where correctness is characterized in terms of an ideal deliberation (in the original position), and a justification of actual political discourse is that it tracks the idealized discourse.4

The important issue here is the relation between \( D \) and \( D^* \). If, as in the case of Rawls’s theory, \( D \) and \( D^* \) are very different types of deliberation—as different as the actual politics (\( D \)) and the original position (\( D^* \))—then it seems that the criterion of correctness really is distinct from \( D \). On the other hand, if \( D^* \) is a modest idealization of \( D \), or \( D \) minus its manifest failures, then, while it is strictly true that the criterion of correctness is independent of \( D \), it also seems, in a pretty obvious way, dependent on \( D \) as well. Let us call public constructivist views according to which the actual deliberation \( D \) is significantly constitutive of public reason deliberative constructivist theories of public reason. To be sure, this is vague, and there is room for considerable disagreement about what is a “significant” constitutive relation, but it is, I think, clear enough to allow us to proceed.
2.5 Deliberative constructivist theories of public reason are, I submit, quintessentially deliberative democratic. I do not deny that versions of deliberative democracy can be based on other theories—Rawls (1997: 772) declared that he was a deliberative democrat—but deliberative constructivist theories make the most sense of the insistence that actual deliberation be taken as far as possible towards actual consensus. If actual discourse is only taken as a path to a deliberation-independent criterion, as in collaborative reasoning, there really is little reason to continue the deliberation until all agree. Although rational believers have reason to reconsider their views when others disagree, when most believe $\alpha$ and a relative few believe $\beta$, there is relatively little epistemic reason to continue deliberations until all accept $\alpha$. If, however, we embrace a deliberative constructivist theory of public reason, absence of agreement is a much deeper worry. Even if most have embraced $\alpha$, that a small minority still embraces $\beta$ is a serious concern, for on this account public reason is supposed to be a vector of all competent deliberators' views. Consequently, we shall see that whether public reason is conceived as collaborative reasoning or as deliberatively constructed has significant consequences for the choice of voting rules.

3. The Aim of Voting

3.1 Voting can be understood as either evidence of what public reason is, or as constitutive of it. Collaborative conceptions of public reason, which possess a deliberation-independent criterion of correctness, do not conceive of voting as constitutive of public reason: that is, they deny that the voting process simply constitutes public reason.
Collaborative conceptions of deliberative democracy, then, will see voting as evidentiary of public reason. That a vote has taken place and \( \alpha \) has won is, on this view, evidence that \( \alpha \) is the deliberation-independent correct view of the matter. On this view, then, deliberation is a way to enhance the inputs into the voting procedure. Since deliberation tends to produce agreement on the correct view of the matter, votes taken after deliberation have a higher evidentiary value.

3.2 Deliberative constructivist accounts of public reason are apt to be confused with views that conceive of public reason as constructed through voting. Some followers of Rousseau—or, at least, some interpretations of Rousseau—have made a similar error. Rousseau certainly had a constructive account of the general will: the general will was formed from set of individual wills, suitably idealized (Gaus, 1997). Rousseau describes the general will as the will of the people. He also, though, insists that if the people deliberate under the proper conditions, their decision will reveal the general will (Rousseau, 1997 [1762]: 59-60). But, of course, Rousseau also held that voting reveals the general will. William Riker (1988: 11) interprets Rousseau thus:

The fundamental notion goes at least back to Rousseau. There is a social contract, which creates a “moral and collective body” that has “life” and “will,” that is, the famous “general will,” the will of the incorporated people, the Sovereign....The way to discover the general will...is to compute it by consulting the citizens.

Riker, drawing on the findings of social choice theory, argues that since no voting system is guaranteed to give a rational choice while meeting minimum conditions of fairness and logicality, no voting system yields the uniquely correct way to aggregate individual votes
into an overall social choice. And so it makes no sense to claim that voting constitutes the popular will. Social choice theory effectively undermines any claim that voting constitutes a uniquely correct way to amalgamate individual preferences. But deliberative constructivist accounts of public reason do not accord this role to voting: because it is deliberation that constructs public reason, it cannot also be voting that does so. Unless it just so happened that the construction of public reason through voting was equivalent with the construction through deliberation, these are inconsistent ways of constructing public reason. Therefore, like the collaborative conception of public reason, the deliberative constructivist view must take voting as evidentiary of public reason.

