



Stanford Encyclopedia of Philosophy

## Voting Methods

*First published Wed Aug 3, 2011*

Think back to the last time you needed to make a decision as a member of a group. This may have been when you voted for your favorite political candidate during the last election. On a smaller scale, it may have been when you took part in a committee that needed to choose the best candidate for a job or a student to receive a special award. What method, or procedure, did the group use to make the final decision? Many interesting issues arise when we carefully examine our group decision-making processes. Consider a simple example of a group of friends deciding where to go for dinner. If everyone agrees on which restaurant is best, then it is obvious where to go. But how should the friends decide where to go if they have different opinions about which restaurant is best? Is there always a choice that is “fair” taking into account everyone's opinions? Or are there situations in which one person must be chosen to act as a “dictator” by making a unilateral decision?

This article introduces and critically examines a number of different voting methods. The goal is not to provide a general overview of social choice theory or even a comprehensive account of voting theory. Rather, my objective is to highlight and discuss key results and issues that underlie phenomena that we observe when decision makers come together to make a collective decision. So, some topics will only briefly be mentioned, while others will not be discussed at all: Notable omissions include the extensive literature on the discursive dilemma (see List 2006, and references therein) and an overview of the work on voting power indices (Felsenthal and Machover 1998). To learn more about these topics, consult Nurmi (1998) and Saari (2001) for general introductions to voting theory and Brams and Fishburn (2002) and Saari (1995) for technical introductions and analysis of the vast literature.

- [1. The Problem: Who Should be Elected?](#)
  - [1.1 Notation](#)
- [2. Examples of Voting Methods](#)
- [3. Voting Paradoxes](#)
  - [3.1 Condorcet's Paradox](#)
    - [3.1.1 Electing the Condorcet Winner](#)

- [3.2 Failures of Monotonicity](#)
  - [3.3 The Multiple Districts Paradox](#)
  - [3.4 The Multiple Elections Paradox](#)
  - [4. Topics in Voting Theory](#)
    - [4.1 Strategizing](#)
    - [4.2 Characterization Results](#)
    - [4.3 Voting to Track the Truth](#)
  - [5. Concluding Remarks: from Theory to Practice](#)
  - [Bibliography](#)
  - [Academic Tools](#)
  - [Other Internet Resources](#)
  - [Related Entries](#)
- 

# 1. The Problem: Who *Should* be Elected?

The central question of this article is:

Given a group of people faced with some decision, how should a central authority combine the individual opinions so as to best reflect the “will of the group”?

A complete analysis of this question would incorporate a number of different issues ranging from central topics in political philosophy (e.g., how should we define the “will” of the people? what is a democracy?) to the psychology of decision making. In this article, I focus on one aspect of this question: the formal analysis of specific voting methods (see, for example, Riker 1982; Mackie 2003, for a more comprehensive analysis of the above question, incorporating many of the issues raised in this article).

I start with a concrete example to illustrate the type of analysis surveyed in this article. Suppose that there is a group of 21 people, or voters, who need to make a decision about which of four candidates, or options, should be elected, or chosen. Let A, B, C and D denote the four different candidates. The first step is to decide how to represent the voters' *opinions* about the set of candidates. Many different approaches have been explored in the voting theory literature. One approach is to assume that each voter has an ordinal preference ordering over the set of candidates, describing the relative rankings of the candidates. A second approach assumes that voters assign to each candidate a cardinal value describing how much that voter prefers or *values* the candidate. Finally, one can describe an underlying space of issues, how much each voter “cares” about each issue and the degree to which each candidates supports the different issues. Unless otherwise stated, I follow much of the voting theory literature and assume that the voters' opinions are described by linear rankings of the set

of candidates (describing the voters' ordinal preference orderings).

For this example, assume that each of the voters has one of four possible rankings of the candidates. The information about the rankings of each voter is given in the following table.

# Voters			
<b>3</b>	<b>5</b>	<b>7</b>	<b>6</b>
A	A	B	C
B	C	D	B
C	B	C	D
D	D	A	A

Read the table as follows: Each column represents a ranking in which candidates in lower rows are ranked lower. The numbers at the top of each column indicate the number of voters with that particular ranking. Suppose that you are an outside observer without any interest in the outcome of this election. Which of the candidates best represents the “will” of this group? If there were only two candidates to choose from, there is a very intuitive answer: The winner should be the candidate or option that is supported by more than 50 percent of the voters (cf. the discussion below about May's Theorem in Section 4.2). However, if there are more than two candidates, as in the above example, the statement “the candidate that is supported by more than 50 percent of the voters” can be interpreted in different ways, leading to different ideas about who should win the election.

One candidate who, at first sight, seems to be a good choice to win the election is candidate A. Candidate A is ranked first in more of the voters' rankings than any other candidate. (A is ranked first by eight voters, B is ranked first by seven; C is ranked first by six; and D is not ranked first by any of the voters.) That is, more people think that A is better than any other candidate.

Of course, 13 people rank A *last*, so a much larger group of voters will be unsatisfied with the election of A. So, it seems clear that A should *not* be elected. None of the voters rank D first, which suggests that D is also not a good choice. The choice, then, boils down to B and C. Here, there are good arguments for each of B and C to be elected. This echoes an 18th-century debate between the two founding fathers of voting theory, Jean-Charles de Borda (1733–1799) and M.J.A.N. de Caritat, Marquis de Condorcet (1743–1794). For a precise history of voting theory as an academic discipline, including Condorcet's and Borda's writings, see McClean and Urken (1995). I sketch the

intuitive arguments for the election of B and C below.

*Candidate C should win.* Initially, this might seem like an odd choice since C received the fewest number of first-place rankings (6). However, C is a strong choice because he beats every other candidate in a one-on-one election. To see this, we need to examine how the population would vote in the various two-way elections:

# Voters			
3	5	7	6
A	A	B	C
B	C	D	B
C	B	C	D
D	D	A	A

13 rank C above A; 8  
rank A above C

# Voters			
3	5	7	6
A	A	B	C
B	C	D	B
C	B	C	D
D	D	A	A

11 rank C above B;  
10 rank B above C

# Voters			
3	5	7	6
A	A	B	C
B	C	D	B
C	B	C	D
D	D	A	A

15 rank C above D; 7  
rank D above C

The idea is that C should be declared the winner since he beats every other candidate in one-on-one elections. A candidate with this property is called a **Condorcet winner**. (We can similarly define a **Condorcet loser**. In fact, in the above example, candidate A is the Condorcet loser since she loses to every other candidate in head-to-head elections.)

*Candidate B should win.* Consider B's performance in head-to-head elections.

# Voters			
3	5	7	6
A	A	B	C
B	C	D	B
C	B	C	D
D	D	A	A

# Voters			
3	5	7	6
A	A	B	C
B	C	D	B
C	B	C	D
D	D	A	A

# Voters			
3	5	7	6
A	A	B	C
B	C	D	B
C	B	C	D
D	D	A	A

13 rank B above A; 8  
rank A above B

10 rank B above C;  
11 rank C above B

21 rank B above D; 0  
rank D above B

Candidate B performs the same as C in a head-to-head election with A, loses to C by only one vote and beats D in a landslide (everyone prefers B over D). Arguably, we should take into account *all* of these facts when determining who should represent the will of the people. Borda's idea is to assign each candidate a score that reflects *all* of this information. Both Condorcet and Borda suggest comparing candidates in one-on-one elections in order to determine the winner. While Condorcet tallies how many of the head-to-head races each candidate wins, Borda suggests that one should look at the margin of victory or loss. According to Borda, each candidate should be assigned a score representing how much support he or she has among the electorate. One way to calculate the score for each candidate is as follows (I will give an alternative method, which is easier to use, in the next section):

- A receives 24 points (8 votes in each of the three head-to-head races)
- B receives 44 points (13 points in the competition against A, plus 10 in the competition against C plus 21 in the competition against D)
- C receives 38 points (13 points in the competition against A, plus 11 in the competition against B plus 14 in the competition against D)
- D receives 20 points (13 points in the competition against A, plus 0 in the competition against B plus 7 in the competition against C)

The candidate with the highest score (in this case, B) is the one who should be elected.

The conclusion is that in voting situations with more than two candidates, there may not always be one obvious candidate that “best reflects the will of the people.” The remainder of this entry will discuss different methods, or procedures, that can be used to determine the winner of an election.

