

# Perspectives on Preference Aggregation

Michel Regenwetter

University of Illinois at Urbana-Champaign

**ABSTRACT**—*For centuries, the mathematical aggregation of preferences by groups, organizations, or society itself has received keen interdisciplinary attention. Extensive theoretical work in economics and political science throughout the second half of the 20th century has highlighted the idea that competing notions of rational social choice intrinsically contradict each other. This has led some researchers to consider coherent democratic decision making to be a mathematical impossibility. Recent empirical work in psychology qualifies that view. This nontechnical review sketches a quantitative research paradigm for the behavioral investigation of mathematical social choice rules on real ballots, experimental choices, or attitudinal survey data. The article poses a series of open questions. Some classical work sometimes makes assumptions about voter preferences that are descriptively invalid. Do such technical assumptions lead the theory astray? How can empirical work inform the formulation of meaningful theoretical primitives? Classical “impossibility results” leverage the fact that certain desirable mathematical properties logically cannot hold in all conceivable electorates. Do these properties nonetheless hold true in empirical distributions of preferences? Will future behavioral analyses continue to contradict the expectations of established theory? Under what conditions do competing consensus methods yield identical outcomes and why do they do so?*

Social choice procedures (i.e., voting or consensus methods) have a long tradition in economics and political science, but they have also received limited attention in psychological science. Economics and political science have perfected the mathematical analysis of “what is possible” and “what is not universally possible.” Scientists in these fields have highlighted a number of voting paradoxes and impossibility theorems. This

Address correspondence to Michel Regenwetter, Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, IL 61820; e-mail: regenwet@illinois.edu.

mathematically sophisticated literature has drawn a complex picture. Two essential and widely endorsed conclusions can be simply paraphrased as follows: (a) Competing notions of rational social choice appear intrinsically irreconcilable, and (b) universal democratic decision making, based purely on aggregation of voter preferences, may be a mathematical impossibility.

Theoretical and empirical work in psychological science takes a more optimistic perspective. Psychologists have found little empirical evidence for voting paradoxes and, so far, little behavioral support for the famed incompatibility of social choice methods. These findings suggest that technical assumptions or theoretical problem formulations in social choice theory might bias our understanding. I discuss a number of major open theoretical and empirical questions about how to reconcile rational social choice theory with empirical data. Resolving these questions requires the interdisciplinary development of a comprehensive descriptive theory of social choice behavior, with psychological science plausibly in a leadership role.

Social choice procedures are mathematical tools. Here, I introduce key mathematical notions via examples, not formal definitions. For full-fledged technical treatments, see Regenwetter, Grofman, Marley, and Tsetlin (2006) and its predecessor papers or Regenwetter, Ho, and Tsetlin (2007) and Regenwetter and Rykhlevskaia (2007).

## MATHEMATICAL REPRESENTATION OF INDIVIDUAL PREFERENCES

Throughout this article, our mathematical representation of individual preferences consists of *preference relations*. Consider four choice alternatives— $A$ ,  $B$ ,  $C$ , and  $D$ —and suppose that a decision maker strictly prefers  $A$  to  $B$ , strictly prefers  $B$  to  $C$ , and strictly prefers  $A$  to  $C$ , but has no other strict preferences. I will denote such a decision maker’s preference as a preference relation,  $\mathcal{R}$ —namely as a collection of ordered pairs:  $\mathcal{R} = \{(A, B), (B, C), (A, C)\}$ . For simplicity, I only consider strict preference: If a person prefers  $X$  to  $Y$ , then they cannot also prefer  $Y$  to  $X$ . In the example above, the decision maker has no strict preference among  $C$  and  $D$ . This is captured by the fact that  $\mathcal{R}$  contains neither  $(C, D)$  nor  $(D, C)$ .