3.3 What, then, is the aim of voting for a deliberative constructivist? What is it evidence of? One obvious proposal is that the perfect voting procedure would give us a totally reliable answer as to whether some law $L$ accords with public reason on matter $I$. However, there is strong reason for a deliberative democrat to reject this idea. To say that the application of the ideal voting rule would, with total reliability, yield the correct judgment as to whether $L$ accords with public reason on $I$ is to say that the voting rule is a perfect indicator of whether $L$ conforms to $D^*$, the idealized deliberation. This, though, would be to transform our constructivist account into a collaborative conception, for it seeks voting rules that reveal a $D$-independent criterion of correctness. To put it bluntly: on this view actual deliberation ($D$) is no longer constitutive of public reason, since the voting procedure directly tracks $D^*$. A deliberative constructivist account must tie public reason closer to the outcome of $D$ itself (while still admitting that $D$ is itself an approximation of $D^*$). Perhaps a deliberative constructivist would wish the voting
procedure to be a totally reliable indicator of the outcome of $D$—what $D$ would conclude were $D$ carried to its conclusion. But this too seems inconsistent with—or, at least, in tension with—deliberative constructivism. According to deliberative constructivism, public reason on matter $I$ is a vector of suitably idealized citizens' beliefs, desires and values concerning $I$. Discussion and deliberation will alter many of these views, so that at the conclusion of the deliberation most people's, perhaps everyone's, set of beliefs and values is apt to be different from that with which they started. Public reason thus constructed is almost certainly path dependent. Supposing a certain initial set of individual beliefs, etc., $\{b_1 \ldots b_N\}$, the final set of beliefs $\{b^*_1 \ldots b^*_N\}$ is apt to depend on the order in which alterations to the initial set are made. An early concession by $b_1$ may change the terms of the deliberation, producing a different path to a consensus than would have occurred had $b_1$ resisted alterations of his belief until late in the day.

Because of this, it would be exceedingly difficult—really impossible—to predict where this process will lead. Suppose we take a vote at time $t$; participants not only have to know the state of their deliberations about $I$ at $t$, but be good predictors of what paths deliberation will follow after $t$, so that they could predict the ultimate projected end of the deliberations. Because such estimates are terribly hard to make, the deliberative constructivist stresses the need for actual discourse; it is the actual path $D$ takes that (at least partly) determines the publicly correct position on $I$. In light of this, a vote ought to provide evidence of the state of deliberation at the time of the vote. If public reason is to be seen as the outcome of $D$, and if voting is a way of coping with the fact that we must act before $D$ has reached its outcome, a deliberative democrat should want the vote to reflect how far $D$ has proceeded. Given path dependence, all that really can be known
about the requirements of public reason at $t$ is the state of $D$ at $t$.

Yet this raises a puzzle. If we accept that voting “cuts short” deliberation, then we know that $L$ has not yet been vindicated by $D$; the deliberation has not been completed. If we are not trying to predict what would happen if the deliberation continued to its conclusion, what is the point of voting now? Let us call $L$ quasi-vindicated at $t$ if no consideration has been advanced at $t$ that shows that $L$ has failed as a candidate for the publicly justified position on $I$. A proposal that is quasi-vindicated at $t$ may fail to be vindicated later in the deliberation: perhaps $L$ will have to be modified to meet an objection (again we confront path dependence), or dropped altogether. But given that this cannot be known, the best voting procedure for which the constructivist deliberative democrat can reasonably hope would be one according to which, if a vote is taken at time $t$, $L$ will pass if and only if it has been quasi-vindicated at $t$. To ask more—that it predicts whether at the end of the deliberation $L$ will be finally vindicated—is to ask that the voting procedure outpaces actual deliberation.

3.4 Suppose, then, that a deliberative constructivist conceives of voting as a reliable indicator of whether $L$ has been quasi-vindicated by $D$ at the time of the vote. There are still two ways to conceive of what each voter is up to. We may see each voter as seeking to answer the question: “What is the current state of $D$? Has $L$ been quasi-vindicated in our deliberation?” Given the previous discussion, this would seem an attractive interpretation. If the aim of voting is to indicate the state of $D$ at the time of the vote, it would appear that each voter should seek to give her best judgment of whether, given the overall state of $D$, $L$ has been quasi-vindicated, i.e., thus far seems to conform to
public reason. However, there is decisive reason to reject this conception of what voters ought to be doing: the problem is dispersion of information. As F.A. Hayek showed in his analysis of the market, each participant has extensive personal and local knowledge, but no one has comprehensive knowledge of the state of the total system (see Hayek, 1991; Barnett, 1998: 30-35). My “local knowledge” concerns facts available to me that others do not have (for example, what arguments I have heard); “personal knowledge” includes facts personal to me, such as my values or comprehensive doctrine — such knowledge is relevant to my evaluation of a proposal. As a general rule, no one but me has adequate knowledge of my values and how they relate to the acceptability of a proposal.

