The Limits of Epistemic Democracy

Martin van Hees*
University of Groningen
Department of Philosophy
Oude Boteringestraat 52
9712 GL Groningen Netherlands
Martin.van.Hees@rug.nl

October 27, 2005

Abstract

The so-called doctrinal paradox reveals that a jury that decides by majority on the truth of a set of propositions, may come to a conclusion that is at odds with a legal doctrine to which they all subscribe. The doctrinal paradox, and its subesequent generalization by List and Pettit (2003), reveal the logical difficulties of epistemic democracy. This paper presents several generalizations of the paradox that are formulated with the use of many-valued logic. The results show that allowing the individual or the collective judgements to be formulated in terms of degrees of beliefs does not ensure the possibility of collective epistemic decision making.

1 Introduction

In Plato's dialogue *Euthyphro*, Socrates cross-examines Euthyphro, who claims to know what it is that makes an act pious or impious. In his familiar scrutinizing way, Socrates shows that the certainty of his opponent is unwarranted - things are not as clear as Euthyphro supposes them to be. One of the definitions of piety that Socrates examines - and rejects - is that it consists of doing what the gods agree with. Socrates notes that the gods disagree about many things. As he puts it, the 'same things will be hated by the gods and loved by them'. Consequently, according to Socrates, defining a pious act as an act with which the gods agree and an impious act as one

*This paper was presented at the workshop on Judgement Aggregation and the Discursive Dilemma at the University of Konstanz, June 18-19, 2004. I am very grateful to the participants for their comments. Furthermore, I would like to thank Jelle Gerbrandy, Marc Pauly, and Allard Tamminga for their helpful comments.

with which they disagree entails that some actions can be said to be both pious and impious - they are pious because (some of) the gods agree with them and impious because (some of) the gods also disagree with them.

Socrates' argument suggests a particular view of the nature of the convictions that we can ascribe to a group of agents. On the one hand, a collective judgement must somehow be based on the judgements made by the agents constituting the collectivity. A decision about what 'the' gods think should be derived from the opinions of the various individual gods. If there is unanimity among the individual gods - and such an assumption is made in the rest of the dialogue - we can indeed talk about the judgement or the belief of 'the' gods. On the other hand, if they disagree then this disagreement will be reflected on the collective level. An inconsistency arises: an action is judged to be both pious and impious.

When we assign an epistemic judgement to a collective - e.g. when we speak about the opinion of a court or the judgement of a cabinet - we do indeed often assume that such a judgement should somehow be a function of the judgements made by the agents (court members, ministers) constituting the collectivity (the court, the cabinet). However, we do not share Socrates' pessimism that in case of disagreement the resulting judgement necessarily has to be inconsistent. Although we are familiar with the notion of a hung jury, just as we are familiar with cabinets that are unable to take a stance, in many cases it seems that we can assign beliefs or judgements to a collective without there being unanimity among its members. After all, such corporate bodies often use some aggregation procedure through which individual judgements are 'translated' into a collective judgement. Why should the same not be true of the gods on Mount Olympus?

This paper studies such aggregation procedures, and the main results show that Socrates' fear for inconsistency on the collective level is more justified than we might think. Aggregation procedures that yield collective decisions have, of course, been the object of study within social choice theory. However, social choice theorists study aggregation procedures that take individual *preferences* as their input, while the procedures that we refer to here take individual *beliefs* or *opinions* as their input. It is possible to formally connect these two different approaches. One way of doing so was proposed and explored by List and Pettit (2004) and consists of taking the propositions not as primitives but as describing the preferences of individuals. Thus the beliefs or judgements that individuals submit are beliefs or judgements about their own preferences, and the social judgement would be interpreted as a judgement about the social preferences. Under such an approach, one may be able to deduce the Arrovian results in the propositional framework used here.

Despite the challenging prospect of obtaining such a general framework, there are also reasons to focus on the aggregation of individual judgements separately from that of individual preferences, reasons which derive from the view that the distinction between *epistemic* decision making as examined here and *preferential* decision making as studied within social choice theory is not without importance. First of all, the input information has a different logical structure in the two settings. The preferences studied in social choice theory are described by binary relations over social states, that is, over alternatives that completely specify states of affairs and which are mutually exclusive. Beliefs or opinions, however, are represented by propositions and by the truth-values that the believers assign to these propositions. If we interpret these propositions as expressing beliefs about the state of the world (rather than about preferences) we see that most propositions describe only *a part* of a social state. Thus all sorts of logical relationships can exist between them. Moreover, and more substantively, it has been argued that the formal analysis of democratic decision making has focused too much on a preferential approach, at the expense of ignoring the fact that the exchange of information or the specification of reasons for one's preferences plays an important role in democratic decision making. Because of this neglect of the epistemic aspects of collective decision making, the aggregation problems highlighted by social choice theory have been taken to be less relevant to an epistemic conception of democracy (Cohen 1986, Coleman and Ferejohn 1986).

Interest in the aggregation problems of epistemic decision making was sparked off by the doctrinal paradox (Kornhauser and Sager 1993) and particularly by its subsequent generalization by List and Pettit (2002). The doctrinal paradox reveals that a jury that decides by majority on the truth of a set of propositions may come to a conclusion that is at odds with a legal doctrine to which they all subscribe (hence 'doctrinal' paradox). Take the illustration from Kornhauser and Sager (1993): a court has to decide whether a defendant is liable under a charge of breach of contract, and the three judges have to answer three questions: Was the contract valid (p)? Is there a breach (q)? And is the defendant liable (r)? The legal doctrine that constrains their decision making is that the defendant is liable if, and only if, the contract was valid and if there was indeed a breach ($r \leftrightarrow (p \wedge q)$). The members of the court make the following judgements (where for instance 'p' stands for 'accepts p' (or 'believes p to be true') and '$\neg p$' for 'rejects p' ('believes p to be false')):

1: $p, \neg q, \neg r$
2: $\neg p, q, \neg r$
3: p, q, r

Although each of the court members, in his or her opinion, respects the legal doctrine (that is, $r \leftrightarrow (p \wedge q)$), if they take a majority vote on each proposition the resulting collective decision will be p, q, and $\neg r$, which is at odds with the doctrine.