Consider  $n$  choice alternatives—I refer to these also as objects, candidates, competitors, or options. For any preference relation  $\mathcal{R}$ , one can, for each object  $X$ , derive a (generalized) rank of  $X$  in  $\mathcal{R}$  as follows. First, start with the number  $n + 1$  and add to it the number of choice alternatives that are strictly preferred to  $X$  in  $\mathcal{R}$ . Now, subtract how many choice alternatives  $X$  is strictly preferred to in  $\mathcal{R}$ . The (generalized) rank is this quantity divided by two (Regenwetter & Rykhlevskaia, 2004). For example, if there are  $n = 3$  choice alternatives,  $A$ ,  $B$ , and  $C$ , and the preference relation is the collection  $\mathcal{R} = \{(A, B), (B, C), (A, C)\}$ , then  $A$  has rank  $\frac{(3+1)+0-2}{2} = 1$ ,  $B$  has rank  $\frac{(3+1)+1-1}{2} = 2$ , and  $C$  has rank  $\frac{(3+1)+2-0}{2} = 3$ . An important feature of the (generalized) rank is that it allows assignment of numerical ranks to all choice alternatives even in cases in which decision makers lack strict preferences among some alternatives or have incoherent (e.g., “intransitive”) preferences. Multiple options may share the same rank, but only one option, at most, can have Rank 1, meaning that the individual strictly prefers this option to all alternatives. Only one option can have rank  $n$ , and this is precisely the case if the decision maker strictly prefers every other option to this option. Ranks are multiples of  $\frac{1}{2}$ .

## SIX SOCIAL CHOICE METHODS

I now review six social choice methods that have received varying degrees of scientific attention. Throughout, assume  $N$  decision makers who have preference relations  $\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_N$ , respectively. The first method implements pairwise contests among choice alternatives.

### Condorcet Criterion

According to this criterion, choice alternative  $X$  is socially preferred to choice alternative  $Y$  if more people strictly prefer  $X$  to  $Y$  than  $Y$  to  $X$ —that is, if the number of relations containing  $(X, Y)$  exceeds the number of relations containing  $(Y, X)$ . A candidate is the Condorcet winner if it beats each competitor in this fashion. The *Condorcet paradox* denotes the fact that a Condorcet winner need not exist. Although, arguably, the Condorcet criterion is generally viewed as the most natural implementation of democratic aggregation, a substantial theoretical literature has predicted an omnipresent threat of the Condorcet paradox, and leading textbooks strongly caution against using this method (e.g., Shepsle & Bonchek, 1997). Condorcet is rarely used in multicandidate elections. The next three methods I examine are called “scoring rules” (or “positional voting methods”) because they assign overall scores to candidates based on their ranks in the individual preference relations.

### Plurality

Each candidate receives a point from each individual preference relation in which this candidate’s rank equals 1. Preference relations are discarded if no option has Rank 1. In practice, this

usually amounts to the world’s most common voting method: Each valid ballot contains one vote for one candidate (which is counted as having Rank 1). Candidates are socially ordered by decreasing total score.

### Borda

For each preference relation, each option scores  $n$  minus its rank in that relation. The social order sorts the options from highest to lowest total Borda Score. Equivalently one can compute the arithmetic mean of each candidate’s ranks across the given collection of preference relations and socially order the alternatives by increasing mean rank. The Borda Score has been the main conceptual competitor of the Condorcet criterion for over two centuries. Like Condorcet, it is rarely used.

### Antiplurality

Each option receives a penalty point from each preference relation in which its rank equals  $n$ . Preference relations are discarded if they do not allocate a rank of  $n$  to anyone. In practice, this usually amounts to the rarely used voting procedure in which each valid ballot contains a single vote against a single candidate. The aggregate order is by increasing total number of penalty points.

A major theoretical critique of scoring rules is that the social preference among two choice alternatives could be altered through the introduction (or removal) of third candidates. The following iterative single-seat election method is heavily used around the world. It also appears to be the most promoted (and frequently adopted) method for electoral reform for all levels and types of government in the United States.

### Alternative Vote, Hare System, Instant Runoff

To be elected, a candidate must have Rank 1 on more than 50% of all ballots. As long as no candidate meets this quota, the candidates with the least Rank 1 votes are sequentially eliminated, and all ranks are recomputed on the ballots until a candidate meets the quota. Notably, the American Psychological Association uses this method to elect its president.