## 1.1 Notation

In this article, I will keep the formal details to a minimum; however, it is useful at this point to settle on some terminology. Assume that there is a finite set of voters  $V$  and a finite set of candidates  $X$ . I use lowercase letters  $i, j, k, \dots$  to denote elements of  $V$  and uppercase letters  $A, B, C, \dots$  to denote elements of  $X$ . Different voting methods require different types of information from the voters as input. For example, some methods ask voters to select a single candidate or a set of candidates, while other methods ask voters to linearly rank all of the candidates. The input requested from the voters are called **ballots**. A **profile** is

a sequence of ballots, one from each voter. The second component of a voting procedure is the method used to calculate the winner, given a profile of ballots.

As noted above, one underlying assumption is that the voters' actual *desires* about who should win the election are represented as linear preference relations over the set of candidates. Given a set of candidates  $X$ , let  $L(X)$  denote the set of linear orderings on  $X$  (that is, relations on  $X$  that are irreflexive, transitive, and complete). These orderings are intended to represent the voters' *ordinal* preferences about the relative rankings of each of the candidates (see the entry on preferences, Hansson and Grüne-Yanoff 2009, for an extended discussion of these properties and other issues surrounding formal modeling preferences). We use  $P_i$  to denote voter  $i$ 's preference ordering over  $X$ . It is important to note that these orderings do not reflect any cardinal information (for example, the intensity of the preference of one candidate over another). For instance, suppose that there are three candidates  $X=\{A,B,C\}$ . Then, the assumption is that a voter's "preference" can be any one of the six possible linear orderings over  $X$ :

Preference	$P_1$	$P_2$	$P_3$	$P_4$	$P_5$	$P_6$
	A	A	B	B	C	C
	B	C	A	C	A	B
	C	B	C	A	B	A
# Voters	$n_1$	$n_2$	$n_3$	$n_4$	$n_5$	$n_6$

I can now be more precise about the definition of a Condorcet winner (loser). The key notion here is the **majority relation**, which is the ranking of candidates in terms of how they perform in one-on-one elections. Formally, we write  $A >_M B$ , provided that more voters rank candidate A above candidate B than the other way around (we write  $\geq_M$  if there are ties). So, if the distribution of preferences is given in the above table, we have:

- $A >_M B$  just in case  $n_1+n_2+n_5 > n_3+n_4+n_6$  (otherwise  $B \geq_M A$ )
- $A >_M C$  just in case  $n_1+n_2+n_3 > n_4+n_5+n_6$  (otherwise  $C \geq_M A$ )
- $B >_M C$  just in case  $n_1+n_3+n_4 > n_2+n_5+n_6$  (otherwise  $B \geq_M C$ )

Candidate A is called the **Condorcet winner** if A is maximal in the majority ordering  $>_M$ . The **Condorcet loser** is the candidate that minimizes this ordering.

I conclude this section with a few comments on the relationship between the ballots and the voters' opinions about the candidates. Two issues are important

to keep in mind. First, the ballots of a particular voting method are intended to reflect *some* aspect of the voters' opinions about the desirability of the different candidates. Some types of ballots are intended to represent all or part of the voter's preference ordering, while other types represent information that cannot be inferred directly from the voter's *ordinal* preference ordering (for example, by describing *how much* a voter likes a particular candidate). Second, it is important to be precise about the type of considerations voters take into account when selecting a ballot. One approach is to assume that voters choose *sincerely* by selecting the ballot that best reflects their view about the desirability of the different candidates. A second approach assumes that voters choose *strategically*. In this case, a voter selects a ballot that she *expects* to lead to her most desired outcome given the information she has about how the other members of the group will vote. Strategic voting is an important topic in voting theory and social choice theory (see Taylor 2005, for a discussion and pointers to the literature), but in this article, unless otherwise stated, I assume that voters choose sincerely.

## 2. Examples of Voting Methods

A voting procedure is a way of aggregating the individual's preferences in order to come to a collective decision. A quick survey of elections held in different democratic societies throughout the world reveals a wide variety of methods. In this section, I discuss some of the key procedures that have been analyzed in the voting theory literature. These procedures may be of interest because they are widely used (e.g., plurality rule or plurality rule with runoff) or because they are of theoretical interest (e.g., Dodgson's method). I do not provide a comprehensive overview of the different methods that have been discussed in the literature (see Brams and Fishburn 2002, for a systematic overview of different voting methods). Rather, I focus on methods that either are familiar or help illustrate important ideas. I start with the most widely used method:

### **Plurality Rule:**

Each voter selects one candidate (or none if voters can abstain), and the candidate(s) with the most votes win. So, the ballots are simply the set of candidates  $X$  and, given voter  $i$ 's true preference ordering  $P_i$ , the unique sincere ballot for voter  $i$  is  $top(P_i)$  (the maximal element in the ordering  $P_i$ ).

Plurality rule is a very simple method that is widely used despite its many problems. The most pervasive problem is the fact that plurality rule can elect a Condorcet loser. Borda (1784) observed this phenomenon in the 18th century.

# Voters
----------

	<b>1</b>	<b>7</b>	<b>7</b>	<b>6</b>
A	A	A	B	C
B	B	C	C	B
C	C	B	A	A

Candidate A is the Condorcet loser (both B and C beat candidate A, 13 – 8); however, A is the plurality rule winner. In fact, the plurality ranking (A is first with eight votes, B is second with seven votes and C is third with six votes) reverses the majority ordering  $C >_M B >_M A$ . But there are other (more basic) reasons to criticize plurality rule. For instance, the very simple plurality ballots severely limit what the voters can express about their opinions of the candidates. **Ranked voting procedures** ask for much more information from the voter: the ballots are linear orderings of the candidates. The most well-known example of such a procedure is Borda Count:

### **Borda Count:**

Each voter provides a linear ordering of the candidates. Each candidate is assigned a score (the Borda score) as follows: If there are  $n$  candidates, give  $n-1$  points to candidates ranked first,  $n-2$  points to candidates ranked second,..., 1 point to a candidate ranked 2nd to last and 0 points to candidates ranked last. So, the Borda score of  $A$ , denoted  $BS(A)$ , is calculated as follows (where  $\#U$  denotes the number elements in the set  $U$ ):

$$\begin{aligned}
 BS(A) = & (n-1) \times \#\{i \mid i \text{ ranks } A \text{ first}\} + \\
 & (n-2) \times \#\{i \mid i \text{ ranks } A \text{ second}\} + \dots + \\
 & 1 \times \#\{i \mid i \text{ ranks } A \text{ second to last}\} + \\
 & 0 \times \#\{i \mid i \text{ ranks } A \text{ last}\}
 \end{aligned}$$

The candidate with the highest Borda score wins.

Recall the example discussed in the introduction to Section 1. We can calculate the Borda score for each of the candidates as follows:

$$\begin{aligned}
 BS(A) &= 3 \times 8 + 2 \times 0 + 1 \times 0 + 0 \times 13 = 24 \\
 BS(B) &= 3 \times 7 + 2 \times 9 + 1 \times 5 + 0 \times 0 = 44 \\
 BS(C) &= 3 \times 6 + 2 \times 5 + 1 \times 10 + 0 \times 0 = 38 \\
 BS(D) &= 3 \times 0 + 2 \times 7 + 1 \times 6 + 0 \times 8 = 20
 \end{aligned}$$

Borda Count requires the voters to come up with a linear ranking of *all* the candidates. This can be rather demanding when there are a large number of candidates (as it can be difficult for voters to make distinctions between some of



the more obscure candidates). A second way to make a voting method sensitive to more than the voters' top choice is to hold “multi-stage” elections. The different stages can come in the form of actual “runoff” elections in which voters are asked to choose from a reduced set of candidates; or they can be built in to the way the winner is calculated by asking voters to submit linear orderings over the set of all candidates. The following are the most well-known examples of multi-stage voting methods:

### **Plurality with Runoff:**

Start with a plurality vote to determine the top two candidates (or more if there are ties). Then, there is a runoff between these candidates, and the candidate with the most votes wins. Sometimes, a runoff can be avoided if the top candidate gets a sufficiently large percentage of the votes (for example, if she gets an absolute majority: more than 50 percent of the votes).

Rather than focusing on the top two candidates, one can also iteratively remove the candidate(s) with the fewest first-place votes:

### **The Hare Rule:**

The ballots are linear orders over the set of candidates. Repeatedly delete the candidate or candidates that receive the fewest first-place votes, with the remaining candidate(s) declared the winner (or winners in the case of ties).