List and Pettit (2002) have shown that the majority rule is but one member of a larger class of aggregation procedures that fails to ensure that a collection of sets of consistent individual judgements can always be aggregated into a consistent set of collective judgements. Pauly and Van Hees (2004) have further generalized the doctrinal paradox by showing that there is an even larger set of aggregation procedures for which this is true. The generalization is not only based on a weakening of some of the properties that List and Pettit impose on aggregation procedures; it also abandons the assumption that individual and collective beliefs necessarily have a binary nature ('true' or 'false').[1] That is, they allow for the possibility that the individuals as well as the collective can express degrees of belief that lie between 'true' and 'false'.

When deriving their results Pauly and Van Hees made use of the many-valued logic of propositions formulated by Post (1921). An obvious advantage of using Post's logical system is that it is functionally complete, which means that any logical operation can be defined in terms of the Post connectives. However, the price paid for obtaining this functional completeness is the rather counter-intuitive nature of one of its connectives, viz. the negation operator. In order to establish the exact implications of the various results established by Pauly and Van Hees, one could define a new negation operator in terms of the Post-connectives, and then examine the extent to which similar results can be derived with that new operator. Rather than pursuing such a route, which remains within the confines of the Post system, this paper explores the possibility of deriving general results about the possibility and impossibility of epistemic decision making in a many-valued logic that is both weaker (since it is not functionally complete) and intuitively more appealing than the Post system (since it makes use of a more appealing negation operator).

The structure of this paper is as follows. Section 2 presents the formal framework. Section 3 shows that demanding that aggregation functions be 'systematic' - a condition introduced by List and Pettit - narrows the set of aggregation procedures down to a class of dictatorial mechanisms. Since systematicity is a very strong demand, Section 4 examines what happens if we weaken the demand to that of 'independence of irrelevant alternatives'. We show that with a strengthening of one other condition we again end up with the existence of a dictator. The concluding section discusses the implications for the analysis of epistemic decision making.

[1] Alternative recent generalizations of the doctrinal paradox for the two-valued case are given by Dietrich (2004) and Nehring and Puppe (2004). Gärdenfors (2004) presents a generalization that is also based on a two-valued logic but which allows individuals to abstain from expressing their beliefs.

2 Definitions and Notation

Let $N = \{1, 2, \ldots, n\}$ denote a finite set of individual decision makers ($|N| \geq 1$), and Φ_0 a (finite or infinite) set of atomic propositions p, q, etc. The set of all propositions Φ is obtained by closing Φ_0 under the standard propositional connectives of conjunction ($\wedge$) and negation ($\neg$). Thus $\Phi_0 \subseteq \Phi$ and for all formulas $\phi, \psi \in \Phi$: $\neg\phi \in \Phi$ and $(\phi \wedge \psi) \in \Phi$.

The set of propositions about which the decision makers actually have to make a decision is denoted by Ψ and is some non-empty subset of Φ. $\Psi_0 = \Psi \cap \Phi_0$ is the set of atomic propositions in Ψ. A *literal* ϕ is an atomic proposition or the negation of one. Literals are said to be different (designated by '$\not\approx$') if, and only if, they do not involve the same atomic proposition.

Let $T = \{0, 1, \ldots, t-1\}$ be the set of truth values, where we assume that $|T| = t > 1$. Intuitively, we may think of $t-1$ as 'true', 'agree' or 'accept', and of 0 as 'false', 'disagree' or 'reject'. The values between 0 and $t-1$ (if there are any) then represent degrees of truthfulness (agreement or acceptance).

A *global valuation* is any function $v^* : \Phi \to T$ satisfying:

1. $v^*(\neg\phi) = t - 1 - v^*(\phi)$;
2. $v^*(\phi \wedge \psi) = \min\{v^*(\phi), v^*(\psi)\}$.

We let V^* be the set of all global valuations. Note that for the particular case of 2-valued logic with $T = \{0, 1\}$, the connectives defined do indeed correspond to standard negation and conjunction.

We assume that individuals only express judgements about the propositions belonging to the agenda. However, such judgements will sometimes implicitly refer to judgements that are not part of the agenda. Suppose for instance that the agenda is $\{p \wedge q, \neg p \wedge r\}$ and assume that some i believes the proposition $p \wedge q$ to be true (assigns it a truth value of $t-1$). It is then reasonable to assume that he judges $\neg p \wedge r$ to be false: the truth of $p \wedge q$ 'implicitly' entails the truth of p which in turn entails the rejection of $\neg p \wedge r$.

To accomodate for such implicit logical relations between propositions belonging to the agenda, we assume that the judgements that individuals submit (and the collective judgements that are derived from it) should be compatible with some global valuation. That is, we define a *valuation* as any function $v : \Psi \to T$ for which there is some $v^* \in V^*$ such that the restriction of v^* to Ψ equals v. We let V denote the set of all such valuations. An *aggregation function* $A : V^N \to V$ returns for every profile of valuations $(v_1, \ldots, v_n)$ an aggregated valuation $A(v_1, \ldots, v_n)$. Note that we assume not only that both the individual and the social judgements are consistent, but also that they are complete: a truth value is assigned to each element of the agenda. Furthermore, note that we assume that aggregation functions have *universal domain*, i.e., they are defined on all possible valuation profiles.

