Learning, Choice Consistency, and Individual Differences in How People Think Elections Should be Decided

by

Zachary A. Caddick

A.A. in Social and Behavioral Sciences, Chaffey College, 2011

B.A. in Psychology, California State University San Bernardino, 2013

M.A. in Experimental and Research Psychology, San José State University, 2016

M.S. in Cognitive Psychology, University of Pittsburgh, 2020

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts and Sciences

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2022

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Zachary Alan Caddick

It was defended on

December 8, 2022

and approved by

Scott H. Fraundorf, Ph.D., Associate Professor, Department of Psychology

Kevin R. Binning, Ph.D., Associate Professor, Department of Psychology

Nicholas R. Miller, Ph.D., Professor Emeritus, Department of Political Science, University of Maryland Baltimore County

Dissertation Director: Benjamin M. Rottman, Ph.D., Associate Professor, Department of Psychology

Learning, Choice Consistency, and Individual Differences in How People Think Elections Should be Decided

Zachary Alan Caddick, PhD

University of Pittsburgh, 2022

There are ongoing debates about whether the U.S. should switch from plurality voting to alternative systems (e.g., cardinal or ranked-choice voting) and debates about the relative fairness and ease of learning different systems. To address these issues, we developed the 'Who Won the Election Task' (WWET) in which participants were shown the results of a hypothetical election in which a group of people were voting on which candidate to hire. The WWET had participants determine elections from raw data and allowed us to calculate the degree to which participants' choices agreed with the three voting systems. In four studies, we evaluated how participants' preferences about voting systems, the consistency in these preferences when measured in different ways, and whether their understanding of the voting systems and individual differences predicted their voting system preferences. Additionally, we tested educational interventions, which improved participants' understanding of the voting systems. Across all the studies, participants' choices in the WWET were most consistent with plurality voting. However, participants tended to view ranked-choice voting as fairer than plurality. In Studies 3 and 4 participants even sometimes viewed cardinal voting as fairer than plurality. In general, we found low consistency in voting system preferences when measured in different ways. One reason this may occur is because participants struggled to comprehend the alternative voting systems and were not adequately self-assessing their own knowledge. This research has

implications for persuading the public to change voting systems for elections as well as how groups should make collective decisions (e.g., hiring decisions).

Table of Contents

List of Tables

List of Figures

Preface

If there is one thing that is clear to me as I reach the end of this long journey it is this: I would not be here without the help of innumerable people. What might appear to some as an individual accomplishment is truly the success of a village. I have been especially fortunate to have so many wonderful mentors and allies along my journey. The sheer volume of professors, researchers, and fellow learners who have—for reasons often beyond my understanding—taken an interest in my development and helped nurture and direct my growth leaves me speechless. For that I will be forever changed, and I lack the ability to effectively communicate my gratefulness in words. I will seek to repay the balance of their investment by paying it forward and investing in others.

I want to specifically thank my advisor Ben Rottman, for his guidance on my dissertation project as well as throughout my tenure at Pitt. I know that I am a sharper thinker and better scientist for having had the experience of being Ben's advisee. I also want to thank my dissertation committee who have helped guide and shape this project in important ways. I want to thank Danielle Nebril, Claire Pamerleau, Thomas Saba, and Jackson Weiler for their excellent contributions as research assistants on this project.

There are many people in which if I imagine a counterfactual world where they were not a part of my life, I would not be here today—not least of which are my causal reasoning lab mates: Kevin Soo, Cory Derringer, Ciara Willett, Yiwen Zhang, and Sara Jaramillo. The first year here Kevin and Cory's programming assistance made a world of difference in my life. So did all our lunches, the competitive games of GeoGuessr, breaks in the office where we would watch government hearings, the endless puns and general lightheartedness. Although we may "live in a

d-cell world after all"—your *presence* certainly had an *effect* on my own trajectory. I am thankful for their camaraderie and friendship.

I also want to thank my family for their support. Prior to starting the program, my dad joined me in San Jose, California to begin a ~3,000 mile road trip to Pittsburgh (with a detour in Yellowstone). There's a famous quote that goes: "'You can never cross the ocean unless you have the courage to lose sight of the shore." I think it is worth noting that sometimes both the courage and means to reach new places are the result of the support of others. I am acutely aware of that lesson as my time in graduate school comes to an end.

If there was any one person who I cannot imagine completing the doctoral program without their presence in my life it is my partner Erin Gustafson, my ally and confidante. I am both looking forward to the next chapter and confident in our ability to conquer obstacles after everything we have been through together the last few years.