If there are only three candidates, then the above two procedures are the same (removing the candidate with the least number of votes is the same as keeping the top two candidates). The following example shows that these two procedures can conflict when there are more than three candidates:

# Voters			
7	5	4	3
A	B	D	C
B	C	B	D
C	D	C	A
D	A	A	B

Candidate A is the plurality-with-runoff winner: Candidates A and B are the top two candidates, receiving seven and five votes, respectively, in the first round. In the runoff election, the groups voting for candidates C and D give their support to candidate B and A, respectively, with A winning 10 – 9.

However, Candidate D wins with the Hare rule: In the first round, candidate C is eliminated after receiving only three votes. But then this group's votes are transferred to D, giving her seven votes. This means that in the second round, candidate B has the fewest votes (five votes) and so is eliminated. After the elimination of candidate B, candidate D has an absolute majority with 12 total votes (note that in this round the group in the second column transfers all their votes to D since C was eliminated in an earlier round).

One final procedure is Coombs rule, which iteratively removes the candidates with the most last-place votes.

**Coombs Rule:**

Each voter submits a linear ordering over the set of candidates. Candidates who are ranked last by the most voters are iteratively removed. The last candidate(s) to be removed are the winner(s).

In the above example, candidate B wins the election using Coombs rule. In the first round, A, with nine last-place votes, is eliminated. The next candidate to be eliminated is D, with 12 last-place votes. Finally, C, with 16 last place votes, is eliminated.

The next type of procedures ask voters to submit ballots that represent information that cannot be inferred directly from their ordinal preference orderings. The first example gives voters the option to either select a candidate that they want to vote *for* (as in plurality rule) or to select a candidate that they want to vote *against*.

**Negative Voting:**

Each voter is allowed to choose one candidate to either vote *for* (giving the candidate one point) or to vote *against* (giving the candidate -1 points). The winner(s) is(are) the candidate(s) with the highest score(s) (i.e., the most positive votes).

Negative voting is tantamount to allowing the voters to support either a single candidate or all but one candidate (taking a point away from a candidate C is equivalent to giving one point to all candidates except C). That is, the voters are asked to choose a set of candidates that they support, where the choice is between sets consisting of single candidates or sets consisting of all except one candidate. The next procedure generalizes this idea by allowing voters to choose *any* subset of candidates:

**Approval Voting:**

Each voter selects a *subset* of the candidates (where the empty set means the voter abstains) and the candidate(s) with the most votes wins.

Approval voting has been extensively discussed by Steven Brams and Peter Fishburn (Brams and Fishburn 2007; Brams 2008). See, also, the recent collection of articles devoted to approval voting (Laslier and Sanver 2010).

Approval voting forces voters to think about the decision problem differently: They are asked to determine which candidates they *approve* of rather than determining the relative ranking of the candidates. That is, the voter is asked which candidates are above a certain “threshold of acceptance”. (See Brams and Sanver 2009, for examples of voting procedures that ask voters to both select a set of candidates that they approve *and* to (linearly) rank the candidates.) The final type of procedures I introduce in this section allow voters to express their *intensity* of preference among the candidates.

### **Cumulative Voting:**

Each voter is asked to distribute a fixed number of points, say ten, among the candidates in any way they please. The candidate(s) with the most points wins the election.

This general idea was taken further in a recent proposal for a new method of voting by Michel Balinski and Rida Laraki (2007). The general idea of their new method (**majoritarian judgment**) is that voters assign grades to each candidate from a commonly accepted *grading language*. Once the grades are assigned, each candidate is assigned her *median grade*. The winner(s) is(are) the candidate(s) with the highest median grade. The details of this procedure are beyond the scope of this article, but they can be found along with axiomatic characterizations in Balinski and Laraki (2010).

This section introduced a number of different procedures that can be used to make a group decision. One striking fact is that many of the different procedures give conflicting results on the same input. This raises an important question: How should we *compare* the different procedures? Can we argue that some procedures are better than others? There are a number of different criteria that can be used to compare and contrast different voting methods:

1. **Pragmatic concerns:** Is the procedure easy to *use*? Is it *legal* to use a particular voting procedure for a national or local election? The importance of “ease of use” should not be underestimated: Despite its many flaws, plurality rule (arguably the simplest voting procedure to use and understand) is, by far, the most commonly used method (cf. the discussion by Levin and Nalebuff 1995, p. 19).
2. **Behavioral considerations:** Do the different procedures *really* lead to different outcomes in practice? An interesting strand of research, *behavioral social choice*, incorporates empirical data about actual elections into the general theory of voting (This is discussed briefly in Section 5. See Regenwetter *et al.* 2006, for an extensive discussion).

3. **Information required from the voters:** What type of information do the ballots convey? While ranked procedures (e.g., Borda Count) require the voter to compare *all* of the candidates, it is often useful to ask the voters to report something about the “intensities” of their preferences over the candidates. Of course, there is a trade-off: Limiting what voters can express about their opinions of the candidates often makes a procedure much easier to use and understand.
4. **Axiomatic characterization results and voting paradoxes:** Much of the work in voting theory has focused on comparing and contrasting voting procedures in terms of abstract principles that they satisfy. The goal is to characterize the different voting procedures in terms of *normative* principles of group decision making. See Sections 3 and 5.2 for discussions.

### 3. Voting Paradoxes

In this section, I introduce and discuss a number of *voting paradoxes* — i.e., anomalies that highlight problems with different methods. See Saari (1995, 2001) and Nurmi (1999) for penetrating analyses that explain the underlying mathematics behind the different voting paradoxes.

#### 3.1 Condorcet's Paradox

A very common assumption is that a *rational* preference ordering must be *transitive* (i.e., if A is preferred to B, and B is preferred to C, then A must be preferred to C. See the entry on preferences (Hansson and Grüne-Yanoff 2009) for an extended discussion of the rationale behind this assumption). Indeed, if a voter's preference ordering is not transitive, allowing for cycles ( $A > B > C > A$ ), then there is no candidate that the voter can be said to actually support (for each candidate, there is another candidate that the voter prefers). Such voters have contradictory opinions about the candidates and, arguably, should be ignored or eliminated by any voting system. Many authors argue that such voters with cyclic preference orderings have inconsistent opinions about the candidates and should be *ignored* by any voting procedures (in particular, Condorcet forcefully argued this point). A key observation of Condorcet (which has become known as Condorcet's Paradox) is that even if each voter's preference ordering is transitive, the majority ordering may not be transitive.

Condorcet's original example was more complicated, but the following situation with three voters and three candidates illustrates the phenomenon:

# Voters
----------

<b>1</b>	<b>1</b>	<b>1</b>
A	C	B
B	A	C
C	B	A

Note that we have:

- Candidate A beats candidate B 2-1 in a one-on-one election.
- Candidate B beats candidate C 2-1 in a one-on-one election.
- Candidate C beats candidate A 2-1 in a one-on-one election.

Thus, we have a **majority cycle**  $A >_M B >_M C >_M A$ , and so there is no Condorcet winner. One interpretation is that, although each of the individual voters has a rational preference ordering, the group's preference ordering (defined as the majority ordering) is not rational. This simple, but fundamental observation has been extensively studied (see Gehrlein 2006, for an overview of the literature).

### 3.1.1 Electing the Condorcet Winner

Condorcet's Paradox shows that there may not always be a Condorcet winner in an election. However, one natural requirement for a voting rule is that if there is a Condorcet winner, then that candidate should be elected. Voting procedures that satisfy this property are called **Condorcet consistent**. Many of the procedures introduced above are not Condorcet consistent. I already presented an example showing that plurality rule is not Condorcet consistent (in fact, plurality rule may even elect the Condorcet *loser*).

The example from Section 1 shows that Borda Count is not Condorcet consistent. In fact, this is an instance of a general phenomenon that Fishburn (1974) called **Condorcet's other paradox**. Consider the following voting situation with 81 voters and three candidates from Condorcet (1785).