### Median Generalized Rank

This final method is based on statistical measures of central tendency. For each candidate, one determines the median of its ranks across preference relations, and socially orders the choice alternatives according to the magnitude of that median rank.

## SOME ESTABLISHED WISDOM IN SOCIAL CHOICE THEORY

In this section, I focus on three major concerns of social choice theory: impossibility theorems, the Condorcet paradox, and the concept of Condorcet efficiency.

The famous impossibility theorem of Nobel Laureate Kenneth Arrow (1951) shows that, logically, a social choice procedure that satisfies a certain set of prerequisite conditions for rational aggregation of preferences cannot exist. One of its conditions forbids outcomes like the Condorcet paradox. Another requirement of Arrow's theorem is that any acceptable method must satisfy the list of conditions in all conceivable distributions of preferences, whether they occur in real life or not. A number of other impossibility results in the theoretical literature are of a similar nature.

Riker (1982) provided a summary table for extensive prior research on the likelihood of the Condorcet paradox. According to that table, the paradox should be omnipresent, especially for large numbers of candidates and large numbers of voters. Riker assumed that every decision maker's preference translates into a "complete" ranking (i.e., no two options have equal rank). The table makes the *impartial culture assumption*, according to which every electorate is a random sample from a uniform distribution over such complete rankings. Social choice theorists readily acknowledge that the real world looks nothing like an impartial culture. Yet, Riker's table is standard fare in basic courses on social choice.

Despite the presumed threat of the Condorcet paradox, the Condorcet criterion continues to be appealing. The Condorcet efficiency of, say, the Borda score, is the conditional probability that the Borda winner matches the Condorcet winner, given that a Condorcet winner exists in samples drawn typically from an impartial culture. The overall picture painted by this literature reinforces the literature on impossibility theorems: Competing social choice procedures are based on logically incompatible principles, and, under standard theoretical assumptions, they cannot be expected to yield mutually consistent outcomes.

Saari (1999, 2000a, 2000b, 2001) has developed a sophisticated mathematical framework with the potential to construct preference distributions that yield nearly any prespecified pattern of disagreements among a range of social choice methods, as long as such a disagreement is mathematically possible.

## BEHAVIORAL SOCIAL CHOICE: EMPIRICAL FINDINGS

I now turn to the behavioral work on the Condorcet paradox and the consensus among consensus methods.

In previous research (Regenwetter et al., 2006), my colleagues and I developed a methodological framework for behavioral social choice in which the mathematical properties of social choice procedures can be evaluated against empirical behavioral data. The original book and its predecessor papers placed a major emphasis on the empirical hunt for the Condorcet paradox. In Regenwetter et al. (2006), and subsequently Regenwetter, Ho, and Tsetlin (2007) and Regenwetter, Kim, Kantor, and Ho (2007), we considered a variety of real and experimental ballot and national election survey data sets. We found virtually no evidence for the empirical occurrence of the

paradox. The only cases where the paradox could not be ruled out were cases that lacked statistical identifiability or statistical replicability. List and Goodin (2001), Regenwetter et al. (2006), Tangian (2000) and others also discussed theoretical explanations for the absence of the paradox. Most notably, the predictions in Riker's famous (1982) table completely collapse as soon as the underlying theoretical assumptions deviate ever so slightly from the impartial culture. Others, including Feld and Grofman (1992) and Mackie (2003), have previously questioned from a variety of angles whether the famed prevalence of the Condorcet paradox is empirically valid. Dryzek and List (2003) and List, Luskin, Fishkin, and McLean (2007) suggest that deliberation among decision makers may help to eliminate the paradox.

