[Christian List](#)

When to defer to supermajority testimony – and when not

Working paper

When to defer to supermajority testimony – and when not

Christian List[1]

20 October 2006

Pettit (2006) argues that deferring to majority testimony is not generally rational: it may lead to inconsistent beliefs. He suggests that "another ... approach will do better": deferring to supermajority testimony. But this approach may also lead to inconsistencies. Here I identify the conditions under which deference to supermajority testimony ensures consistency, and those under which it does not. I also introduce the new concept of 'consistency of degree *k*', which is weaker than full consistency by ruling out only 'blatant' inconsistencies in an agent's beliefs while permitting less blatant ones, and show that, while super-majoritarian deference often fails to ensure full consistency, it is a route to consistency in this weaker sense.

1. The problem

Philip Pettit has recently argued that although it is sometimes rational to defer to majority testimony on perceptual matters – say, whether a car went through the traffic lights on the red – this is not generally the case with matters more deeply embedded in one's web of belief – say, whether abortion is wrong (Pettit 2006). A key problem is that deference to majority testimony may lead to inconsistent beliefs. For example, suppose one agent believes that p and q are both true, a second believes that p is true and q is false, and a third believes that p is false and q is true. Then p, q and *not*-(p&q) are each believed by a majority, and thus deference to these majorities would lead to inconsistent beliefs.

Pettit suggests that "[t]here is another … approach that will do better … This is not to allow just any majoritarian challenge to reverse a belief but to allow only a certain sort of supermajoritarian challenge to do so" (Pettit 2006, p. 184). As an illustration, he observes that, assuming consistent individual beliefs, there can never be supermajorities of 70% believing each of p, q and *not*-(p&q) to be true. If there were such supermajorities, the inconsistency would have to show up in the beliefs of at least one individual agent.

It is easy to see, however, that a 70% supermajority requirement is insufficient to prevent an inconsistency between a larger number of propositions. In a group of four agents, for

[1] I am grateful to Franz Dietrich, Philip Pettit and Wlodek Rabinowicz for stimulating discussions and advice. Address: C. List, Dept. of Goverment, London School of Economics, London WC2A 2AE, U.K.

example, there can easily be 75% supermajorities for each of *p*, *q*, *r* and *not*-(*p*&*q*&*r*), even when each agent holds individually consistent beliefs, such as when the first agent accepts all but the first of these four propositions, the second accepts all but the second, and so on, as shown in table 1.

Table 1: A supermajoritarian inconsistency

	p	*q*	*r*	*not*-(*p*&*q*&*r*)
Agent 1	false	true	true	true
Agent 2	true	false	true	true
Agent 3	true	true	false	true
Agent 4	true	true	true	false
Supermajority of 75%	true	true	true	true

When does deference to supermajority testimony guarantee consistency, and when not? In this short paper, I answer this question in full generality. Drawing on recent results (Dietrich and List 2006, generalizing List 2001) from the theory of judgment aggregation (List and Pettit 2002), I state necessary and sufficient conditions not only for achieving consistency through supermajoritarian deference but also for achieving something less than full consistency: namely what I call consistency of degree *k*, or in short *k*-consistency. This is the requirement that inconsistencies in an agent's beliefs, if there are any, should not be too blatant, where *k* is an integer number capturing the degree of 'blatancy' of the inconsistencies ruled out, in a sense to be made precise.

My argument generalizes, but also qualifies, the observation that deference to supermajority testimony can sometimes be rational.