Now in the market, the price system functions as a way of communicating personal and local knowledge throughout the system. Attempts to act directly on comprehensive knowledge of the system, as in central planning, fail: no single actor, or small group of actors, possesses the extensive knowledge that is summed up through the price system. For the same reasons that central planning fails, so too would a deliberative democratic vote that asked each voter to give her opinion of the overall state of \( D \) at \( t \).

That information is tremendously hard to get; any evaluation of it is apt to be grossly simplified and distorted. What each voter does possess—again, this is precisely parallel to the market case—is personal and local knowledge of her evaluation of the current state of \( D \) from her perspective: i.e., whether or not, in her opinion, \( L \) has thus far been vindicated in \( D \). Given that the constructivist vindication of \( L \) requires that her set of beliefs vindicate \( L \), a crucial piece of each voter’s local and personal knowledge is whether \( L \) does that. The important point is that each voter employs her most extensive knowledge when she judges not whether the current state of \( D \) endorses \( L \), but whether her own beliefs and
encounters lead her to endorse L.

4. The First Desideratum: Avoiding Oppression

4.1 A primary desideratum of a deliberative constructivist voting rule is that our laws respect everyone: no one is subject to a law that her reason cannot endorse. So the first desideratum is for the rule to avoid false positives: laws that pass the test of the voting system but that are not quasi-vindicated. False positives are oppressive: they subject some citizens to laws which the public deliberation already has shown are objectionable given their reasoning.

Assume that (i) all voters have a strict preference regarding L—no one votes who is in fact unable to say whether L has or has not been vindicated by her personal and local knowledge. Assume also that (ii) the population of actual voters is a large proportion of potential voters (i.e., competent adult citizens). This second assumption supposes that D has proceeded sufficiently far that most citizens have come to at least a qualified judgment on L (the proportion of citizens who refrain from voting is modest), and that the citizens are not apathetic, disinterested, or seeking to free ride on the participation of others. The overall political theory of deliberative democracy clearly endorses these assumptions as reasonable characterizations of a polity in which deliberative democracy obtains. If there is widespread non-participation, then clearly public deliberation has failed; and if the vote occurs before the deliberation has proceeded to the point where most people have, at least tentatively, formed an opinion, there is little point in characterizing the vote as a part of deliberative democracy.

Given these assumptions, we can construct functions that relate the percentage of
voters who vote for \( L \) to the likelihood that \( L \) is a false positive: that it is not quasi-vindicated in \( D \). Call these curves the \textit{Tendency to Avoid Oppression Functions}. Figure 1 below gives three prima facie plausible relations.

4.2 According to Curve A, the “tendency to avoid oppression” (that is, the tendency to avoid false positives) is very high when over 50 percent of the voters endorse \( L \).\textsuperscript{6} Almost all the costs in terms of possible oppression have been eliminated by the majority rule \( (N/2+1) \) mark: after that, increased support for \( L \) does very little to further reduce the possibility of oppressive laws. So if our aim is simply to avoid oppressive laws, \( N/2+1 \) would, according to the reasoning behind curve A, tend to get most of the benefits. The Condorcet jury theorem would seem to endorse the general shape of A—curves that are flat at (what I shall call) the \textit{top end} (i.e., from 51%-100%). For example, if we suppose that each voter is, say, a bit more likely to be correct than incorrect as to whether \( L \) has
been quasi-vindicated, a simple majority rule with ten thousand voters would select L only if it was quasi-vindicated with over a ninety-nine percent reliability (Kuflik, 1997: 306). This would yield a curve that is essentially flat at the top end.

This line of reasoning may well be relevant to a collaborative conception of public reason. It will be recalled that, according to collaborative accounts, the point of public reasoning is to arrive at a deliberation-independent conception of correctness (§2.1). Condorcet's jury theorem seeks to show that majority rule is a highly efficient way to arrive at shared true judgments. To be sure, the literature is filled with doubts about the relevance of Condorcet-type reasoning for democracy; I shall not seek to pursue these doubts here (see Gaus, 1997, but also Goodin and Estlund, 2004). My point is that those who advocate collaborative public reasoning are apt to find Condorcet-type arguments highly relevant; Condorcet's model of the problem seems to conform to the crux of collaborative reasoning.