# Voters					
<b>30</b>	<b>1</b>	<b>29</b>	<b>10</b>	<b>10</b>	<b>1</b>
A	A	B	B	C	C
B	C	A	C	A	B
C	B	C	A	B	A

The majority ordering is  $A >_M B >_M C$ , so A is the Condorcet winner. Using the

Borda rule, we have:

$$BS(A) = 2 \times 31 + 1 \times 39 + 0 \times 11 = 101$$

$$BS(B) = 2 \times 39 + 1 \times 31 + 0 \times 11 = 109$$

$$BS(C) = 2 \times 11 + 1 \times 11 + 0 \times 59 = 33$$

So, candidate B is the Borda winner. Condorcet pointed out something more: The only way to elect candidate A using *any* scoring method is to assign more points to candidates ranked second than to candidates ranked first. A scoring method, which generalizes the Borda score, is defined by first fixing a nondecreasing sequence of real numbers  $s_0 \leq s_1 \leq \dots \leq s_{n-1}$  with  $s_0 < s_{n-1}$ . The idea is to assign a score to each candidate by multiplying the number of  $j$ th-place votes they receive by a  $s_{j-1}$ , and then adding all the results together over all values of  $j$ . To simplify the calculation, assume that candidates ranked first receive two points, and candidates ranked last receive no points. Let  $v$  be the number of points assigned to candidates ranked second. Then, the scores assigned to candidates A and B are as follows:

$$\text{Score}(A) = 2 \times 31 + v \times 39 + 0 \times 11$$

$$\text{Score}(B) = 2 \times 39 + v \times 31 + 0 \times 11$$

So, in order for  $\text{Score}(A) > \text{Score}(B)$ , we must have  $2 \times 31 + v \times 39 > 2 \times 39 + v \times 31$ , which implies that  $v > 2$ . But, of course, it is counterintuitive to give more points for being ranked second than for being ranked first. Peter Fishburn generalized this example as follows:

**Theorem** (Fishburn 1974).

For all  $m \geq 3$ , there is some voting situation with a Condorcet winner such that every weighted scoring rule will have at least  $m-2$  candidates with a greater score than the Condorcet winner.

So, no scoring rule is Condorcet consistent, but what about other methods? The following example from Steven Brams (2008, Chapter 3) shows that there are situations in which no fixed voting rule can elect a Condorcet winner. A **fixed voting rule** (or  $k$ -Approval Voting) is a method by which the voters choose a predetermined number of candidates. For example, plurality is a “vote for one” fixed rule. Consider the following voting situation with five voters and four candidates:

# Voters		
2	2	1
A	B	C
D	D	A

B	A	B
C	C	D

Candidate A is the unique Condorcet winner (the majority orderings is  $A >_M B >_M D >_M C$ ), but no fixed-rule voting procedure will *guarantee* that A is elected.

- Using vote for 1 (plurality rule), candidates A and B are tied for the win.
- Using vote for 2, candidate D is elected.
- Using vote for 3, candidates A and B are tied for the win.

Of course, approval voting may elect candidate A (for example, if everyone approves of A and all candidates they rank higher than A). In fact, Brams (2008, Chapter 2) proves that if there is a unique Condorcet winner, then that candidate may be elected under approval voting (assuming that all voters vote *sincerely*: see Brams 2008, Chapter 2, for a discussion). Note that approval voting may also elect other candidates (perhaps even the Condorcet loser).

A number of voting procedures were devised specifically to *guarantee* that a Condorcet winner will be elected, if one exists. I discuss four examples to give a flavor of how such Condorcet consistent procedures work. (See Brams and Fishburn 2002, and Taylor 2005 for more examples.)

#### **Condorcet Rule:**

Each voter submits a linear ordering over all the candidates. If there is a Condorcet winner, then that candidate wins the election. Otherwise, all candidates tie for the win.

#### **Copeland's Rule:**

Each voter submits a linear ordering over all the candidates. A *win-loss record* for candidate B is calculated as follows:

$$WL(B) = \#\{C \mid B >_M C\} - \#\{C \mid C >_M B\}$$

The Copeland winner is the candidate that maximizes WL.

The next method was proposed by Charles Dodgson (better known by the pseudonym Lewis Carroll). Interestingly, this is an example of a procedure in which it is computationally difficult to compute the winner (that is, the problem of calculating the winner is NP-complete). See Bartholdi *et al.* (1989) for a discussion.

#### **Dodgson's Method:**

Each voter submits a linear ordering over all the candidates. For each candidate, determine the fewest number of pairwise swaps needed to

make that candidate the Condorcet winner. The candidate(s) with the fewest swaps is(are) declared the winner(s).

### Black's Procedure:

Each voter submits a linear ordering over all the candidates. If there is a Condorcet winner, then that candidate is the winner. Otherwise, let the winners be the Borda Count winners.

These procedures (and the other Condorcet consistent procedures) guarantee that a Condorcet winner, if one exists, will be elected. But, *should* a Condorcet winner be elected? There are strong intuitions that a Condorcet winner (if one exists) is the candidate that best reflects the will of the voters and that there is something amiss with a voting procedure that does not always elect such a candidate. However, there are arguments against these intuitions. The most persuasive argument comes from the work of Donald Saari (1995, 2001). Consider the following example of 81 voters (this example was originally discussed by Condorcet).

# Voters					
30	1	29	10	10	1
A	A	B	B	C	C
B	C	A	C	A	B
C	B	C	A	B	A

This is another example that shows that Borda's method need not elect the Condorcet winner. The majority ordering is

$$A >_M B >_M C,$$

while the ranking given by the Borda score is

$$B >_{\text{Borda}} A >_{\text{Borda}} C.$$

However, there is an argument that candidate B is the best choice for this electorate. Saari's central observation is to note that the 81 voters can be divided into three groups:

# Voters		
10	10	10
A	B	C

# Voters		
1	1	1
A	C	B

# Voters	
20	28
A	B



B	C	A
C	A	B
Group 1		
C	B	A
B	A	C
Group 2		
B	A	
C	C	
Group 3		

Groups 1 and 2 constitute majority cycles with the voters evenly distributed among the three possible orderings. That is, these groups form a perfect symmetry among the linear orderings. So, within each of these groups, the voters' opinions cancel each other out; therefore, the decision should depend only on the voters in group 3. In group 3, candidate B is the clear winner.

3.2 Failures of Monotonicity

A voting procedure is **monotonic** provided that moving up in the rankings does not adversely affect a candidate's chances to win an election. This property captures the intuition that receiving more support from the voters is always better for a candidate. For example, it is easy to see that plurality rule is monotonic: The more votes a candidate receives, the better chance the candidate has to win. Surprisingly, there are voting methods that do not satisfy this natural property. The most well-known example is plurality with runoff. Consider the two tables below. Note that the only difference between the two tables is the preference orderings of the fourth group of voters. This group of two voters ranks B above A above C in the table on the left and swaps B and A in the table on the right (so, A is now their top-ranked candidate; B is ranked second; and C is still ranked third).

# Voters			
6	5	4	2
A	C	B	B
B	A	C	A
C	B	A	C

# Voters			
6	5	4	2
A	C	B	A
B	A	C	B
C	B	A	C

Candidate A is the plurality-with-runoff winner

Candidate C is the plurality-with-runoff winner

In the election on the left, candidate C, with five votes, is eliminated in the first round. Then, C's votes are all transferred to candidate A, giving her a total of 11 to win the election. However, in the election on the right, even after moving up in the rankings of the fourth group (A is now ranked first by this group), candidate A does *not* win this election. In fact, by trying to give more support to the winner of the election on the left, rather than solidifying A's win, the last

group's least-preferred candidate ended up winning the election! In the election on the right, rather than C being eliminated in the first round, it is candidate B, with only four votes, who is eliminated. Once B is eliminated, candidate C beats candidate A (C receives nine votes while A receives eight).

The above example is surprising since it suggests that, when using plurality with runoff, it may not always be beneficial for a candidate to receive extra votes in the first round. A second example of a failure of monotonicity is the **no-show paradox** of Fishburn and Brams (1983), as the following example illustrates. Suppose that there are three candidates, and the population is divided into the following groups:

# Voters					
417	82	143	357	285	324
A	A	B	B	C	C
B	C	A	C	A	B
C	B	C	A	B	A

In the first round, candidate C wins the election with 609 votes (but this is not an absolute majority); candidate B receives 500 votes and candidate A receives 499 votes. Thus, candidate A is eliminated in the first round. In the second round, 417 votes are transferred to candidate B and 82 votes are transferred to candidate C. Thus, candidate B wins the election with 917 votes (candidate C receives a total of 691 votes). Now, suppose that there are two voters with the ranking  $A > B > C$  who did not take part in the above election. These two voters rank A first, and so, they certainly would prefer that their support for candidate A be taken into account. But, consider what happens when these two voters are added to the population:

# Voters					
419	82	143	357	285	324
A	A	B	B	C	C
B	C	A	C	A	B
C	B	C	A	B	A

In this election, candidate C still wins the first round with 609 votes, but candidate B is eliminated since A now receives 501 votes while B receives only 500 votes. But this means that candidate C wins the election (C receives 966 votes and A receives 644 votes). So, by showing up to the election, these two

extra voters actually caused their least-preferred candidate to win!