In Regenwetter et al. (2006) and its predecessor papers, we compared the two most important historical competitors, Condorcet and Borda, using a variety of techniques to infer preference distributions from approval voting ballots and from national election survey data of three countries. In Regenwetter, Ho, and Tsetlin (2007), we inferred Condorcet, Borda, and plurality outcomes from approval voting ballots using more sophisticated inference techniques. Finally, in Regenwetter et al. (2007) we compared Condorcet, Borda, plurality, and the alternative vote using four sets of American Psychological Association ballots that used the alternative vote. In all cases, we found striking agreements between rival social choice methods. Tangian (2000) discussed theoretical situations in which Condorcet and Borda will tend to agree.

In important pioneering work using 37 election data sets, Felsenthal, Maoz, and Rapoport (1993) concluded that different versions of multiseat plurality, Borda, the alternative vote, and two other methods were very similar. They reported that a type of plurality method showed the greatest signs for violations of rationality principles. More recently, in psychology, Hastie and Kameda (2005) studied nine social choice rules, including Condorcet, Borda, plurality, an averaging rule, and a median rule. These authors used a simulated test bed of a hunter-gatherer society that had to carry out a signal-detection-like task. They found high agreement of all methods with the Condorcet criterion. For instance, in 12-member groups, six out of eight methods agreed with the Condorcet criterion regarding the winner, with probability exceeding .85.

## BEHAVIORAL SOCIAL CHOICE: METHODOLOGICAL ISSUES

Several methodological issues complicate theoretical and empirical social choice research. I review three problems: model dependence, statistical replicability, and the role of strategic behavior.

Theoretical assumptions can have profound implications on social choice results. For instance, we have seen that deviating ever so slightly from the impartial culture assumption can

literally reverse theoretical predictions. Like rational choice theory, behavioral social choice also requires theoretical modeling and often technical assumptions. For instance, inferences from incomplete or imperfect data are usually only possible with the help of some technical assumptions. These assumptions, should they be invalid, could distort empirical findings. It is therefore crucial to analyze the same data from multiple and fundamentally different angles and to investigate the extent to which modeling assumptions may impact substantive conclusions. In past research, several colleagues and I systematically varied our assumptions and concluded that, although inferred social orders sometimes changed with assumptions, the empirical absence of the Condorcet paradox and the amazing congruence among social choice methods carried through across their different approaches (Regenwetter et al., 2006; Regenwetter, Ho, & Tsetlin, 2007; Regenwetter, Kim, et al., 2007).

Theoretical social choice research in economics and political science usually treats ballots as deterministic and flawless representations of the decision makers' intended votes. The Florida recount of the 2000 U.S. presidential election has invalidated this assumption on a large scale. In Regenwetter et al. (2006), Regenwetter, Ho, and Tsetlin (2007), Regenwetter and Rykhlevskaia (2007), and Regenwetter and Tsetlin (2004), we motivated and developed maximum likelihood and Bayesian statistical inference tools for probabilistic (random sample) ballot or survey data. In Regenwetter, Kim, et al. (2007), we used bootstrap methods for evaluating the statistical replicability ("confidence") of inferred social choice outcomes.

A caveat with early empirical work is that data sets tended to be sparser (tens to hundreds of observations) than are some recently investigated data sets (tens of thousands of observations). Inferences made from sparse data could be sensitive to small perturbations in the data (i.e., statistical replicability may be low).

It remains extremely difficult to formulate statistically testable models with identifiable parameters in a way that allows one to infer preference distributions from empirical data when decision makers distort their preferences strategically. In Regenwetter, Ho, and Tsetlin (2007), we developed such a method for approval voting data under the assumption that individual preferences satisfy a general type of Thurstone model (Thurstone, 1927). We provided statistical evidence that decision makers were not voting strategically. Much of the other behavioral work cited in this article relies on assumptions about sincere expression of voter preferences. A major challenge for future work is to incorporate strategic components into other models without requiring overly artificial subsidiary technical assumptions when dealing with survey or ballot data.