Before we proceed with the analysis, a few remarks about the logic we use are in order. As stated in the introduction, the system of many-valued logic used here differs from the Post system used in Pauly and Van Hees (2004). In particular, it differs with respect to the rule for the negation operator - our definition is intuitively much more appealing and is in fact the standard one in many-valued logic.[2]

Our logic is not functionally complete but can be made so by introducing additional operators. It should be noted, however, that, provided the principle of truth-functionality is not abandoned (that is, the principle that the truth value of any complex proposition is a function of the values of its component parts), such an enrichment of the logic will not affect the main results derived here. After all, our results show the conditions under which the truth values that society assigns to propositions are completely determined by a particular individual (the dictator). Since the dictator thus also determines the truth values that society assigns to atomic propositions, the truth value of any proposition that involves operators other than the ones of conjunction and negation will also coincide with the truth value assigned to it by the dictator.

Finally, we note that the assumption of truth-functionality entails that the degrees of beliefs that are expressed by the truth values should *not* be interpreted as reflecting degrees of uncertainty. Indeed, in such a context a probabilistic approach, in which truth-functionality does not hold, would be more appropriate. Stated differently, our logic is much closer to a 'fuzzy' interpretation of degrees of belief than to a probabilistic one.[3]

3 Systematicity

In order to describe the List-Pettit and the Pauly-Van Hees generalization of the doctrinal paradox for a two-valued logic, we present several conditions. The first one is about the agenda.

Condition 3.1 (Minimal Agenda Richness) *The agenda Ψ contains at least two distinct atomic propositions p, q as well as $p \wedge q$ and $\neg(p \wedge q)$.*

The second condition states that each individual has the same influence on the final decision.

[2]See Rescher (1969). As indicated in the introduction, the functional completeness of the Post system implies that one can define the negation operator that we use in terms of the Post operators. However, the resulting definition (see Urquhart, 2001, p. 265, for a way to derive it) is so complex that a Post system of logic that incorporates our negation operator can only be applied to the analysis of epistemic decision making if very strong agenda assumptions are made.

[3]See, however, Urquhart (2001) for a critical discussion of the usefulness of either fuzzy or many-valued logic for representing degrees of belief.

Condition 3.2 (Anonymity) *For any permutation $f : N \to N$, any valuation profile $(v_1, \ldots, v_n) \in V^n$ and any proposition $\phi \in \Psi$, $A(v_1, \ldots, v_n)(\phi) = A(v_{f(1)}, \ldots, v_{f(n)})(\phi)$.*

The next condition, systematicity, states that decision making about the propositions is done in a uniform way. Define a *decision method* as a scalar function D which maps any vector belonging to the n-fold Cartesian product of $\{0, \ldots, t-1\}$ (denoted by T^n) into $T = \{0, \ldots, t-1\}$. For $x \in T^N$, we write x_i for $x(i)$.

Condition 3.3 (Systematicity) *There is a decision method D such that for all $(v_1, \ldots, v_n) \in V^n$ and all $\phi \in \Psi$, $A(v_1, \ldots, v_n)(\phi) = D(v_1(\phi), \ldots, v_n(\phi))$.*

Each of these conditions was introduced by List and Pettit (2003). They showed, for the case in which $t = 2$, that there is no aggregation function that satisfies them simultaneously. It follows from Theorem 4 in Pauly and Van Hees (2004) that the List-Pettit result can be strengthened by weakening anonymity to the demand that there is no dictator. A dictator is thereby defined as follows.

Condition 3.4 (Dictatorship) *There is some $i \in N$ such that for all $(v_1, \ldots, v_n) \in V^n$ and all $\phi \in \Psi$, $A(v_1, \ldots, v_n)(\phi) = v_i(\phi)$.*

Theorem 3.1 (Pauly and Van Hees 2004) *Let $t = 2$ and let minimal agenda richness be satisfied. An aggregation function is systematic if and only if it is a dictatorship.*

As stated, this theorem is a corollary of a more general result established by Pauly and Van Hees. That more general result applies to all cases in which $t > 1$. However, a similar general result cannot be derived in the framework employed here since, as explained in the introduction, it is based on a different many-valued logic. That no similar result for $t > 2$ can be obtained is shown by the following example.

Example. Let $t > 2$. If t is odd define D as a decision method such that $D(x_1, \ldots, x_n) = (t-1)/2$ for all $(x_1, \ldots, x_n) \in T^n$. If t is even let D be a decision method such that for some $j \in N$, for all $x_1, \ldots, x_n$, $D(x_1, \ldots, x_n) = 0$ if $x_j < t/2$ and $D(x_1, \ldots, x_n) = t-1$ otherwise. Define A as the procedure such that for all $(v_1, \ldots, v_n) \in V^n$ and all $\phi \in \Psi$, $A(v_1, \ldots, v_n)(\phi) = D(v_1(\phi), \ldots, v_n(\phi))$. It can be checked that, for any agenda, and hence also for those that satisfy minimal agenda richness, A is systematic but there is no dictator.[4]

Although the example shows that a dictatorship can be avoided when $t > 2$, the decision method that is used to illustrate this is still rather unattractive.

[4]Note that A thus defined does form a dictatorial rule if $t = 2$.

It either (if t is odd) always yields the same decisions, or (if t is even) the decision is unaffected by the opinion of all but one individual. Our main result shows that this unattractiveness cannot be avoided. Before we describe that result we present a very useful lemma.