1.0 Introduction

When people come together to make a group decision, they must decide how the preferences of group members will be collected and aggregated. The process of taking input from individuals and aggregating these preferences into a single shared decision is a pervasive phenomenon known as collective decision-making. Collective decision-making exists in many domains, including how a group of friends decide which restaurant to eat at, how a hiring committee decides which job candidate to hire, how voters elect politicians, and many more. Collective decision-making also occurs in other species (e.g., bees; Seeley, 1991). Collective decision-making can have many benefits. Groups of individuals often make better decisions than any one individual (e.g., wisdom of the crowds; Kameda, Toyokawa, & Tindale, 2022). Group decisions can also lead to greater buy-in and consensus among group members (e.g., democracy). Perhaps the most salient example of collective decision-making in society is voting. Voting systems are formalized methods for aggregating preferences and can be used for electing political candidates but also in other situations to reach a group decision.

The specific voting system used during collective decision-making has important implications for the resulting decision. First, the scale used for voting can vary. Individuals may choose a single option out of two or more options, they may rank the options in order of preference, or they may rate the options by preference strength. Most commonly in U.S. elections, voters provide a single choice for a preferred candidate and the candidate who receives the most votes wins, a method known as *plurality voting*. In contrast, in *cardinal voting*, voters provide a rating of the strength of their preferences, which allows them to provide more nuanced information about many options. Second, beyond the inputs for a collective decision, the rules for aggregating preferences to reach a decision can vary innumerably and can alter outcomes.

Different ways of collection (nominal, ordinal, metric) are somewhat tied to different methods of aggregation (e.g., plurality, cardinal voting), but not in a perfect one-to-one mapping. For example, in both ranked-choice voting (RCV) and the Borda method of voting, voters rank candidates. However, the way the votes count and the method of aggregating is quite different across these two voting systems. In the Borda method of voting the ranking given to each candidate is associated with a different weighted score. The scoring formula is *N*-*R*, where *N* is the total number of options (e.g., candidates) and *R* is the rank given by a voter. For example, in a three-candidate election, the first ranked candidate on a ballot will receive 2 points, the second–1 point, and the third–0 points. The Borda aggregation method involves summing the points for each candidate with the winner being the candidate with the most points. Contrastingly, RCV works by first checking if any candidate has at least 50% support of the highest rank. If so, this candidate wins. If not, then the candidate with the least amount of support according to the highest rank is eliminated and those votes are re-assigned to the voter's next favorite candidate. The votes are then counted again to see if a candidate has over 50% support. If so, this candidate wins. If not, the process is repeated.

Different voting systems can lead to different group decisions. There are situations in which, if a single group of voters is asked to submit their votes using two different voting systems, and they vote sincerely both times, the two voting systems could result in a different winner. Ideally, both candidates and voters would feel that so long as the voting system has been followed, that the election is fair, even if the candidate themselves or a voter's preferred candidate loses. But, the very notion of fairness is potentially undermined–and that the very least, made more complex–by the fact that different voting systems can lead to different outcomes.

If people disagree about which voting system is the most fair they may feel that their

preferred candidate only lost because one voting system was used instead of another, which could feel arbitrary or potentially even systematically biased against one candidate or one party. Furthermore, less formal situations of collective decision making (e.g., deciding where to go to dinner, or which candidate to hire), the decision makers may not have even thought about how they would aggregate votes in advance, and then wind up in a situation in which there are multiple options, which may lead to different outcomes. In sum, picking a voting system for a collective decision is vital because it sets out the rules for making a 'fair' decision. However, individuals may disagree over the voting system or may not even realize that they need to pick one or that there are multiple options.

The current paper is motivated by the question of what voting systems people prefer and why. Making progress towards this question is important so that groups can pick voting systems that feel fair to them. This question is deceptively simple and raises a number of other questions that need to be addressed. The goal of this paper is not to seek a definitive answer to the question of what voting system people prefer, which would require a representative survey, or assessing if different groups have different preferences. Instead, the goal is to develop methods to understand how to go about studying this question. In an effort to do so, we examine four interrelated questions. First, are voting system preferences consistent when comparing different ways of measuring preferences? Second, how well do individuals understand different voting systems and is understanding related to individuals' preferences? Third, what is the best way to teach people about different voting systems and when people learn more about the voting systems, do their preferences change? Fourth, are there individual difference measures that predict and may explain voting system preferences? Though it would be convenient if these questions could be addressed independently, we do not feel that it is possible. For example, understanding which

voting system is preferred by a group may depend on how the preferences are measured and how much knowledge the individuals have about the voting system. A researcher must then decide which way to go about measuring preferences, and how much information to convey about the voting systems.