#### AN ILLUSTRATIVE BEHAVIORAL ANALYSIS

Following the approach of Regenwetter, Kim, et al. (2007), I provide a brief empirical illustration on American Psychological

Association presidential election ballots. The 2004 election featured five presidential candidates and 16,245 voters. If the social orders were to disagree, the huge number of observations should give us ample statistical power to determine such disagreements with ease. For purposes of brevity, I have treated the ballots as so-called "weak orders" from which one can readily compute the Condorcet, plurality, Borda, antiplurality, alternative vote, and median generalized rank outcomes using the definitions I gave earlier in the article. The alternative vote elects one option—it does not socially rank the candidates. I report the degree of agreement among Condorcet, Plurality, Borda, Antiplurality, Alternative Vote, and Median Generalized Rank as to who the winner is and the degree of agreement among Condorcet, Plurality, Borda, Antiplurality, and Median Generalized Rank on who the loser is. By "winner(s)" I mean the candidate(s) with the lowest generalized rank in the social order, and by "loser" I mean a (unique) candidate, if there is one, who has the generalized Rank 5 in the social order.

We bootstrapped 1,000 samples ( $N = 16,245$  for each sample, with replacement) from the original data. All 1,000 of these samples yielded identical and unique winners for Condorcet, plurality, Borda, antiplurality, and the alternative vote. In 895 samples, this winner was further matched by a unique and identical winner under the median generalized rank system. In the remaining 105 samples, the median yielded additional, tied winners. None of these 1,000 samples yielded a Condorcet paradox. All 1,000 samples yielded identical losers across all five methods. I omit alternative models and methods for brevity.

Is this perfect consensus among archival consensus methods surprising? The social choice literature on the impartial culture would draw samples with replacement from a uniform distribution, rather than an empirical distribution. To circumnavigate the problem of ties in the median generalized rank, consider the winners under Condorcet, plurality, Borda, antiplurality, and the alternative vote and the losers under Condorcet, plurality, Borda, and antiplurality. Although these outcomes matched 100% in our bootstrap analysis, simulated samples from an impartial culture led to such an agreement 3% of the time. In samples from a uniform distribution over allowable APA ballots (using the weak-order approach), we also found such agreement only 3% of the time. The standard theoretical benchmarking methods suggest extreme disagreement among methods: Out of 1,000 simulated draws from an impartial culture, 25% yielded outcomes in which a candidate had Rank 1 in the social order of one method and Rank 5 in the social order of another method. For a uniform distribution over allowable APA ballots, the corresponding proportion was 33%. In our behavioral analysis, such disagreements occurred not even once in 1,000 draws!

#### PERSPECTIVES ON PREFERENCE AGGREGATION

The apparent rarity of the Condorcet paradox and the apparent empirical consensus among mutually incompatible social

choice procedures call for a descriptive theory of behavioral social choice that provides explanations for these puzzling findings. Countless interesting open questions regarding these two important phenomena remain: When and why are social choice procedures empirically congruent? Does it matter whether we study preferential data reflecting personal tastes, in which there may be no objective “right answer,” versus judgmental data, in which objective optimal choices exist? For judgmental data, what is the role of ecological cues for the consensus among consensus methods? Similarly, what are the roles of group deliberation, group size, persuasion campaigns, and strategic behavior? To what extent do properties of the choice alternatives drive the agreement among social choice rules, and to what extent may institutions play a role? What role do computational complexity, as well as heuristics and biases in information processing, play? What types of decision environments foster consensus among consensus methods and thus permit uncontestedly “democratic” aggregation? How does the practical nonmanipulability (Tideman, 2006) affect the empirical agreement among competing methods? Although many of these cognitive, social, and political processes have been studied extensively in other contexts, future research should synergize these fields to make better sense of social choice behavior.

Finally, a large open question concerns the exact policy implications of this research paradigm if, indeed, widespread consensus among consensus methods is validated and begins to be explained.