### 3.3 Multiple-Districts Paradox

Suppose that a population is divided into districts. If a candidate wins each of the districts, one would expect that candidate to win the election over the entire population of voters. This is certainly true for plurality vote: If a candidate is ranked first by a majority of the voters in in each of the districts, then that candidate will also be ranked first by a majority of voters over the entire population. Interestingly, though, this is not true for plurality rule with runoff, as the following example from Fishburn and Brams (1983) shows.

District 1						District 2					
# Voters						# Voters					
160	0	143	0	0	285	257	82	0	357	285	39
A	A	B	B	C	C	A	A	B	B	C	C
B	C	A	C	A	B	B	C	A	C	A	B
C	B	C	A	B	A	C	B	C	A	B	A

Candidate A wins both districts:

**District 1:** There are a total of 588 voters in this district. Candidate B receives the fewest first-place votes, and so is eliminated in the first round. In the second round, candidate A is now the plurality winner with 303 total votes.

**District 2:** There are a total of 1020 voters in this district. Candidate C receives the fewest first-place votes (324), and so is eliminated in the first round. In the second round, 285 votes are transferred to candidate A and 39 are transferred to candidate C. In the second round, Candidate A is the plurality winner with 644 votes.

However, note that if you combine the two districts, then Candidate B is the winner (the combined districts give us the example discussed above in Section 3.2).

This paradox is an example of a more general phenomenon known as Simpson's

Paradox (Malinas and Bigelow 2009). See Saari (2001, Section 4.2) for a discussion of Simpson's Paradox in the context of voting theory.

### 3.4 The Multiple Elections Paradox

This paradox, first introduced by Brams, Kilgour and Zwicker (1998), has a somewhat different structure from the paradoxes discussed above. Voters are taking part in a *referendum*, where they are asked their opinion directly about various propositions. So, voters must select either “yes” (Y) or “no” (N) for each proposition. Suppose that there are 13 voters who cast the following votes for three propositions (so voters can cast one of eight possible votes):

Propositions	YYY	YYN	YNY	YNN	NYY	NYN	NNY	NNN
# Votes	1	1	1	3	1	3	3	0

When the votes are tallied for each proposition separately, the outcome is N for each proposition (N wins 7–6 for all three propositions). Putting this information together, this means that NNN is the outcome of this election. However, there is *no support* for this outcome in this population of voters.

A similar issue is raised by **Anscombe's paradox** (Anscombe 1976), in which:

It is possible for a majority of voters to be on the losing side of a majority of issues.

This phenomenon is illustrated by the following example with five voters voting on three different issues (the voters either voter ‘yes’ or ‘no’ on the different issues).

	Issue 1	Issue 2	Issue 3
Voter 1	yes	yes	no
Voter 2	no	no	no
Voter 3	no	yes	yes
Voter 4	yes	no	yes
Voter 5	yes	no	yes
Majority	yes	no	yes

However, a majority of the voters (voters 1, 2 and 3) do *not* support the majority outcome on a majority of the issues (note that voter 1 does not support the

majority outcome on issues 2 and 3; voter 2 does not support the majority outcome on issues 1 and 3; and voter 3 does not support the majority outcome on issues 1 and 2)!

The issue is more interesting when the voters do not vote directly on the issues, but on candidates that take positions on the different issues. Suppose there are two candidates A and B who take the following positions on the three issues:

	Issue 1	Issue 2	Issue 3
Candidate A	yes	no	yes
Candidate B	no	yes	no

Candidate A takes the majority position, agreeing with a majority of the voters on each issue, and candidate B takes the opposite, minority position. Under the natural assumption that voters will vote for the candidate who agrees with their position on a majority of the issues, candidate B will win the election (each of the voters 1, 2 and 3 agree with B on two of the three issues, so B wins the election 3-2)! This version of the paradox is known as **Ostrogorski's Paradox** (Ostrogorski 1902). (See Kelly 1989; Rae and Daudt 1976; Wagner 1983, 1984; and Saari 2001, Section 4.6 for analyses of this paradox and Pigozzi 2005, for relationships to judgment aggregation literature.)

## 4. Topics in Voting Theory

### 4.1 Strategizing

In the discussion above, I have assumed that voters select ballots *sincerely*. That is, the voters are simply trying to communicate their opinions about the candidates under the constraints of the chosen voting method. However, in many contexts, voters would rather choose *strategically*. One need only look to recent U.S. elections to see concrete examples of strategic voting. The most often cited example is the 2000 U.S. election: Many voters who ranked third-party candidate Ralph Nader first voted for their second choice (typically Al Gore). A detailed overview of the literature on strategic voting is beyond the scope of this article (see Taylor 2005 for a discussion and pointer to the relevant literature; also see Poundstone 2008 for an entertaining and informative discussion of the occurrence of this phenomenon in many actual elections). I will explain the main issues, focusing on specific voting rules.

In general, there are two general types of manipulation that can be studied in the context of voting. The first is manipulation by a chairman or outside party that has the authority to set the agenda or select the voting method that will be

used. So, the outcome of an election is not manipulated from within by unhappy voters, but, rather, it is *controlled* by an outside authority figure. To illustrate this type of control, consider a population with three voters whose preferences over four candidates are given in the table below:

# Voters		
1	1	1
B	A	C
D	B	A
C	D	B
A	C	D

Note that everyone prefers candidate B over candidate D. Nonetheless, a chairman can ask the right questions so that candidate D ends up being elected. The chairman proceeds as follows: First, ask the voters if they prefer candidate A or candidate B. Since the voters prefer A to B by a margin of two to one, the chairman declares that candidate B is no longer in the running. The chairman then asks voters to choose between candidate A and candidate C. Candidate C wins this election 2–1, so candidate A is removed. Finally, in the last round the chairman asks voters to choose between candidate C and candidate D. Candidate D wins this election 2–1 and is declared the winner.

A second type of manipulation focuses on how the voters themselves can manipulate the outcome of an election by *misrepresenting* their preferences. Consider the following two seven-voter, three-candidate election scenarios:

# Voters		
3	3	1
A	B	C
B	A	A
C	C	B

Election Scenario 1

# Voters		
3	3	1
A	B	C
B	C	A
C	A	B

Election Scenario 2

The only difference between the two scenarios is that the middle group of voters swapped their ordering of their bottom ranked candidates (A and C). In the first election scenario, candidate A is the Borda count winner. However, in the second election scenario, candidate B is the Borda count winner. So, if we assume that scenario 1 represents the “true” preferences of the electorate, it is

in the interest of the middle group to misrepresent their preference and rank C second, followed by A, since the outcome will result in their most-preferred candidate (B) being elected. This is an instance of a general result known as the **Gibbard-Satterthwaite Theorem** (Gibbard 1973; Satterthwaite 1975): Under natural assumptions, there is no voting method that *guarantees* that voters will choose their ballots sincerely (for a precise statement of this theorem and an extensive analysis, see Taylor 2005).

There is a growing literature that characterizes voting methods in terms of how *computationally complex* they are to manipulate. A discussion of this literature is beyond the scope of this article; however, I refer the reader to Bartholdi et al. (1989); Conitzer et al. (2007); Faliszewski and Procaccia (2010); and Faliszewski et al. (2010) for an introduction and pointers to the relevant literature.

## 4.2 Characterization Results

Much of the literature on voting theory (and, more generally, social choice theory) is focused on so-called *axiomatic characterization results*. The main goal here is to characterize different voting methods in terms of abstract normative principles of collective decision making. So, the “axioms” discussed in this literature are intended to describe properties that a group decision method *should* satisfy. It is worth pointing out that this is different from the way a mathematician or logician uses the word “axiom”: To mathematicians or logicians, “axioms” are basic principles that a mathematical theory or logical system *do* satisfy. That is, “axioms” are being used in a *descriptive* sense. (See Endriss 2011, for an interesting discussion of characterization results from a logician's point-of-view.)

I will not attempt to provide a general overview of axiomatic characterizations in social choice theory here (see Gaertner 2006, for an introduction to this vast literature). Rather, I informally discuss a few key axioms and results and how they relate to the voting methods and paradoxes discussed above. I start with three core properties.