Lemma 3.1 *Let A be an aggregation function and for all $\phi \in \Psi$ let there be a decision method D_ϕ such that for all $(v_1, \ldots, v_n) \in V^n$, $A(v_1, \ldots, v_n)(\phi) = D_\phi(v_1(\phi), \ldots, v_n(\phi))$. Then the following properties hold:*

1. *For every literal ϕ such that $\phi, \neg\phi \in \Psi$ and for all $x \in T^N$, we have $D_{\neg\phi}(t-1-x_1, \ldots, t-1-x_n)) = t-1-D_\phi(x_1, \ldots, x_n)$.*
2. *For all literals $\phi \not\approx \psi$ such that $\phi, \psi, \phi \wedge \psi \in \Psi$ and for all $x, y \in T^N$, $\min\{D_\phi(x_1, \ldots, x_n), D_\psi(y_1, \ldots, y_n)\} = D_{\phi\wedge\psi}(\min\{x_1, y_1\}, \ldots, \min\{x_n, y_n\})$.*

Proof. For the first claim, consider any $x = (x_1, \ldots, x_n) \in T^N$. Since we can construct a v_i on the basis of each possible assignment of truth values to the various atomic propositions, there is for any x_i at least one v_i and one literal ϕ such that $v_i(\phi) = x_i$. Then $v_i(\neg\phi) = t-1-x_i$, and hence $D_{\neg\phi}(t-1-x_1, \ldots, t-1-x_n) = D_{\neg\phi}(v_1(\neg\phi), \ldots, v_n(\neg\phi))$ which by our assumption must equal $A(v_1, \ldots, v_n)(\neg\phi)$. By definition of v we have $A(v_1, \ldots, v_n)(\neg\phi) = t-1-A(v_1, \ldots, v_n)(\phi) = t-1-D_\phi(v_1(\phi), \ldots, v_n(\phi))$ which in turn equals $t-1-D_\phi(x_1, \ldots, x_n)$.

For the second claim consider any $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. Again, since for each possible truth value assignment to atomic propositions there exists a corresponding v_i, there are literals ϕ and ψ and some v_i such that $v_i(\phi) = x_i$ and $v_i(\psi) = y_i$. In that case we have $\min\{D_\phi(x), D_\psi(y)\}$ $= \min\{D_\phi(v_1(\phi), \ldots, v_n(\phi)),\ D_\psi(v_1(\psi), \ldots, v_n(\psi))\}$ which by assumption must equal $\min\{A(v_1, \ldots, v_n)(\phi), A(v_1, \ldots, v_n)(\psi)\} = A(v_1, \ldots, v_n)(\phi \wedge \psi)$. Again by our assumption, this term must equal

$$D_{\phi\wedge\psi}(\min\{v_1(\phi), v_1(\psi)\}, \ldots, \min\{v_n(\phi), v_n(\psi)\})$$

which by definition equals $D_{\phi\wedge\psi}(\min\{x_1, y\}, \ldots, \min\{x_n, y_n\})$. □

In the context of a systematic aggegregation function and an agenda satisfying minimal agenda richness, we can derive the following corollary.

Corollary 3.1 *Let A be a systematic aggregation function and let minimal agenda richness be satisfied. Then there exists a decision method D such that*

1. *For all $x \in T^N$, we have $D(t-1-x_1, \ldots, t-1-x_n) = D(x_1, \ldots, x_n)$.*
2. *For all $x, y \in T^N$, $\min\{D(x_1, \ldots, x_n), D(y_1, \ldots, y_n)\} = D(\min\{x_1, y_1\}, \ldots, \min\{x_n, y_n\})$.*

Proof. By minimal agenda richness there are at least two atomic propositions (and hence also two distinct literals) p, q such that p, q, $(p \wedge q)$ and $\neg(p \wedge q)$ belong to the agenda. Systematicity states that there exists a D such that for all $(v_1, \ldots, v_n) \in V^n$ and all $\phi \in \Psi$, $A(v_1, \ldots, v_n)(\phi) = D(v_1(\phi), \ldots, v_n(\phi))$. Since the agenda contains at least two distinct literals the result follows directly from Lemma 3.1. □

Define the effective range $E(A)$ of A as $E(A) = \{k \mid \text{for some } (v_1, \ldots, v_n) \in V^n \text{ and some } \phi \in \Psi,\ A(v_1, \ldots, v_n)(\phi) = k\}$.

Condition 3.5 (Restricted Dictatorship) *There exists an $i \in N$ such that for all $k \in E(A)$ there is some $k' \in T$ such that $v_i(\phi) = k'$ implies $A(v_1, \ldots, v_n)(\phi) = k$ for all $(v_1, \ldots, v_n) \in V^n$ and all $\phi \in \Psi$.*

Theorem 3.2 *Let minimal agenda richness be satisfied. If an aggregation function is systematic, then it is a restricted dictorship.*

The proof is preceded by two lemmas. For any $x = (x_1, \ldots, x_n) \in T^n$, if for some i and some integers f, g we have $x_i = f$ and $x_j = g$ for all $j \neq i$, we shall also denote x by $(f^i, \overline{g})$. The set of all elements x of T^n in which $x_i = f$ is denoted by T_i^f.

Lemma 3.2 *Let minimal agenda richness be satisfied, and let A be systematic. For all $k \in E(A)$ there is some individual $i \in N$ and some $c \in T$ such that $D(c^i, \overline{t-1}) = k = D(c^i, \overline{0})$.*

Proof.

Step 1. First we show that for any $k \in E(A)$ there is some $i \in N$ and $c \in T$ such that $D(c^i, \overline{t-1}) = k$. Take an arbitrary integer $k \in E(A)$. By definition of $E(A)$ there is some $x = (x_1, \ldots, x_n)$ such that $D(x) = k$. Let x be 'maximal' with respect to k, i.e., there is no x^* such that $D(x^*) = k$ and $x_i^* \geq x_i$ for all i and $x_j^* > x_j$ for some j. Suppose that there are at least two distinct i, j such that $x_i \neq\ t-1\ \neq x_j$ and consider the profile $\tilde{x}$ defined as $\tilde{x}_i = x_i + 1$ and $\tilde{x}_g = x_g$ for all $g \neq i$ and the profile $\hat{x}$ in which $\hat{x}_j = x_j + 1$ and $\hat{x}_g = x_g$ for all $g \neq j$. By construction

$$(\min\{\tilde{x}_1, \hat{x}_1)\}, \ldots, \min\{\tilde{x}_n, \hat{x}_n\}) = x.$$

Since $D(x) = k$, it follows from Lemma 3.1 that $\min\{D(\tilde{x}), D(\hat{x})\} = k$. This, however, contradicts the fact that x is maximal with respect to k.