2. Minimal inconsistent sets and supermajority testimony

What are the simplest inconsistencies that can arise in an agent's belief set?[2] Call a set of propositions minimal inconsistent if it is inconsistent but all its proper subsets – obtained

[2] I here think of a belief set as a set of propositions accepted by an agent. Propositions are represented by sentences in a suitable logic, such as standard propositional or predicate logic. Generally, any logic satisfying self-entailment, monotonicity, completability and compactness, as defined in Dietrich (2006), is suitable. Apart from standard propositional and predicate logics, many modal, conditional and deontic logics are examples of logics to which the present analysis applies.

by removing one or more propositions from the set – are consistent. For example, the sets $\{p, q, not\text{-}(p\&q)\}$ and $\{p, q, r, not\text{-}(p\&q\&r)\}$ are each minimal inconsistent: they become consistent as soon as one or more propositions are removed. By contrast, the set $\{p, p\&q, not\text{-}p\}$, although inconsistent, is not minimal inconsistent: even if one of p or $p\&q$ is removed from it, it remains inconsistent. Any inconsistent set of propositions has at least one, and possibly many, minimal inconsistent subsets. It follows that any agent with inconsistent beliefs has at least one minimal inconsistent set of propositions among his or her beliefs. Conversely, any agent whose beliefs include no minimal inconsistent set of propositions is consistent throughout.

Under what conditions can deference to supermajority testimony lead an agent to believe a minimal inconsistent set of propositions?

> **Fact 1.** It is possible for a minimal inconsistent set of k propositions to be each supported by a supermajority among agents with individually consistent beliefs if and only if the supermajority size is less than or equal to ${}^{k-1}/_{k}$.

To prove this fact, consider any minimal inconsistent set of k propositions. Call them p_1, p_2, …, p_k. I first show that supermajorities of size ${}^{k-1}/_{k}$ (and by implication of smaller sizes) among agents with individually consistent beliefs can support each of these propositions. Take any set of agents divisible into k subsets of equal size. Suppose the agents in the first subset believe all of the k propositions except p_1, the agents in the second subset believe all except p_2, and so on. As every proper subset among p_1, p_2, …, p_k is consistent – in particular, every subset obtained by dropping precisely one of these propositions – any such agent holds consistent beliefs. But now each of p_1, p_2, …, p_k – that is, each proposition in a minimal inconsistent set of size k – is supported by a supermajority of size ${}^{k-1}/_{k}$.

Conversely, I show that supermajorities of size greater than ${}^{k-1}/_{k}$ among agents with individually consistent beliefs can never support all of p_1, p_2, …, p_k. Assume, for a contradiction, that there are k such supermajorities. For any two of these supermajorities, even if maximally distinct, the overlap must exceed ${}^{k-1}/_{k} - (1 - {}^{k-1}/_{k}) = {}^{k-2}/_{k}$. For any three, the overlap must exceed ${}^{k-2}/_{k} - (1 - {}^{k-1}/_{k}) = {}^{k-3}/_{k}$. Continuing, for all k supermajorities, the overlap must exceed ${}^{k\text{-}k}/_k = 0$. So the supermajorities must have a non-empty overlap, implying that at least one agent lies in their intersection. But this would mean that this agent holds inconsistent beliefs, contradicting the assumption that agents have individually consistent beliefs. This completes the proof.

3. Ensuring consistency

What, in light of fact 1, could a rational policy of deference to supermajority testimony look like? In particular, what could such a policy look like from the perspective of preventing inconsistency in our beliefs?

Suppose the aim is to arrive at fully consistent beliefs. Consider the entire set of propositions on which beliefs are to be formed or revised. This set could, for example, contain all those propositions that occur somewhere in an agent's web of belief. Let k be the size of a largest minimal inconsistent set constructible from these propositions and their negations. To illustrate, if the only propositions the agent forms or revises beliefs on are p, *if p then q*, and q, then the largest minimal inconsistent set constructible from these propositions and their negations would be {p, *if p then q*, *not-q*}, and thus k would be 3. If the underlying set of propositions is larger and more complex, of course, k can be significantly larger.

Now fact 1 implies immediately that the policy of adopting all and only those beliefs held by a supermajority of size greater than ${}^{k-1}/_k$ can never lead to an inconsistency. If it did, the resulting inconsistent belief set would have to include a minimal inconsistent set of propositions; but that set would contain at most k propositions (as k is the size of the largest minimal inconsistent set constructible from the given propositions and their negations), and fact 1 implies that no such set can be supported by supermajorities of size greater than ${}^{k-1}/_k$ among agents with individually consistent beliefs. Thus the following holds (Dietrich and List 2006, generalizing List 2001).