It is not, however, appropriate to deliberative constructivism. As I have explicated it, deliberative constructivism does not construe voting in terms of diverse answers to the same question, in which your answer can be construed as an “observation” that gives me reason to reconsider mine. Each of us have differing personal knowledge: my personal values ("conceptions of the good") are relevant to what laws I can rationally embrace, and (we can say as a general rule), others do not generally have knowledge of them, nor are they deliberating on the basis of them. Given this, to some extent, we are all asking different questions: how does L fare vis a vis my values, concerns etc.? Hence Condorcet reasoning — which is about many people answering the same question— gets no real foothold.
Because of this, curve A is an implausible conjecture about the relation of the tendency to avoid oppression and degree of voter support, since the reasoning underlying curve A expects little additional information from the personal and local knowledge of the voters at the top end. They are essentially extra, and redundant, observations. (Think of the law of large numbers: once one has an adequate sample from a large pool, there is little information to be gained by continued sampling.) If all voters had essentially the same information, this would be a plausible conjecture, as the voters at the top end would be essentially in surplus, and so largely unnecessary, observations. However, our societies are characterized by a rich plurality of values and experiences. Each individual is apt to bring new information to bear on the deliberation and decision.

It might appear that curve A can be defended if we reinterpret what voters are up to. Rather than each deciding on his personal (and local) knowledge of the actual deliberation to which he has been party, we might revert to the interpretation according to which each seeks to answer the same comprehensive question: “Overall, has L been quasi-vindicated in D?” (§3.4). Two objections confront such a proposal. First, although this interpretation may perhaps help alleviate the burden of each appealing to his own personal knowledge, we still have severe problems of different local knowledge: given the segmented nature of the information, even if each was seeking to answer the same question, each would be doing so on different informational bases. Secondly, and more importantly, if each is really trying to answer this global question—summing up not simply his experiences but an overall comprehensive judgment about the current state of D—then the reliability of each of these judgments is apt to be extremely low. Again, Hayekian-like considerations would lead us to doubt whether votes can make accurate
estimates of the global state of knowledge. Given this, the Condorcet jury theorem, which prima facie seems to endorse curves shaped like A, would actually support curves that move in the opposite direction of A. As is well known, if the average competency of each voter is less than 50 percent, Condorcet reasoning shows that reliability plunges downwards as votes are aggregated, just as dramatically as aggregation produces an upward sweep in reliability when average voter reliability is above 50 percent.

4.3 Curve B, positing a simple linear relation between degree of support and tendency to avoid oppression, does not make the mistake of A in overly discounting the information of those at the top end of the curve. Its error, which it shares with A, is that it places far too much weight on low levels of support for \( L \) as an indicator that \( L \) might be quasi-vindicated (and so not oppressive). For \( L \) to be quasi-vindicated, it must be the case that at the time of the vote, \( D \) has not revealed any decisive objection to \( L \) that has led one or more normatively competent deliberators to reject \( L \) on the grounds that it does not account for her reasons relating to \( I \). That a small number of deliberators have found no problem with \( L \) is slim evidence that there is no problem with \( L \). We would expect almost any proposal to find acceptance among a small set of citizens with shared sectarian beliefs and values; even when trying to deliberate according to the idealized deliberative norms, small groups are apt to quasi-vindicate a wide variety of proposals that would not be quasi-vindicated by full public deliberation. Thus, for example, that 25 percent of the voters find a proposal acceptable does not imply that a voting rule that allowed a proposal to be passed with 25 percent of the vote would select quasi-vindicated proposals 25 percent of the time. The chance of such a rule revealing whether \( L \) has been
quasi-vindicated is, of course, much less than that.

4.4 Curve C thus points to the right general type of relation. Low levels of support for $L$ say very little about the probability that $L$ has been quasi-vindicated; even 51 percent support is very uncertain evidence that no idealized (i.e., deliberatively sound) deliberator’s local and personal knowledge has revealed a problem with $L$. As we approach unanimity, our confidence that the result of the vote reliably tracks $D$ at the time of the vote drastically increases.

However, curve C too should be rejected insofar as it indicates that a unanimity rule is significantly more reliable than near-unanimity rules. Remember, $L$ is quasi-vindicated if no deliberators who have conformed to the idealized norms of deliberation have found a decisive objection to $L$. We also know, though, that in any group of voters, there will be some deliberators who fail to measure up to these norms, even if we suppose, as deliberative democrats typically do, that widespread deliberative competency is a precondition for deliberative democracy. Let us call $m$ the set of malfunctioning voters; these are voters who either manifestly fail to accord with deliberative norms, or who are cavalier in weighing reasons, or who approach voting in narrow self-interested or, as Habermas would say, “strategic” terms. Call $N$ the total set of voters. The unanimity rule, which requires that $N$ voters approve, thus requires that group $m$ endorses $L$. But the inputs of group $m$ are not apt to significantly add to the tendency of the voting rule to filter out proposals that are not quasi-vindicated, since $m$ voters add little if any reliable local and personal knowledge. A more reasonable hypothesis, depicted by Figure 2, is that, after achieving the support of $N-m$ voters (call this group $k$), the support of additional voters does not
significantly improve the tendency of the voting rule to meet our first desideratum.