- **Anonymity:** The names of the voters do not matter: If two voters change votes, then the outcome of the election is unaffected.
- **Neutrality:** The names of the candidates, or options, do not matter: If two candidates are exchanged in every ranking, then the outcome of the election changes accordingly.
- **Universal Domain:** The voters are free to have any opinion about the candidates. In other words, no preference ordering over the candidates can be ignored by a voting method. Formally, this means that voting methods must be total functions on the space of all profiles (recall that a

profile is a sequence of ballots, one from each voter. Here, I am assuming, as is typical for this literature, that the ballots are the linear orderings over the set of candidates).

These properties ensure that the outcome of an election depends only on the voters' opinions, with all the voters being treated equally. Other properties are intended to rule out some of the paradoxes and anomalies discussed above. In section 4.1, there is an example of a situation in which a candidate is elected, even though *all* the voters prefer a different candidate. The next principle rules out such situations:

- **Unanimity:** If candidate A is preferred to candidate B by *all* voters, then candidate B should not win the election.

Section 3.2 discussed examples in which candidates end up losing an election as a result of more support from some of the voters. Intuitively, a voting procedure is monotonic if moving up in the rankings (all else being equal) should not cause a candidate to lose the election. There are many ways to make this precise. The following strong version (called **Positive Responsiveness** in the literature) is used to characterize majority rule when there are only two candidates:

- **Positive Responsiveness:** If candidate A is tied for the win and moves up in the rankings, then candidate A is the unique winner.

I can now state our first characterization result. Note that in all of the examples above, it is crucial that there are three or more candidates (for example, Condorcet's paradox depends critically on there being three or more candidates). In fact, when there are only two candidates, or options, then majority rule (choose the option with the most votes) can be singled out as "best":

**Theorem** (May 1952).

A social decision method for choosing between two candidates satisfies neutrality, anonymity and positive responsiveness if and only if the method is majority rule.

See May (1952) for a precise statement of this theorem and Asan and Sanver (2002), Maskin (1995), and Woeginger (2003) for generalizations and alternative characterizations of majority rule. With more than two candidates, the most important result is Ken Arrow's celebrated impossibility theorem (1963). Arrow showed that there is no *social welfare function* (a social choice function maps the voters' linear preference orderings to a single social preference ordering) satisfying universal domain, unanimity, non-dictatorship (the social ordering is defined to be the ordering of a single individual) and the following key property:



- **Independence of Irrelevant Alternatives:** The social ranking (higher, lower, or indifferent) of two candidates A and B depends only on the relative rankings of A and B for each individual.

This means that if the voters' rankings of two candidates A and B are the same in two different election scenarios, then the social rankings of A and B must be the same. This is a very strong property that has been extensively criticized (see Gaertner 2006, for pointers to the relevant literature). It is beyond the scope of this article to go into detail about the proof and the ramifications of Arrow's theorem, but I note that many of the voting methods we have discussed do not satisfy the above property. A striking example of a voting method that does not satisfy independence of irrelevant alternatives is Borda count. Consider the following two election scenarios:

# Voters		
3	2	2
A	B	C
B	C	A
C	A	B

Election Scenario 1

# Voters		
3	2	2
A	B	C
B	C	X
C	X	A
X	A	B

Election Scenario 2

Notice that the relative rankings of candidates A, B and C are the same in both election scenarios. In the second scenario, a new (undesirable) candidate is added (i.e., an “irrelevant alternative”). The ranking of the candidates according to their Borda score in scenario 1 puts A first with eight points, B second with seven points and C last with six points. With candidate X in the election (scenario 2), this ranking is reversed: Candidate C is first with 13 voters; candidate B is second with 12 points; candidate A is third with 11 points; and candidate X is last with six points. So, even though the relative rankings of candidates A, B and C do not differ in the two scenarios, the presence of candidate X reverses the Borda rankings.

Finally, I discuss characterizations of all scoring rules (any method that calculates a score based on weights given to different candidates according to where they fall in the ranking; see Section 3.1.1 for a definition) and Approval voting. One defining property of these methods is that they do not suffer from the multiple-districts paradox.

- **Reinforcement:** Suppose that  $N_1$  and  $N_2$  are disjoint sets of voters facing the same set of candidates. Further, suppose that  $W_1$  is the set of winners

for the population  $N_1$ , and  $W_2$  is the set of winners for the population  $N_2$ . If there is at least one candidate that wins both elections, then the winner(s) for the entire population (including voters from both  $N_1$  and  $N_2$ ) is  $W_1 \cap W_2$ .

The reinforcement property explicitly rules out multiple-districts paradoxes (so, candidates that win all sub-elections are guaranteed to win the full election). In order to characterize all scoring rules, one additional technical property is needed:

- **Continuity:** Suppose that a group of voters  $N_1$  elects a candidate A and a disjoint group of voters  $N_2$  elects a different candidate B. Then there must be some number  $m$  such that the population consisting of the subgroup  $N_2$  together with  $m$  copies of  $N_1$  will elect A.

We then have:

**Theorem** (Young 1975).

A social decision method satisfies anonymity, neutrality, reinforcement and continuity if and only if the method is a scoring rule.

This result was generalized by Myerson (1995) by dropping the requirement that voters have linear preferences. Additional axioms have been suggested that single out Borda count among all scoring methods (Young 1974; Nitzan and Rubinstein 1981). In fact, Saari has argued that “any fault or paradox admitted by Borda's method also must be admitted by all other positional voting methods” (Saari 1989, pg. 454). For example, it is often remarked that Borda count (and all scoring rules) can be easily manipulated by the voters. Saari (1995, Section 5.3.1) shows that among all scores rules Borda count is the least susceptible to manipulation (in the sense that it has the fewest profiles where a small percentage of voters can manipulate the outcome).

I conclude this brief discussion of characterization results with Fishburn's characterization of approval voting (see Xu 2010, for an overview of the different characterizations of approval voting).

**Theorem** (Fishburn 1978).

A social decision method is approval voting if and only if the method satisfies anonymity, neutrality, reinforcement and the following technical property:

- If there are exactly two voters who approve of disjoint sets of candidates, then the methods selects as winners all the candidates chosen by the two voters (i.e., the union of the ballots

chosen by the voters).

### 4.3 Voting to Track the Truth

The voting methods discussed above have been judged on *procedural* grounds. This “proceduralist approach to collective decision making” is defined by Coleman and Ferejohn (1986, p. 7) as one that “identifies a set of ideals with which any collective decision-making procedure ought to comply. ... [A] process of collective decision making would be more or less justifiable depending on the extent to which it satisfies them.” The authors add that a distinguishing feature of proceduralism is that “what justifies a [collective] decision-making procedure is strictly a necessary property of the procedure — one entailed by the definition of the procedure alone.” Indeed, the characterization theorems discussed in the previous section can be viewed as an implementation of this idea (cf. Riker 1982). The general view is to analyze voting methods in terms of “fairness criteria” that ensure that a given method is sensitive to *all* of the voters' opinions in the right way.

However, one may not be interested only in whether a collective decision was arrived at “in the right way,” but in whether or not the collective decision is *correct*. This *epistemic* approach to voting is nicely explained by Joshua Cohen (1986):

An epistemic interpretation of voting has three main elements: (1) an independent standard of correct decisions — that is, an account of justice or of the common good that is independent of current consensus and the outcome of votes; (2) a cognitive account of voting — that is, the view that voting expresses beliefs about what the correct policies are according to the independent standard, not personal preferences for policies; and (3) an account of decision making as a process of the adjustment of beliefs, adjustments that are undertaken in part in light of the evidence about the correct answer that is provided by the beliefs of others. (p. 34)

Under this interpretation of voting, a given method is judged on how well it “tracks the truth” of some objective fact (the truth of which is independent of the method being used). A comprehensive comparison of these two approaches to voting touches on a number of issues surrounding the justification of democracy (cf. Christiano 2008); however, I will not focus on these broader issues here. Instead, I briefly discuss an analysis of majority rule that takes this epistemic approach.

The most well-known analysis comes from the writings of Condorcet (1785). The following theorem, which is attributed to Condorcet and was first proved

formally by Laplace, shows that if there are only two options, then majority rule is, in fact, the best procedure from an epistemic point of view. This is interesting because it also shows that a proceduralist analysis and an epistemic analysis both single out majority rules as the “best” voting method when there are only two candidates.

Assume that there are  $n$  voters that have to decide between two alternatives. Exactly one of these alternatives is (objectively) “correct” or “better.” The typical example here is a jury deciding whether or not a defendant is guilty. The two assumptions of the Condorcet jury theorem are:

- **Independence:** The voters' opinions are probabilistically independent (so, the probability that two or more voters are correct is the product of the probability that each individual voter is correct).
- **Voter Competence:** The probability that a voter makes the right decision is greater than  $1/2$ , and this probability is the same for all voters.