Hence there is at most one $i \in N$ such that $x_i \neq t-1$. If such an i indeed exists, let $c = x_i$. If there is no such i let $c = t-1$ and take arbitrary i. In either case there is some i and some integer c (viz. x_i and $t-1$, respectively), such that $D(c^i, \overline{t-1}) = k$.

Step 2. We next show that for any $k \in E(A)$, there is some $i \in N$ and $c \in T$, such that if $D(c^i, \overline{t-1}) = k$, then $D(c^i, \overline{0}) = k$. We do so by contradiction: we assume that there is some $k \in E(A)$ such that for all $i \in N$ and all $c \in T$ such that $D(c^i, \overline{t-1}) = k$ we have $D(c^i, \overline{0}) \neq k$. In particular, let k be the smallest element of $E(A)$ for which this is true, that is, for all $g \in E(A)$, if $g < k$ then there is a j and d such that $D(d^j, \overline{t-1}) = D(d^j, \overline{0}) = g$.

Case 1: $k > t-1-k$. By definition of k there is an individual j and a d such that $D(d^j, \overline{0}) = D(d^j, \overline{t-1}) = t-1-k$. From this it follows by (part 1 of) Lemma 3.1 that $D((t-1-d)^j, \overline{t-1}) = D((t-1-d)^j, \overline{0}) = k$, which contradicts the assumption that no i and c exists for which $D(c^i, \overline{t-1}) = D(c^i, \overline{0}) = k$.

Case 2: $t-1-k \geq k$. By Step 1 there is some i and some c for which $D(c^i, \overline{t-1}) = k$. We then have by (part 1) of Lemma 3.1 $D((t-1-c)^i, \overline{0}) = t-1-k$. Since $t-1-k \geq k$, we have by (part 2 of) Lemma 3.1

$$k = \min\{D(c^i, \overline{t-1}), D((t-1-c)^i, \overline{0})\} =$$

$$D(\min\{t-1, 0\}, \ldots, \underbrace{\min\{c, t-1-c\}}_{i}, \ldots, \min\{t-1, 0\}) =$$

$$D(\min\{c, (t-1-c)\}^i, \overline{0}).$$

Subcase 2.1: $c \leq t-1-c$. We then have $D(c^i, \overline{0}) = k$ and $D(c^i, \overline{t-1}) = k$, contradicting the way k was defined.

Subcase 2.2: $c > t-1-c$. Then $D((t-1-c)^i, \overline{0}) = k$. Since we also saw that $D((t-1-c)^i, \overline{0}) = t-1-k$, it must be true that $k = t-1-k$.

Let $D(c^i, \overline{0}) = g \neq k$. Since $D(c^i, \overline{t-1}) = k$, and since $(\min\{t-1, 0\}, \ldots, \underbrace{\min\{c, c\}}_{i}, \ldots, \min\{t-1, 0\}) = (c^i, \overline{0})$, we must have $g < k$ by part 2 of Lemma 3.1. From $D(c^i, \overline{0}) = g$, it follows by Lemma 3.1 that $D((t-1-c)^i, \overline{t-1}) = t-1-g$. Since $c > t-1-c$ and since $g < k = t-1-k$ implies $g < t-1-g$, we have by Lemma 3.1

$$g = \min\{D((t-1-c)^i, \overline{t-1}), D(c^i, \overline{0})\} =$$

$$D(\min\{t-1, 0\}, \ldots, \underbrace{\min\{t-1-c, c\}}_{i}, \ldots, \min\{t-1, 0\}) =$$

$$D((t-1-c)^i, \overline{0}).$$

Applying Lemma 3.1 once again shows that $D(c^i, \overline{t-1}) = t-1-g$, which contradicts the initial assumption that $D(c^i, \overline{t-1}) = k$. □

Lemma 3.3 *Assume minimal agenda richness is satisfied and let the aggregation rule A be systematic. For all integers c and all i, if $D(c^i, \overline{0}) = D(c^i, \overline{t-1})$, then $D(x) = D(y)$ for all $x, y \in T_i^c$.*

Proof. Let $D(c^i, \overline{0}) = D(c^i, \overline{t-1}) = k$ $(0 \leq k < t)$. Take any $(x_1, \ldots, x_n)$ in which $x_i = c$. Since

$$(\min\{0, x_1\}, \ldots, \min\{c, x_i\}, \ldots, \min\{0, x_n\}) = (0, \ldots, 0, \underbrace{c}_{i}, 0, \ldots, 0),$$

$D(x_1, \ldots, x_n) \geq k$ by part 2 of Lemma 3.1. Moreover, because

$$(\min\{t-1, x_1\}, \ldots, \min\{c, x_i\}, \ldots, \min\{t-1, x_n\}) = (x_1, \ldots, x_n),$$

$D(x_1, \ldots, x_n) \leq k$ by Lemma 3.1. Thus for any $(x_1, \ldots, x_n)$ in which $x_i = c$, we have $D(x_1, \ldots, x_n) = k$ □

Proof of Theorem 3.2 If $E(A)$ is as singleton set, say $E(A) = \{k\}$, the result is trivially true: any proposition is always assigned the value k. Assume therefore that $E(A)$ contains at least two distinct elements, say g and k. Lemma 3.2 and 3.3 imply that there are individuals i, j and integers c, d such that $D(x) = k$ for all $x \in T_i^c$ and $D(x) = g$ for all $x \in T_j^d$. If $i \neq j$, take a profile x in which $x_i = c$ and $x_j = d$. Since $x \in T_i^c \cap T_j^d$, $i \neq j$ would entail that $D(x) = g$ and $D(x) = k$, which is a contradiction. Hence, $i = j$. Since this is true for any $g, k \in E(A)$, i is a restricted dictator. □

If the effective range of an aggregation function contains only one alternative, then it assigns the same value to each proposition. Clearly, if the agenda contains at least one proposition and its negation, Lemma 3.1 implies that such an aggregation function can only assign $(t-1)/2$ as the value to all propositions. Note that this value lies exactly mid-way between $t-1$ ('true') and 0 ('false'). In a way, the aggregation procedure therefore refrains from making a judgement about the propositions. For this reason we call such an aggregation procedure *trivial*.