> **Fact 2.** Let k be the size of a largest minimal inconsistent set of propositions constructible from the propositions on which beliefs are to be formed or revised and their negations. Then the set of propositions supported by a supermajority of size greater than ${}^{k-1}/_k$ among agents with individually consistent beliefs is consistent.

However, for any supermajority size below unanimity, the set of propositions supported by supermajorities of that size is not guaranteed to be deductively closed: the propositions receiving the requisite supermajority support may entail other propositions that fail to receive such support (for general results, see Dietrich and List 2006). But this means that deferring to supermajority testimony on propositions on which there is the requisite supermajority agreement while holding on to one's prior beliefs on all other propositions may not be a rational policy: those prior beliefs may conflict with the supermajority beliefs elsewhere. Unless one is willing to raise the supermajority threshold to unanimity, a rational policy of supermajoritarian deference would therefore require not only deferring to supermajority testimony when such testimony is above the relevant threshold, but also revising other beliefs in light of it.

Moreover, as the set of propositions on which beliefs are to be formed or revised increases in size and complexity, the value of k – the size of the largest minimal inconsistent set constructible from these propositions and their negations – typically increases as well, and thus the supermajority threshold required to ensure consistency approaches unanimity.

4. Avoiding blatant inconsistencies

Achieving full consistency in one's beliefs may not always be feasible. Indeed, it is perhaps unrealistic to expect the beliefs of a normal human agent to be consistent. On the other hand, we do expect those beliefs to be free at least from the most blatant inconsistencies. When is an inconsistency blatant? An agent who believes a single proposition that is self-contradictory, such as *p*&(*not-p*), is clearly blatantly inconsistent. An agent who simultaneously believes a proposition and its negation, such as *p* and also *not-p*, is also fairly blatantly inconsistent, even if each of *p* and *not-p* is not contradictory by itself. An agent who believes three propositions which are in contradiction, such as *p*, *if p then q*, and *not-q*, is still rather blatantly inconsistent, but not as much as one who believes a self-contradictory proposition or a proposition-negation pair. An agent with inconsistent beliefs across five propositions, such as four logically independent conjuncts and the negation of their conjunction, is still inconsistent, but intuitively less so than any one of the earlier agents.

Now suppose that, although my large set of beliefs is inconsistent in its entirety, it turns out that every combination of 1588 or fewer propositions among my beliefs is consistent and the smallest set over which I hold inconsistent beliefs contains 1589 propositions. Should my beliefs still be described as blatantly inconsistent? Intuitively, the inconsistency here is much less blatant than in any of the earlier cases.

My proposal is to measure the blatancy of an agent's inconsistency by the size of the smallest minimal inconsistent set of propositions believed by the agent. The smaller this size, the more blatant the agent's inconsistency. To be sure, this is a rather simple measure, but I illustrate its usefulness in a moment. In the examples just given, the values of the measure are 1, 2, 3, 5 and 1589, respectively, capturing the intuitive ranking of how blatant the inconsistencies in question are.

Just as the blatancy of an agent's inconsistency can be measured by the size of the smallest minimal inconsistent set of propositions among the agent's beliefs, so the degree of consistency of the agent can be measured in a closely related way. Call an agent whose belief set is free from any minimal inconsistent subset of k or fewer propositions consistent of degree k, or in short k-consistent. For example, an agent who believes no self-contradictory proposition is 1-consistent. An agent who, in addition, believes no proposition-negation pair (and no inconsistent set of similar complexity) is 2-consistent. One who further does not believe any inconsistent set of the form {*p*, *if p then q*, *not-q*} is 3-consistent. And so on. In the contrived example of my less-than-fully-consistent beliefs, I would be 1588-consistent. Full consistency, finally, is the special case of k-consistency for an infinite value of k.

Perhaps the best a human agent can ever hope to achieve is k-consistency for a reasonably large value of k. What could a policy of deference to supermajority testimony look like if the aim were to achieve k-consistency for some finite value of k? The following corollary of fact 1 answers this question.