See Dietrich (2008) for a critical discussion of these two assumptions. The classic theorem is:

### **Condorcet Jury Theorem.**

Suppose that Independence and Voter Competence are both satisfied. Then, as the group size increases, the probability that the majority chooses the correct option increases and converges to certainty.

See Nitzan (2010) for a modern exposition of this theorem. For a generalization of this theorem beyond two candidates, see Young (1995) and List and Goodin (2001). Conitzer and Sandholm (2005) take these ideas further by classifying different voting methods according to whether or not the methods can be viewed as a *maximum likelihood estimator* (for a noise model).

## **5. Concluding Remarks: from Theory to Practice**

As with any mathematical analysis of social phenomena, questions abound about the “real-life” implications of the theoretical analysis of the voting methods given above. The main difficulty is whether the voting paradoxes are simply features of the formal framework used to represent an election scenario or formalizations of real-life phenomena. This raises a number of subtle issues about the scope of mathematical modeling in the social sciences, many of which fall outside the scope of this article. I conclude with a brief discussion of two questions that shed some light on how one should interpret the above analysis.

### **How *likely* is a Condorcet Paradox or any of the other voting paradoxes?**

There are two ways to approach this question. The first is to calculate the

probability that a majority cycle will occur in an election scenario. There is a sizable literature devoted to analytically deriving the probability of a majority cycle occurring in election scenarios of varying sizes (see Gehrlein 2006, and Regenwetter *et al.* 2006, for overviews of this literature). The calculations depend on assumptions about the distribution of preference orderings among the voters. One distribution that is typically used is the so-called **impartial culture**, where each preference ordering is possible and occurs with equal probability. For example, if there are three candidates, and it is assumed that the voters' preferences are represented by linear orderings, then each linear ordering can occur with probability  $1/6$ . Under this assumption, the probability of a majority cycle occurring has been calculated (see Gehrlein 2006, for details). Riker (1982, p. 122) has a table of the relevant calculations. Two observations about this data: First, as the number of candidates and voters increases, the probability of a majority cycles increases to certainty. Second, for a fixed number of candidates, the probability of a majority cycle still increases, though not necessarily to certainty (the number of voters is the independent variable here). For example, if there are five candidates and seven voters, then the probability of a majority cycle is 21.5 percent. This probability increases to 25.1 percent as the number of voters increases to infinity (keeping the number of candidates fixed) and to 100 percent as the number of candidates increases to infinity (keeping the number of voters fixed). *Prima facie*, this result suggests that we should expect to see instances of the Condorcet and related paradoxes in large elections. Of course, this interpretation takes it for granted that the impartial culture is a realistic assumption. Many authors have noted that the impartial culture is a significant idealization that almost certainly does not occur in real-life elections. Tsetlin et al. (2003) go even further arguing that the impartial culture is a worst-case scenario in the sense that *any* deviation results in lower probabilities of a majority cycle (see Regenwetter *et al.* 2006, for a complete discussion of this issue).

A second way to argue that the above theoretical observations are robust is to find supporting empirical evidence. For instance, is there evidence that majority cycles have occurred in actual elections? While Riker (1982) offers a number of intriguing examples, the most comprehensive analysis of the empirical evidence for majority cycles is provided by Mackie (2003, especially Chapters 14 and 15). The conclusion is that, in striking contrast to the probabilistic analysis referenced above, majority cycles typically have not occurred in actual elections. However, this literature has not reached a consensus about this issue (cf. Riker 1982): The problem is that the available data typically does not include voters' opinions about *all* pairwise comparison of candidates, which is needed to determine if there is a majority cycle. So, this information must be *inferred* (for example, by using statistical methods) from the given data.

**How do the different voting methods compare in actual elections?** In this

article, I have analyzed voting methods under highly idealized assumptions. But, in the end, we are interested in a very practical question: Which method should a given society adopt? Of course, any answer to this question will depend on many factors that go beyond the abstract analysis given above. An interesting line of research focuses on incorporating *empirical evidence* into the general theory of voting. Evidence can come in the form of a computer simulation, a detailed analysis of a particular voting method in real-life elections (for example, see Brams 2008, Chapter 1, which analyzes Approval voting in practice), or as *in situ* experiments in which voters are asked to fill in additional ballots during an actual election (Laslier 2010, 2011).

However, the most striking results here can be found in the work of Michael Regenwetter and his colleagues. They have analyzed datasets from a variety of elections, showing that many of the usual voting methods that are considered irreconcilable (e.g., plurality, Borda count and methods that choose the Condorcet winner) are, in fact, in perfect agreement. This suggests that the “theoretical literature may promote overly pessimistic views about the likelihood of consensus among consensus methods” (Regenwetter *et al.* 2009, p. 840). See Regenwetter *et al.* 2006 for an introduction to the methods used in these analyses and (Regenwetter *et al.* 2009) for the current state-of-the-art.

## Bibliography

- Anscombe, G. E. M., 1976, “On Frustration of the Majority by Fulfillment of the Majority's Will,” *Analysis*, 36(4): 161–168.
- Arrow, K., 1963, *Social Choice and Individual Values*, New Haven: Yale University Press, 2nd edition.
- Asan, G. and R. Sanver, 2002, “Another Characterization of the Majority Rule,” *Economics Letters*, 75(3): 409–413.
- Balinski, M. and R. Laraki, 2007, “A theory of measuring, electing and ranking,” *Proceeding of the National Academy of Sciences*, 104(21): 8720–8725.
- , 2010, *Majority Judgement: Measuring, Ranking and Electing*, Boston: MIT Press.
- Bartholdi III, J.J., C. A. Tovey, and M.A. Trick, 1989, “The Computational Difficulty of Manipulating an Election,” *Social Choice and Welfare*, 6(3): 227–241.
- , 1989, “Voting schemes for which it can be difficult to tell who won the election,” *Social Choice and Welfare*, 6(2): 157–165.
- Borda, J.-C. de, 1784, “Mémoire sur les élections au scrutin par M. de Borda” in *Mémoires de l'Académie Royale des Sciences année 1781*, Paris: l'Imprimerie Royale, pp. 657–665; Translated in McClean and Urken 1995, pp. 83–89.
- Brams, S., 2008, *Mathematics and Democracy*, Princeton: Princeton University

Press.

Brams, S. and P. Fishburn, 2007 (2nd Edition), *Approval Voting*, New York: Springer.

---, 2002, "Voting Procedures", in *Handbook of Social Choice and Welfare*, K. J. Arrow, A. K. Sen, and K. Suzumura (eds.), Amsterdam: Elsevier, pp. 173-236.

Brams, S., D. M. Kilgour D. M., and W. Zwicker, 1998, "The paradox of multiple elections," *Social Choice and Welfare*, 15(2): 211-236.

Brams, S. and Sanver, M. R., "Voting Systems That Combine Approval and Preference," in *The Mathematics of Preference, Choice, and Order: Essays in Honor of Peter C. Fishburn*, S. Brams, W. Gehrlein, and F. Roberts (eds.), Berlin: Springer, pp. 215-237.

Christiano, T., "Democracy", *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2008/entries/democracy/>>.

Cohen, J., 1986, "An epistemic conception of democracy," *Ethics*, 97(1): 26-38.

Coleman, J. and J. Ferejohn, 1986, "Democracy and social choice," *Ethics*, 97(1): 6-25.

Condorcet, M.J.A.N. de C., Marque de, 1785, *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Paris: l'Imprimerie Royale; Translated in Mclean and Urken 1995, pp. 91-113.

Conitzer, V., T. Sandholm, and J. Lang, 2007, "When are Elections with Few Candidates Hard to Manipulate?," *Journal of the ACM*, 54(3): Article 14.

Conitzer, V. and T. Sandholm, 2005, "Common Voting Rules as Maximum Likelihood Estimators" in *Proceedings of the 21st Annual Conference on Uncertainty in Artificial Intelligence* (UAI-05), pp. 145-152.

Daudt, H. and D. W. Rae, 1976, "The Ostrogorski paradox: a peculiarity of compound majority decision," *European Journal of Political Research*, 4(4): 391-399.

Dietrich, F., 2008, "The premises of Condorcet's jury theorem are not simultaneously justified," *Episteme - a Journal of Social Epistemology*, 5(1): 56-73.

Dowding, K. and M. Van Hees, 2007, "In Praise of Manipulation," *British Journal of Political Science*, 38(1): 1-15.