Definition 3.1 (Triviality) *An aggregation function A is trivial if for all $(v_1, \ldots, v_n), (v'_1, \ldots, v'_n) \in V^n$ and all $\phi, \psi \in \Psi$, $A(v_1, \ldots, v_n)(\phi) = A(v'_1, \ldots, v'_n)(\psi)$.*

Given Lemma 3.1 and minimal agenda richness, it is easily seen that for any t there is either exactly one such trivial aggregation function (if t is odd) or none at all. Furthermore, note that under a trivial aggregation function, everybody is a restricted dictator. Since there can only be a restricted dictator under an anonymous aggregation function if it is the trivial one, we can derive the following result as a corollary of Theorem 3.2

Corollary 3.2 *Let minimal agenda richness be satisfied. An aggregation function satisfies systematicity and anonymity if, and only if, it is the trivial function.*

Since an aggregation function can only be trivial if t is odd, and thus if $t > 2$, List and Pettit's result follows as a further corollary: given minimal agenda richness, there exists no anonymous and systematic aggregation function if $t = 2$.

Our next theorem concerns aggregation functions in which the effective range coincides completely with the set T of all truth values.

Condition 3.6 (Non-Imposition) *For any $k \in \{0, \dots, t-1\}$ there is some $\phi \in \Psi$ and some $(v_1, \dots, v_n) \in V^n$ such that $A(v_1, \dots, v_n)(\phi) = k$.*

Theorem 3.3 *Let minimal agenda richness be satisfied. An aggregation function satisfies systematicity and non-imposition if, and only if, it is a dictatorship.*

Proof. We only prove sufficiency. Non-imposition and Theorem 3.2 imply that for some $i \in N$ we can define a bijection f from T to T by stipulating that $f(c) = k$ if, and only if, $D(x) = k$ for all $x \in T_i^c$.

For any $c, c' \in T$ we have by Lemma 3.1

$$D(\min\{0,0\}, \dots, \min\{0,0\}, \underbrace{\min\{c, c'\}}_{i}, \min\{0,0\}, \dots, \min\{0,0\}) =$$

$$\min\{D(c^i, \overline{0}), D(c'^i, \overline{0}) =$$
$$\min\{f(c), f(c')\}.$$

It follows that f is a strictly increasing function: for all $c, c' \in T$, $f(c') \geq f(c)$ iff $c' \geq c$. But that in turn implies that for all $c \in T$ we have $f(c) = c$. For suppose to the contrary that for some c, $f(c) \neq c$. In particular, let c be the smallest element of T for which this is true, i.e., $f(c) > c$. Then by the fact that f is a bijection there must be some $d > c$ such that $f(d) = c$, contradicting the strictly increasing nature of f. □

4 Weakening systematicity

This section explores the consequences of weakening the condition of systematicity. Consider the following two conditions, both of which are variants of the similarly named social choice-theoretic conditions.

Condition 4.1 (Independence of irrelevant alternatives) *For all $(v_1, \dots, v_n)$, $(v'_1, \dots, v'_n) \in V^n$ and $\phi \in \Psi$, $A(v_1, \dots, v_n)(\phi) = A(v'_1, \dots, v'_n)(\phi)$ whenever for all $i \in N$ $v_i(\phi) = v'_i(\phi)$.*

The condition states that the collective judgement about a proposition depends only on the individual judgements about that proposition: it is independent of the judgements that individuals make about other propositions. Or, as the following lemma makes clear, independence implies that we can associate a decision method with each proposition.

Lemma 4.1 *An aggregation function A satisfies independence of irrelevant alternatives iff for every $\phi \in \Psi$ there is some decision method D_ϕ such that for all $(v_1, \ldots, v_n) \in V^n$, $A(v_1, \ldots, v_n)(\phi) = D_\phi(v_1(\phi), \ldots, v_n(\phi))$.*

Proof. See Lemma 1 of Pauly and Van Hees (2004).

The next condition is neutrality: it states that a collective judgement is invariant under permutations of the propositions.

Condition 4.2 (Neutrality) *For any permutation $f : \Psi \to \Psi$, any valuation profiles $(v_1, \ldots, v_n)$, $(v'_1, \ldots, v'_n) \in V^n$, if for all i and all $\phi \in \Psi$, $v_i(\phi) = v'_i(f(\phi))$, then $A(v_1, \ldots, v_n)(\phi) = A(v'_1, \ldots, v'_n)(f(\phi))$.*

Proposition 4.1 *If an aggregation function is systematic, then it satisfies independence of irrelevant alternatives and neutrality.*

Proof. The proof is straightforward and is therefore omitted. □

Systematicity is thus a very strong condition: propositions are always treated in the same way. The following example shows that by weakening systematicity to independence of irrelevant alternatives one can obtain possibility results.

Example. Suppose $\Psi = \{\phi, \psi, \phi \wedge \psi, \neg(\phi \wedge \psi)\}$, and define for all $(v_1, \ldots, v_n) \in V^n$, $A(v_1, \ldots, v_n)(\phi) = x$, $A(v_1, \ldots, v_n)(\psi) = y$, $A(v_1, \ldots, v_n)(\phi \wedge \psi) = x$ and $A(v_1, \ldots, v_n)(\neg(\phi \wedge \psi)) = t - 1 - x$, where $0 \leq x < y < t - 1 - x \leq t - 1$. Clearly, the agenda is minimally rich. Moreover, A satisfies independence of irrelevant alternatives but violates systematicity. Obviously, A is not a restricted dictatorship, which shows that independence of irrelevant alternatives (unlike systematicity) does not entail a restricted dictatorship.