5. The Second Desideratum: Avoiding Lost Opportunities for the Moral Life

5.1 Although adding the m voters (those lying between k and N) may marginally increase the voting rule’s ability to meet the first desideratum—to pass L only when L has been quasi-vindicated in D—the k rule does almost as good. Still a rule requiring N does not do worse, and may do a little better. Is there any reason to choose the k rule (N minus the group m) over the N rule?

It looks as if the N and k rules do about the same in meeting our first desideratum—the benefit of avoiding false positives: both do well at avoiding L being selected when it
has not been vindicated in $D$. They both avoid oppressive laws. However, avoiding false positives is not the only aim of a deliberative democratic voting rule. The drawback to $N$ is that it tends to too many \textit{false negatives}—it tends to eliminate proposals that have been quasi-vindicated because erroneous voters (especially, but not only, the $m$ group) reject them. We not only wish to avoid oppression: we want a society that lives according to shared, public, norms. As we increase the number of voters required to pass a measure, we run increased risk of missing out on shared norms verified by public reason: we miss opportunities for a shared moral life. “Malfunctioning” voters can veto a measure that normatively sound voters—those who conform to the proper norms of discourse, weighing reasons and so on—would endorse. Thus laws that truly have been quasi-vindicated by $D$ (our actual deliberations) might be often rejected by an $N$ rule. The attraction of a $k$ rule is that it cuts down on false negatives vis a vis $N$ while doing as well, or nearly as well, avoiding false positives.

However, even the $k$ rule will generate false negatives. Even voters who do not malfunction according to the rules of discourse or epistemic standards of deliberating on reasons can give an erroneous judgment. It would be wrong to identify the class on non-malfunctioning voters with those who have made no errors. One can have a fully justified, yet erroneous belief (see Gaus, 1996: 39). Consequently, while under a $k$ rule a vote of $k-1$ would defeat $L$, it is certainly still possible that $L$ was in fact quasi-vindicated in $D$, and so the rule generated a false negative. False negatives, then, are a cost of a deliberative democratic voting rule that must be considered when seeking to identify the optimal rule.
A decision on the optimal voting rule, then, must also take into account the costs of the voting rule in terms of its tendency to eliminate quasi-vindicated proposals. Figure 3 advances three estimates of the benefits of voting rules in avoiding false negatives as the percentage of votes required for passage rises.

Again we have three possible relations between the number of voters required to pass $L$ and the tendency of the voting system to satisfy the desideratum. Curve A depicts a simple linear relation. If only 1 person is required to pass a piece of legislation, it will, of course, be extremely easy for legislation to pass; consequently, the probability that a truly quasi-vindicated $L$ will erroneously be rejected by the voting system approaches zero. So a one person rule would highly satisfy our second desideratum: it would not miss opportunities for the moral life. At the other extreme, a unanimity rule—requiring
everyone’s assent before $L$ is passed—would do badly on this score: it would have a high tendency to eliminate quasi-vindicated rules and so undermine the possibility for a moral life. But while curve A gets the extremes right, its supposition of a simple linear relation is implausible. The probability that $L$ really is quasi-vindicated in cases where it is supported by only a small minority of voters seems negligible. It certainly would not be the case that, say, a rule that required only 20 percent to approve of $L$ would do much worse than the one person rule in satisfying this second desideratum: in neither case do we expect it is likely that all competent reasoners vindicate $L$ but the vast majority of voters do not realize this, and vote against it. The real possibility of false negatives does not arise until a rule requires for passage that $L$ is supported by a high percentage of votes cast. As we approach decision rules that require large extraordinary majorities, we would expect the benefits of avoiding false negatives to steeply decrease, up to the point in which they are extremely low with the $N$ rule, which rules out proposals vetoed by any one voter.

5.3 Curve B holds that the tendency to satisfy the second desideratum decreases rather quickly after majority rule. The shape of B is certainly more plausible that A, but again, its postulated relation seems dubious. Remember, a law is quasi-vindicated at a certain point in the deliberation only if no competent deliberator has uncovered decisive objections to it; and we are assuming that the large majority of citizens are competent. If so, the supposition that say, 40% of the citizens are very likely to disapprove of $L$ even when it is quasi-vindicated seems remote. Remember, if even a single competent deliberator has decisive objections to $L$, it is not quasi-vindicated, and so eliminating $L$
would not be to lose an opportunity to live a shared moral life. Only as the number of voters approaches the number of fully competent citizens do we run high risks of erroneously eliminating quasi-vindicated laws; even normatively competent deliberators can make mistakes. If the number required for passage is the entire set of normatively competitive reasoners, the risk of false negatives begins to loom large. If we require the approval of more than the set of competent reasoners (on our supposition, if we go beyond a $k$ rule and require support of some of the $m$ voters), the risk of false negatives quickly increases, since these voters are far more apt than the competent voters to register erroneous objections. Curve C thus seems to depict the correct general relation between number of voters and the tendency to satisfy the second desideratum: a slow decrease in the tendency from 1 to near $k$, and then a sudden decrease after that.