Fact 3. For any value of k, the set of propositions supported by a supermajority of size greater than ${}^{k-1}/_{k}$ among agents with individually consistent (or even merely k-consistent) beliefs is k-consistent.

Of course, if the underlying set of propositions on which beliefs are to be formed or revised has no minimal inconsistent subsets of size greater than *k*, then *k*-consistency implies full consistency. In this case, fact 3 reduces to fact 2. Otherwise, fact 3 is more general.

Fact 3 suggests that, while full consistency may often be hard to achieve through deference to supermajority testimony short of unanimity, supermajoritarian deference may nonetheless be a good route to *k*-consistency for a suitable value of *k*. And this remains true even if the agents constituting the supermajorities in question are themselves merely *k*-consistent. Thus, for any value of *k*, deference to supermajorities of size greater than $^{k-1}/_{k}$ preserves *k*-consistency.

In summary, the larger the supermajority threshold we require for the acceptance of a belief, the less blatant the inconsistencies we are liable to run into.

5. Coherence and correspondence

For a sufficiently high threshold, deference to supermajority testimony may yield consistent beliefs; and for lower thresholds, it may yield beliefs that are not too blatantly inconsistent; in both cases, other beliefs, on which there is no sufficient supermajority agreement, may need to be revised accordingly.

Does this make supermajoritarian deference rational? My focus has been on 'coherence' considerations: supermajority testimony is less prone to incoherence than majority testimony. A different set of considerations are 'correspondence' ones. Is supermajority testimony likely to indicate the truth on matters of fact?

Whenever agents fulfil the assumptions of Condorcet's jury theorem – that is, they each have an independent, better-than-random chance of making a correct judgment on the matter in question (Grofman, Owen and Feld 1983) – the answer is essentially positive, just as in the simple majority case discussed by Pettit (2006). But agents need not fulfil these assumptions on matters deeply embedded in their webs of belief (see also Bovens and Rabinowicz 2006). As a simple consequence of the laws of probability, for example, an agent cannot generally be as reliable at detecting the truth of a conjunction as he or she is at detecting the truth of each conjunct. Thus, even if agents were very reliable in their judgments on simple matters – whether perceptual or not – they would have to be less reliable on some other, composite or derivative ones. Just as a majority can be wrong on such matters, so a supermajority can be deeply mistaken as well – and even more confidently so.

An analysis of the truth-indicating reliability of supermajority testimony is beyond the scope of this short paper (for some relevant results, see Feddersen and Pesendorfer 1998 and List 2004). But it seems wise to exercise caution in deferring to such testimony. Before you move to a new ground because a supermajority of other agents stands there, make sure that ground is firm enough to support you too.

References

Bovens, L. and W. Rabinowicz. 2006. Democratic Answers to Complex Questions – An Epistemic Perspective. *Synthese* 150: 131-53.

Dietrich, F. 2006. A generalised model of judgment aggregation. *Social Choice and Welfare*, forthcoming.

Dietrich, F. and C. List. 2006. Judgment Aggregation by Quota Rules: Majority Voting Generalized. *Journal of Theoretical Politics*, forthcoming.

Fedderson, T. and W. Pesendorfer. 1998. Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting. *American Political Science Review* 92: 23-35.

Grofman, B. G., G. Owen and S. L. Feld. 1983. Thirteen theorems in search of the truth. *Theory and Decision* 15: 261-78.

List, C. 2001. Mission Impossible? The Problem of Democratic Aggregation in the Face of Arrow's Theorem. DPhil-thesis, University of Oxford.

List, C. 2004. On the Significance of the Absolute Margin. *British Journal for the Philosophy of Science* 55: 521-44.

List, C. and P. Pettit. 2002. Aggregating Sets of Judgments: An Impossibility Result. *Economics and Philosophy* 18: 89-110.

Pettit, P. 2006. When to defer to majority testimony – and when not. *Analysis* 66: 179-87.