The aggregation function, however, is rather unattractive. Even though it assigns different social values to different propositions, these assignments are completely independent of the individual values regarding those propositions. We would like to impose some responsiveness condition.[5]

Condition 4.3 (Minimal Responsiveness) $A(v)(\phi) = 0$ *and* $A(v')(\phi) \neq 0$ *for some $\phi \in \Psi$ and some $(v_1, \ldots, v_n)$, $(v'_1, \ldots, v'_n) \in V^n$.*

To derive our next result we use a strengthened agenda condition.

Condition 4.4 (Agenda Richness)

1. *Ψ contains at least two distinct atomic propositions p and q;*
2. *for all atomic propositions $p \in \Psi$: $\neg p \in \Psi$;*

[5] Note that the condition of Minimal Responsiveness is weaker than the condition of Responsiveness used by Pauly and Van Hees (2003) but stronger than their condition of Weak Responsivess.

3. *for all literals* $\phi, \psi \in \Psi(\phi \not\approx \psi)$*:* $(\phi \wedge \psi) \in \Psi$ *and* $\neg(\phi \wedge \psi) \in \Psi$*;*

4. *if* p *is an atomic proposition occurring in some* $\phi \in \Psi$*, then* $p \in \Psi$*.*

Theorem 4.1 *Let agenda richness be satisfied. If an aggregation function satisfies minimal responsiveness and independence of irrelevant alternatives, then it is systematic.*

Proof. Suppose to the contrary that A satisfies independence of irrelevant alternatives and minimal responsiveness but that it is not systematic. There are then $\phi, \psi \in \Psi$ such that for some $(v_1, \ldots, v_n) \in V^n$, $A(v_1, \ldots, v_n)(\phi) \neq A(v_1, \ldots, v_n)(\psi)$. In particular, by agenda richness and by definition of a valuation v, we have two literals ϕ and ψ for which this is true. First, suppose that $\phi \approx \psi$. If ϕ and ψ were the only literals for which there is some $(v_1, \ldots, v_n)$ such that $A(v_1, \ldots, v_n)(\phi) \neq A(v_1, \ldots, v_n)(\psi)$, it would be true that for all literals $\omega \not\approx \phi, \psi$ and all $(v_1^*, \ldots, v_n^*) \in V^n$, $A(v_1^*, \ldots, v_n^*)(\omega) = A(v_1^*, \ldots, v_n^*)(\phi)$ and $A(v_1^*, \ldots, v_n^*)(\omega) = A(v_1^*, \ldots, v_n^*)(\psi)$. Since Ψ contains at least one such literal ω distinct from ϕ, this would entail that $A(v_1, \ldots, v_n)(\phi) = A(v_1, \ldots, v_n)(\psi)$, which is a contradiction. Hence, $\phi \not\approx \psi$.

Thus we have $D_\phi(x) = a \neq b = D_\psi(x)$, for some $x \in T^N$. Assume without loss of generality that $a < b$. By Lemma 3.1 and 4.1 we then have $D_{\phi\wedge\psi}(x) = a$. Now consider $D_\phi(t-1, \ldots, t-1)$. Since $D_{\phi\wedge\psi}(x) = a \neq b$ and since $x_i \leq t-1$ for all i, it must be true that $D_\phi(t-1, \ldots, t-1) = a$. Hence, $D_{\neg\phi}(0, \ldots, 0) = t-1-a$. Since $t-1-a > t-1-b$, we have $D_{\neg\psi\wedge\neg\phi}(0, \ldots, 0) = t-1-b$. But then by Lemma 3.1 and 4.1 $D_{\neg\psi}(y) = t-1-b$ for all $y \in T^N$, and hence $D_\psi(y) = b$ for all $y \in T^N$. Since $a < b$, we have $b > 0$.

Now take some $\omega \in \Psi$, such that for some $y, y' \in T^n$, $D_\omega(y) = 0$ and $D_\omega(y') \neq 0$. By minimal responsiveness such a ω exists. By agenda richness and by definition of a valuation v this also holds for some literal δ occurring in ω. Since $D_\psi(0, \ldots, 0) = b$ and $D_\delta(y) = 0$ it must be true that $D_{\psi\wedge\delta}(0, \ldots, 0) = 0$. However, from $D_\psi(0, \ldots, 0) = b$ and $D_\delta(y') \neq 0$ it also follows that $D_{\psi\wedge\delta}(0, \ldots, 0) \neq 0$, which is a contradiction. □

Corollary 4.1 *Let agenda richness be satisfied. If an aggregation function satisfies minimal responsiveness and independence of irrelevant alternatives then it is a restricted dictatorship.*

Theorems 3.3 and 4.1 together imply:

Corollary 4.2 *Let agenda richness be satisfied. An aggregation function satisfies minimal responsiveness, non-imposition and independence of irrelevant alternatives if, and only if, it is a dictatorship.*

5 Conclusion

The results presented in this paper seem to underscore Socrates' argument that the possibility of disagreement among individuals (even if they are gods) entails the possibility of inconsistency at the collective level: if we want to derive a collective judgement on the basis of the judgements of the members constituting that collective, we cannot ensure consistency of the resulting judgement. Indeed, the conclusion that the costs of ensuring consistency of collective judgements comes at the price of having a dictatorial aggregation procedure would probably not have surprised Plato too much. The rulers in his *Republic* were to be the philosophers, i.e., those who are supposed to have access to the truth (and thus presumably have a better insight than the disagreeing gods). If there is exactly one philosopher who possesses such insight, then this person should be the unique ruler. In the terminology of this paper, he should be the dictator. The formal results established here would then seem to underscore Plato's view that knowledge is not to be obtained by a democratic amalgamation of individual beliefs.