5.4 We can now combine desiderata-satisfying functions as in Figure 4:
Total benefit (adding the two curves together) occurs in the area X, which will be in the vicinity of the $k$ decision rule, and this requires far more voters to pass a law than the simply majority rule. Given our analysis, there is no good reason to think that area X includes the simple majority rule. There is, then a strong case that an optimal voting rule, in the sense of one that seeks to pass $L$ if and only if it has been quasi-vindicated in the deliberation up to the time of the vote, is something like the $k$ rule.

6. Non-Neutrality or Indeterminacy?

6.1 A traditional argument against supermajority rules is that they violate the condition of neutrality. K.O. May’s (1952) neutrality condition requires that a rule not be biased
against some alternatives; if \( x \) votes are enough to approve alternative A, \( x \) votes must be enough to approve of alternative B. Famously, supermajority rules often do not meet this condition. Assuming that \( j > N/2 + 1 \) (the \( j \) voting rule is a supermajority rule), we can easily see how it can take \( j \) votes to pass \( L \) but less than \( j \) voters to defeat \( L \). It is thus easier to defeat \( L \), and maintain the status quo, than it is to legislate a change. Thus supermajorities are said to be conservative rules biased toward the status quo. Given that deliberative democrats are typically not conservatives who accord a special status to the status quo, they might thus seem to have a principled objection to supermajority rules. This, no doubt, is one ground for their otherwise puzzling support for majority rule (§1).

We need to pause here and reflect on what we mean by a “bias” towards the status quo. In the context of deliberative democracy it would seem that the most plausible interpretation of the claim that a voting rule is biased in favor of the status quo is that by employing it we shall miss too many opportunities to live a shared moral life. By making it difficult to move away from the current set of laws, a supermajority rule might be too prone to block moves to adopt a law that genuinely satisfies the norms of public reason. That surely would be a great cost to insisting on a supermajority rule. However, we have already taken that cost into account: the second desideratum is all about minimizing such lost opportunities. Of course if the only thing we are concerned with is minimizing the missed opportunities for a shared moral life, a simple majority rule is attractive (though a less than simple majority rule such as plurality does even better). But we also must concern ourselves with minimizing the tendency for laws to be oppressive, and simple majority rules do not do well on that score. The power of the case for the
supermajority rule is that it unites both desiderata: we have seen that under plausible assumptions a $k$ rule optimally combines the dual goals of avoiding oppression and not missing opportunities for a shared moral life. Thus the case is not conservative insofar as this means privileging one sort of mistake (an erroneous change) over another (an erroneous failure to change).

6.2 Not all supermajority rules violate neutrality. Robert E. Goodin and Christian List (2006) point out that supermajority rules may be either asymmetrical or symmetrical. An asymmetrical rule violates neutrality requiring, for example, that more votes are needed to change the current law than to retain it: that $j$ is required to enact, but less than $j$ is required to keep the status quo. If supermajority rules are to be determinate—always giving a result—then they must be asymmetrical, and so non-neutral. This is not a problem for what we might call “liberal democracy.” As I have argued elsewhere, liberal political theory is built on a presumption in favor of non-interference and non-coercion (Gaus, 1996: 162-166). As J.S. Mill put it, “the onus of making a case always lies on the defenders of legal prohibitions” (1963a: 938): “in practical matters, the burden of proof is supposed to be with those who are against liberty” (1963b: 262). At the very foundations of such a view is a moral asymmetry between not-coercing and coercive laws. So non-neutral supermajority rules pose no real challenge to liberal conceptions of democracy: liberals can easily endorse determinate, non-neutral, supermajority rules. In contrast, such asymmetry is deeply troublesome for the deliberative democrat. It supposes precisely what deliberative democrats deny: that prior to deliberation some moral conclusions (say, on the status of coercion and individual rights) have been settled.
Recall that Habermas criticizes “Liberals [who] begin with the legal institutionalization of equal liberties, conceiving these as rights held by individualized subjects. In their view, human rights enjoy normative priority over democracy and the constitutional separation of powers has priority over the will of the democratic legislature” (1997: 44). Because actual deliberation is required to validate norms, it seems that only a political order based on actual public deliberation among the citizens can yield valid norms. If this is so, however, the democratic rules cannot be non-neutral, for that supposes that some moral position is more easily defended than others prior to deliberation.