**Acknowledgments**—The author wrote this article as a guest of Sonderforschungsbereich 504 at the University of Mannheim, Germany, and of the Center for Adaptive Behavior and Cognition at the Max Planck Institute for Human Development in Berlin, Germany. The empirical findings were prepared by Anna Popova, with funding from the Air Force Office of Scientific Research, Cognition, and Decision Program, under Award No. FA9550-05-1-0356 (to Michel Regenwetter). Thanks to the American Psychological Association for access to their ballot data and to Anna Popova, Christian List, Clinton Davis-Stober, Gerry Mackie, Scott Feld, Tatsuya Kameda, and William Messner for critical comments on earlier drafts. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of his colleagues, the University of Mannheim, the Max Planck Society, or the Air Force Office of Scientific Research.

## REFERENCES

- Arrow, K.J. (1951). *Social choice and individual values*. New York: Wiley.
- Dryzek, J., & List, C. (2003). Social choice theory and deliberative democracy: A reconciliation. *British Journal of Political Science*, 33, 1–28.
- Feld, S.L., & Grofman, B. (1992). Who is afraid of the big bad cycle? Evidence from 36 elections. *Journal of Theoretical Politics*, 4, 231–237.
- Felsenthal, D.S., Maoz, Z., & Rapoport, A. (1993). An empirical evaluation of 6 voting procedures: Do they really make any difference? *British Journal of Political Science*, 23, 1–27.
- Hastie, R., & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review*, 112, 494–508.
- List, C., & Goodin, R.E. (2001). Epistemic democracy: Generalizing the Condorcet jury theorem. *Journal of Political Philosophy*, 9, 277–306.
- List, C., Luskin, R.C., Fishkin, J.S., & McLean, I. (2007). *Deliberation, single-peakedness, and the possibility of meaningful democracy: Evidence from deliberative polls*. Manuscript in preparation, London School of Economics and Stanford University, CA.
- Mackie, G. (2003). *Democracy defined*. New York: Cambridge University Press.
- Regenwetter, M., Grofman, B., Marley, A., & Tsetlin, I. (2006). *Behavioral social choice*. Cambridge, United Kingdom: Cambridge University Press.
- Regenwetter, M., Ho, M.-H., & Tsetlin, I. (2007). Sophisticated approval voting, ignorance priors, and plurality heuristics: A behavioral social choice analysis in a Thurstonian framework. *Psychological Review*, 114, 994–1014.
- Regenwetter, M., Kim, A., Kantor, A., & Ho, M.-H. (2007). The unexpected empirical consensus among consensus methods. *Psychological Science*, 18, 559–656.
- Regenwetter, M., & Rykhlevskaia, E. (2004). On the (numerical) ranking associated with any finite binary relation. *Journal of Mathematical Psychology*, 48, 239–246.
- Regenwetter, M., & Rykhlevskaia, E. (2007). A general concept of scoring rules: General definitions, statistical inference, and empirical illustrations. *Social Choice and Welfare*, 29, 211–228.
- Regenwetter, M., & Tsetlin, I. (2004). Approval voting and positional voting methods: Inference, relationship, examples. *Social Choice and Welfare*, 22, 539–566.
- Riker, W.H. (1982). *Liberalism versus populism*. San Francisco: W.H. Freeman.
- Saari, D.G. (1999). Explaining all three-alternative voting outcomes. *Journal of Economic Theory*, 87, 313–355.
- Saari, D.G. (2000a). Mathematical structure of voting paradoxes: I. Pairwise vote. *Economic Theory*, 15, 1–53.
- Saari, D.G. (2000b). Mathematical structure of voting paradoxes: II. Positional voting. *Economic Theory*, 15, 55–101.
- Saari, D.G. (2001). *Decisions and elections: Explaining the unexpected*. Cambridge, United Kingdom: Cambridge University Press.
- Shepsle, K., & Bonchek, M. (1997). *Analyzing politics: Rationality, behavior, and institutions*. New York: Norton.
- Tangian, A. (2000). Unlikelihood of Condorcet’s paradox in a large society. *Social Choice and Welfare*, 17, 337–365.
- Thurstone, L.L. (1927). A law of comparative judgment. *Psychological Review*, 34, 273–286.
- Tideman, N. (2006). *Collective decisions and voting: The potential for public choice*. Williston, VT: Ashgate.