Such a stance, though, does not seem to be very compelling for us - after all, if none of the individuals constituting a group has the wisdom that Plato ascribes to his philosophers, a dictatorial rule for arriving at collective judgements is not very attractive. (Moreover, even if there were some such 'true knower' within a group, how can the group know who it is; that is, how can the group arrive at such a collective judgement? Surely, if they do not yet know who the seer is they cannot use the dictatorial rule.) So how negative are these results?

First, it should be noted that the results do not show that inconsistency *always* arises under a non-dictatorial decision-making procedure satisfying independence or the stronger demand of systematicity. The results point out that we cannot *ensure* consistency of the resulting collective judgements, but it may well be the case that such consistency is arrived at in particular situations. Thus a first line of enquiry would be to examine the circumstances under which possibility results would emerge, that is, we may want to examine the kinds of judgement profiles in which the problems arise or the probability that an inconsistency arises.[6]

Secondly, we may want to give up the condition of independence. One possibility, for instance, is to give some propositions the special status of being *premises*. One then makes collective judgements only about these premises, allowing the collective verdict on other propositions to be dependent on the verdict about these premises (for a discussion see Pettit 2004). Another possibility is to say that some of the individuals are experts on some propositions. If we let each expert's judgement determine

[6] For respective answers for the two-valued case, see Nehring and Puppe (2004) and List (forthcoming).

the collective judgement on that particular proposition, we may also obtain possibility results. A difficulty with such approaches, however, just as with the selection of Plato's 'philosopher-dictator', is that in the absence of an independent criterion for establishing which propositions should be viewed as premises, or which individuals are experts on certain propositions, the choice of those premises or experts will itself be subject to a decision procedure, thereby shifting the problem to a higher level. Moreover, List and Dietrich (2004) recently showed that the existence of experts lead to new impossibilities if we impose a criterion of unanimity, that is, if we demand that if everybody assigns the same truth-value to a proposition, this truth-value should also be assigned to it at the collective level.

A different question related to the purport of the results derived is the question of the extent to which they can be formally related to the results established in social choice theory. We do not yet have a general impossibility result from which we can deduce both the main impossibility results of social choice theory and the negative results on judgement aggregation. One possibility was mentioned in the introduction and consists of taking the propositions to be about the preferences of individuals. Another approach to bridging the gap between the two frameworks would be, as we did, to let the propositions represent *aspects* of social states. Since such an aspect can also be described as a set of social states - viz. the set of all social states that have that same feature - the individual valuation functions can then be seen as preferences over *sets* of social states. The next step would then be to examine the aggregation of individual rankings over such sets into a social ranking of these sets. Since the rankings can be interpreted as either preferences or as 'belief-rankings' we would have a framework in which we can analyse both the preferential approach to collective decision making as well as the epistemic one.

Apart from the quest for a formal framework in which both epistemic and preferential aspects of democratic decision making can be analysed, we should also consider the *substantive* implications of the results on epistemic decision making. If epistemic considerations do indeed play the important role that many democratic theorists ascribe to them, then we need not end up with a genuinely negative message. Instead, it may provide an impetus to analyse the dynamic aspects of decision making processes. For instance, the possible inconsistency at the collective level can be seen as a reason to adjust individual judgements. After such an adjustment has taken place, a collective decision is again aimed at. If a coherent one still does not emerge, a new adjustement takes place, etc. Clearly, the conditions that such a process of 'updating beliefs' or 'deliberation' would have to satisfy, or the circumstances under which possibility results will emerge are open questions. However, as with the avenues of research described above, the questions show that the negative results on epistemic aggregation do not close down the analysis of epistemic decision making but, by indicating its limits, they form its starting point.

References

Cohen, J., 'An epistemic conception of democracy', *Ethics*, 97: 26-38, 1986

Coleman, J., and J. Ferejohn, 'Democracy and social choice', *Ethics*, 97: 6-25, 1986

Dietrich, F., 'Judgment aggregation: (im)possibility theorems', 2004, Working Paper

Dietrich, F. and Chr. List, 'A liberal paradox for judgment aggegation', 2004, Working Paper

Gärdenfors, P., 'A representation theorem for voting with logical consequences', 2004, Working Paper

Kornhauser, L. and L. Sager, 'The one and the many: Adjudication in collegial courts', *California Law Review*, 81: 1-59, 1993.

List, Chr. and F. Dietrich, 'A Liberal Paradox for Judgment Aggregation', 2004, mimeo.

List, Chr. and Ph. Pettit, 'Aggregating sets of judgments: An impossibility result', *Economics and Philosophy*, 18:89–110, 2002.

List, Chr. and Ph. Pettit, 'Aggregating sets of judgments: Two impossibility results compared', *Synthese*, 140: 207-235, 2004

List, Chr., 'The probability of inconsistencies in complex collective decisions', *Social Choice and Welfare*, forthcoming

Nehring, K., and C. Puppe, 'Consistent judgement aggregation: A characterization', Working Paper

Pauly, M. and M. van Hees, 'Logical constraints on judgement aggregation', *Journal of Philosophical Logic*, forthcoming

Pettit, Ph., 'Deliberative democracy and discursive dilemma', *Philosophical Issues*, 11: 268-299, 2001

Post, E., 'Introduction to a general theory of elementary propositions', *American Journal of Mathematics*, 43:163–185, 1921.

Rescher, N., *Many-Valued Logic*, New York: McGraw-Hill, 1969

Urquhart, A., *Basic Many-Valued Logic*, in: D.M. Gabbay and F.Guenther (eds.), *Handbook of Philosophical Logic, 2nd edition, Vol.2*, 249-295, Dordrecht: Kluwer Academic, 2001.