Deliberative democrats, then, seem committed to both supermajority rules and neutrality. As Goodin and List argue, these two commitments can be honored; to do so the deliberative democrat must jettison the claim that voting procedures are determinate. That is, deliberative democrats must deny that a justified voting procedure will always be able to select between any pair of options (L, not-L). Supermajority rules can meet the neutrality condition if they accept indeterminacy, for example requiring that if the current status quo is justified it must have \( j \) (where \( j > N/2 + 1 \)) votes, and if it is to be changed there must be \( j \) votes in favor of change. If no option has \( j \) votes, the procedure is indeterminate: it does not resolve the political conflict. Neutrality is thus guaranteed by allowing that in cases where neither option receive \( j \) votes, there is a tie: the options are just as a good as each other from the perspective of the choice mechanism.

This appears the proper stance of deliberative democracy. Suppose that public deliberation has proceeded, and \( j \) citizens have refused to vote either for the status quo or for \( L \). It looks as if at this time, a sufficient number of citizens dissent from both options and so it is very unlikely indeed that either has been quasi-vindicated in public
deliberation. Remember, we are not conceiving of deliberation as seeking to uncover an already established truth, but constructing public reason through deliberation. That \( j \) citizens neither endorse \( L \) nor the status quo is good reason to suppose that neither option is quasi-validated. For if any of the competent deliberators in \( N-j \) are correct, neither proposal is in accord with public reason as it has thus far been constructed.

How can it be the case that neither \( L \) nor not-\( L \) has been vindicated by public reason? Doesn’t one or the other have to accord with public reason in such a case? Imagine that a country is involved in a war that \( N-j \) think is unjust, and is considering a peace proposal that \( N-j \) also think is unjust (the members of these two groups need not be identical). Surely it does not have to be that case that one or the other is just, or that one or the other must be in accord with public reason. Nor does it even have to be determinate that one or the other is better from the perspective of public reason: we cannot know until the deliberation has proceeded. When confronted with a choice of two options, neither of which recommends itself to public reason, it is by no means necessary that public reason itself demands that one must be chosen or that one is superior.

6.3 But, the deliberative democrat may respond, we must do something, and whichever option receives the most votes is, at least very slightly, more likely to be quasi-vindicated than that which receives less votes. Majority rule might be the worst method—except for the alternatives. It is certainly true that, given Figure 4, majority rules fare better than below 50% rules in terms of the combined benefits. So, given the demands of action, it may seem, the best course is a majority vote. There are, though, two reasons for a
deliberative democrat not to embrace this reply.

First, although according to Figure 4 the combined benefits of a proposal with 51 percent of the vote exceed that with 49 percent, not only are the combined benefits much lower than for supermajority rules, but the risk of oppression is high—most of the benefits are in terms of avoiding missed opportunities for the moral life, and relatively little comes from the tendency to avoid oppression. To urge a voting rule that, in the interest of political determinacy leaves a polity open to oppression seems ill-advised. Given the deep moral commitments of deliberative democracy, this compromise with political reality comes at a high cost indeed.

Second, as Habermas (1996: 307) stresses, there is an internal connection between public deliberation and procedures: the nature of the procedures affect the process of public deliberation. Majoritarian rules, which regularly validate proposals that are unlikely to have been quasi-validated by public reason are to that extent are corrupting: they induce a politics in which “winning the game” does not support the search for quasi-vindicated proposals.

7. Conclusion:
Deliberative Democracy, Political Action and the Burdens of Justification

7.1 The deliberative democrat appears caught in a dilemma. On the one hand, there is a strong case for super-majority voting rules, and this case is not objectionably conservative. But given this case, if the deliberative democrat constructs a voting system that adequately tracks public reason, the outcome will often be indeterminate. And this seems to mean that just when we most need a legitimate way to resolve political
differences—when we are really divided about the right common course and so are in what Jeremy Waldron (1999: 102) calls “the circumstances of politics”—deliberative democracy is apt to be of little help. Neither L nor not-L will be endorsed by its rules. If we decide, we must decide without relying on the ideal of deliberative democracy. Confronted by this unattractive position, we can understand deliberative democrats embracing majority rule as way “to cut off deliberation” and “make a decision.” Thus our initial puzzle (§1) has an answer: embracing majority rule seeks to make deliberative democracy politically relevant in the real world of politics, where intractable disagreement is the norm. But this raises the other horn of the dilemma. The cost of political relevance is often embracing oppression: simple majority rule is likely to endorse laws that have been decisively defeated in our actual deliberations.