Endriss, U., 2011, "Logic and Social Choice Theory," in *Logic and Philosophy Today*, J. van Benthem and A. Gupta (eds.), London: College Publications.

Faliszewski, P. and A. Procaccia, 2010, "AI's War on Manipulation: Are We Winning?," *AI Magazine*, 31(4): 53-64.

Faliszewski, P., E. Hemaspaandra, and L. Hemaspaandra, 2010, "Using complexity to protect elections," *Communications of the ACM*, 53(11): 74-82.

Felsenthal, D. and M. Machover, 1998, *The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes*, Cheltenham Glos: Edward Elgar Publishing.

- Fishburn, P. and S. Brams, 1983, "Paradoxes of Preferential Voting," *Mathematics Magazine*, 56(4): 207-214.
- Fishburn, P., 1974, "Paradoxes of Voting," *American Political Science Review*, 68(2): 537-546.
- , 1978, "Axioms for Approval Voting: Direct Proof," *Journal of Economic Theory*, 19(1): 180-185.
- Gaertner, W., 2006, *A Primer in Social Choice Theory*, Oxford: Oxford University Press.
- Gehrlein, W., 2006, *Condorcet's Paradox*, Berlin: Springer.
- Gibbard, A., 1973, "Manipulation of voting schemes: A general result," *Econometrica*, 41(4): 587-601.
- Hansson, S. O. and Grüne-Yanoff, T., "Preferences", *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2009/entries/preferences/>.
- Kelly, J.S., 1989, "The Ostrogorski's paradox," *Social Choice and Welfare*, 6(1): 71-76.
- Laslier, J.-F., 2011, "Lessons from *in situ* experiments during French elections," in *In Situ and Laboratory Experiments on Electoral Law Reform: French Presidential Elections*, B. Dolez, B. Grofman and A. Laurent (eds.), Berlin: Springer, pp. 91-104.
- , 2010, "Laboratory experiments about approval voting" in *Handbook of Approval Voting*, J.-F. Laslier and R. Sanver (eds.), Berlin: Springer, pp. 339-356.
- Laslier, J.-F. and R. Sanver (eds.), 2010, *Handbook on Approval Voting*, Series: Studies in Choice and Welfare, Berlin: Springer.
- Levin, J. and B. Naelbuff, 1995, "An introduction to vote-counting schemes," *Journal of Economic Perspectives*, 9(1): 3-26.
- List, C., 2006, "The Discursive Dilemma and Public Reason", *Ethics*, 116(2): 362-402.
- List, C. and R. Goodin, 2001, "Epistemic Democracy: Generalizing the Condorcet Jury Theorem," *Journal of Political Philosophy*, 9(3): 277-306.
- Mackie, G., 2003, *Democracy Defended*, Cambridge: Cambridge University Press.
- Malinas, G. and J. Bigelow, "Simpson's Paradox", *The Stanford Encyclopedia of Philosophy* (Fall 2009 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2009/entries/paradox-simpson/>.
- Maskin, E., 1995, "Majority rule, social welfare functions and game forms," in *Choice, Welfare and Development: A Festschrift in Honour of Amartya K. Sen*, K. Basu, P. Pattanaik, K. Suzumura (eds.), Oxford: Oxford University Press, pp. 100-109.
- May, K., 1952, "A set of independent necessary and sufficient conditions for simply majority decision," *Econometrica*, 20(4): 680-684.
- McClellan, I. and A. Urken (eds.), 1995, *Classics of Social Choice*, Ann Arbor: The University of Michigan Press.





- Myerson, R., 1995, "Axiomatic derivation of scoring rules without the ordering assumption," *Social Choice and Welfare*, 12(1): 59-74.
- Nitzan, S., 2010, *Collective Preference and Choice*, Cambridge: Cambridge University Press.
- Nitzan, S. and A. Rubinstein, 1981, "A further characterization of Borda ranking method," *Public Choice*, 36(1): 153-158.
- Nurmi, H., 1987, *Comparing Voting Systems*, Dordrecht: D. Reidel.
- , 1999, *Voting Paradoxes and How to Deal with Them*, Berlin: Springer-Verlag.
- , 2010, "Voting Theory," in *e-Democracy: A Group Decision and Negotiation Perspective*, D. R. Insua and S. French (eds.), Berlin: Springer, pp. 101-124.
- Ostrogorski, M., 1902, *Democracy and the Organization of Political Parties*, London: Macmillan.
- Pigozzi, G., 2005, "Two aggregation paradoxes in social decision making: the Ostrogorski paradox and the discursive dilemma," *Episteme: A Journal of Social Epistemology*, 2(2): 33-42.
- Poundstone, W., 2008, *Gaming the Vote: Why Elections aren't Fair (and What We Can Do About It)*, New York: Hill and Wang Press.
- Regenwetter, M., B. Grofman, A.A.J. Marley, A.A.J. and I. Tsetlin, 2006, *Behavioral Social Choice: Probabilistic Models, Statistical Inference, and Applications*, Cambridge: Cambridge University Press.
- Regenwetter, M., B. Grofman, A. Popova, W. Messner, C. Davis-Stober, and D. Cavagnaro, 2009, "Behavioural Social Choice: A Status Report", *Philosophical Transactions of the Royal Society B*, 364(1518): 833-843.
- Riker, W., 1982, *Liberalism against Populism: A Confrontation between the Theory of Democracy and the Theory of Social Choice*, San Francisco: W. H. Freeman & Co.
- Saari, D., 1989, "A dictionary of voting paradoxes," *Journal of Economic Theory*, 48(2): 443-475.
- , 1995, *Basic Geometry of Voting*, Berlin: Springer.
- , 2001, *Decisions and Elections: Explaining the Unexpected*, Cambridge: Cambridge University Press.
- , 2000, "Mathematical Structure of Voting Paradoxes: II. Positional Voting," *Economic Theory*, 15(1): 55-102.
- Satterthwaite, M., 1975, "Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions," *Journal of Economic Theory*, 10(2): 198-217.
- Taylor, A., 2005, *Social Choice and the Mathematics of Manipulation*, Cambridge: Cambridge University Press.
- Tsetlin, I., M. Regenwetter, and B. Grofman, 2003, "The impartial culture maximizes the probability of majority cycles," *Social Choice and Welfare*, 21(3): 387-398.
- Wagner, C., 1983, "Anscombe's Paradox and the Rule of Three-Fourths," *Theory and Decision*, 15(3): 303-308.

- , 1984, "Avoiding Anscombe's Paradox," *Theory and Decision*, Volume 16(3): 233-238.
- Woeginger, G., 2003, "A new characterization of the majority rule", *Economic Letters*, 81(1): 89-94.
- Young, H.P., 1995, "Optimal Voting Rules," *Journal of Economic Perspectives*, 9(1): 51-64.
- , 1975, "Social Choice Scoring Functions," *SIAM Journal of Applied Mathematics*, 28(4): 824-838.

## Academic Tools

 [How to cite this entry.](#)

 [Preview the PDF version of this entry](#) at the [Friends of the SEP Society](#).

 [Look up this entry topic](#) at the [Indiana Philosophy Ontology Project](#) (InPhO).

 [Enhanced bibliography for this entry](#) at [PhilPapers](#), with links to its database.

## Other Internet Resources

- Three videos created by Donald Saari for the 2008 Mathematics Awareness Month which had the theme "mathematics of voting":
  1. [Mathematics of Voting](#)
  2. [Creating Voting Paradoxes](#)
  3. [What Causes Voting Paradoxes](#)

## Related Entries

[democracy](#) | [preferences](#) | [social choice theory](#)

## Acknowledgements

I would like to thank Ulle Endriss, Uri Nodelman, Rohit Parikh, Edward Zalta and two anonymous referees for many valuable comments that greatly improved the readability and content of this article. This article was written while the author was generously supported by an NWO Vidi grant 016.094.345.

[Copyright © 2011](#) by  
[Eric Pacuit](#) <[E.J.Pacuit@uvt.nl](mailto:E.J.Pacuit@uvt.nl)>

Open access to the Encyclopedia has been made possible, in part, with a

financial contribution from the University of Illinois at Urbana-Champaign Library. We gratefully acknowledge this support.

**Stanford** | Center for the Study of  
Language and Information

The Stanford Encyclopedia of Philosophy is [copyright © 2014](#) by [The Metaphysics Research Lab](#), Center for the Study of Language and Information (CSLI), Stanford University

Library of Congress Catalog Data: ISSN 1095-5054