7.2 As I have argued (Gaus, 2007a) elsewhere in a different context and a different way, deliberative democracy, as a theory of political legitimacy of laws and policies, wilts under what we might call the “burdens of justification.” Under modern conditions of reasonable pluralism, the deliberative democrat seeks to collectively justify laws and policies through direct appeal to deliberation and the actual exercise of public reason: what all deliberators rationally endorse. But such public reason is an extremely demanding standard—it does not allow us to dismiss small, normatively competent, actual minorities as unimportant in the face of majoritarian demands. Because of this, if we are to stay true to the deliberative democrat’s ideal of public reason in its application to actual politics, we must employ demanding voting rules. But as soon as we do so, we see that shouldering these heavy burdens of justification often renders the theory indeterminate,
and so of far less practical relevance than it aspires. Whether other accounts of public reason can better bear the burdens of justification is not my concern here (see Gaus, 2007). And I have only considered one version of deliberative democracy—perhaps other versions do not have these difficulties. I think it is clear, though, that an important version of deliberative democracy—deliberative constructivism—has great difficulty both shouldering its justificatory burdens and making itself relevant to the urgencies of political life.

Philosophy
University of Arizona
Notes

1 For simplicity sake I shall refer simply to moral norms, though “moral norms, political laws and policies” is more accurate. Note that I am simplifying: Habermas has distinct accounts of norms and laws.

2 What I describe here is akin to McMahon’s preferred notion of “collective reasoning.” See McMahon (2001).

3 Rawls’s notion of an overlapping consensus seems to employ convergence justification. Under what he calls “full” justification each suitably idealized citizen affirms that the free-standing political conception (which has been given a “pro tanto” justification via consensus) can be endorsed by their differing comprehensive doctrines, and so become the object of an overlapping consensus. See Rawls 1996 (386-387) and see further Gaus, 2007.

4 For an argument that Rawls’s and Habermas’s procedures are very similar, see McMahon (2002).

5 For an analysis of extending this finding to amalgamating judgments, and its problems for deliberative democracy, see Pettit (2004). For doubts about Pettit’s formal argument, see Chapman (2002).

6 Note that, like Buchanan and Tullock (1962: 68ff), I postulate continuous functions, starting with a decision rule that requires just one person all the way up to N. Mueller (2003: 76-78) argues that decision rules that allow passage of L with less than a majority (say 20%) would allow passage of both L and not-L, and there would be great costs to such rules. These costs are high from 1 to 50%, and are zero at >50% and afterwards. Thus Mueller argues for discontinuous functions (and so curves); he claims that, even accepting Buchanan and Tullock’s understanding of internal and external costs, majority rule is optimal (see also Gaus, 2007b: ch. 6). However, if we are worried by the possibility of chaotic rules below simple majority rule, we can construct functions that concern only size of the majority (i.e., curves that start with greater than 50% and go to N)—basically
curves that only concern the second half of the curves I discuss here. It still can be shown that supermajority rules or unanimity rules are optimal under a variety of conditions (Dougherty and Edwards, 2004). To keep the analysis simple, and to show how majority rule fares on our desiderata vis à vis less-than-majority rules, I employ continuous curves here. The costs of contradictory laws, though of course important, is not crucial to our main concerns. If the reader is worried by contradictory laws for decision rules of less that 50%, assume the curves all begin at majority rule. It should be pointed out, though, that when we have choices over three options, rules that require less than 50% need not lead to chaotic results (e.g., plurality).

7 A related argument has recently been advanced by Robert Goodin (2002). Goodin shows that, if we adopt these Bayesian attitudes towards our votes and those of others, for a remarkably wide range of degrees of confidence in our own judgments, a person ought to believe what the majority believes. If it is rational to believe what the majority believes, then we can expect little additional reliability in the outcome of the vote regarding L by requiring extraordinary majorities (rules that require more voter support of more than N/2), for the beliefs of the minority, if rational, would reflect the majority's. My comments in the text concerning Condorcet's jury theorem apply to Goodin's analysis as well—differing local knowledge is even more of a problem for Goodin's analysis.

8 What is interesting is that it shows us what we are all committed to believing (the truth) but we all do not yet do it.

9 Unless, of course, the vast majority of voters are incompetent—but we have supposed (§4.1) otherwise. We also must suppose they are seeking to sincerely register their judgment, not voting expressively—again, this supposition is basic to liberal democracy. On voting expressively, Brennan and Lomasky (1993: ch. 3).

10 This supposition, of course, is at the heart of Buchanan and Tullock's (1962) classic analysis of voting rules.
References


Rousseau.” *Australasian Journal of Philosophy*, vol. 75 (June): 141-162.