In the remainder of the introduction, we will first discuss how the current context of voting in the U.S. connects to the motivation behind this project. Second, we will discuss relevant contributions from the literature on social choice theory that aid our understanding of voting systems. Third, we discuss different methods of measuring voting system preferences and the importance of understanding if these measurements are consistent (or not). Fourth, we discuss how teaching, learning, and understanding may influence voting system preferences. Lastly, we discuss individual difference measures which may help explain why people prefer certain voting systems.

1.1 Current Context of Voting Systems in the U.S.

The specific way people vote in the U.S. has been under increasing scrutiny, and there is a growing interest in modifying the electoral system in the United States. One reason for this interest is a dissatisfaction with the two-party system, which tends to develop when plurality voting is used in single-winner elections (Duverger's law; Duverger, 1959). Indeed, a month before the 2022 midterm elections, 35% of Americans stated that they do not identify as belonging to either of the two dominant political parties (Gallup, 2022). and 62% of Americans state that a third party is needed in the United States (Jones, 2021). Ranked-choice voting (RCV) has seen recent popularity as an alternative to plurality voting. In 2018, voters in Maine approved a ballot initiative to switch to using RCV. In 2020, voters in Massachusetts failed, and voters in Alaska succeeded, in passing ballot measures to switch to RCV. Additionally, 53 U.S. cities

currently use RCV, with the number expected to grow (FairVote, 2022).

There are many different voting systems to choose from and infinite rules that could govern a given voting process. Political organizations exist that advocate for specific alternative voting systems such as RCV (e.g., FairVote, RankTheVote), approval voting (The Center for Election Science), and range voting (The Center for Range Voting). While there are many proposals for changes to voting systems, there is a lack of psychological evidence about people's preferences and how well people understand different voting systems.

Though there is some polling on such topics, it is quite superficial. For example, Alaskans for Better Elections commissioned an exit poll during a special election in August of 2022, the first use of Alaska's new ranked choice voting system. The survey asked, "In your opinion, how simple or difficult was it for you to fill out your Ranked Choice Voting ballot?" on a four-point scale from very simple to very difficult (Alaskans For Better Elections, 2022). Eight-five percent of respondents answered, 'very simple' or 'somewhat simple', which was interpreted as a success for RCV. However, this poll only assesses respondents' perceived difficulty of filling out the ballot, not whether they understand how the RCV scoring procedure works or whether they find the scoring procedure simple and fair. A voter might find the ballot simple and yet not understand the voting system. Sarah Palin, the second-place finisher in that election has called RCV "untrustworthy" and "cockamamie", and more broadly the election has spurred on the debate about whether RCV is fair and understandable (Herz, 2022). One strength of an exit poll survey is that it is potentially representative of a specific population. In contrast, the current research is not designed to be representative of a given population, but it is designed to dive into the details about what may guide public information campaigns to teach about new systems, how to describe voting systems in ballot measures, and how to persuade people to see benefits of

alternative systems. These goals are relevant to political activist organizations as well as governmental agencies trying to explain how a voting system works. More broadly, understanding how people think that group choices should be made is relevant to any group decision, not just political voting.

1.2 Social Choice Theory

"Social choice theory is the study of collective decision processes and procedures. It is not a single theory, but a cluster of models and results concerning the aggregation of individual inputs" (List, 2013). Social choice theory has a rich history spanning over 200 years, with many notable thinkers under its umbrella (e.g., Charles Dodgson AKA "Lewis Carroll", John Stuart Mill). Some of the earliest work in social choice theory involved mathematically describing strengths and weaknesses of different voting systems. The foundational figure Nicholas Condorcet, for instance, famously described the phenomenon of "Condorcet Cycles" where rock-paper-scissors-like scenarios can play out in election results in which it is unclear which of several candidates is the clear winner (Condorcet, 1785). In these scenarios, candidates both best and are bested by other candidates depending on head-to-head matchups in such a way that it is ambiguous which candidate should be considered the ultimate winner.

Kenneth Arrow famously introduced an "impossibility theorem" that mathematically proved that there is no voting system that fulfills five specific desired criteria ("axioms") for 3 or more options and 2 or more voters (Arrow, 1951). Though there is some disagreement about the importance of all the axioms and how strictly they should be adhered to (Stanford Encyclopedia of Philosophy, 2022), many are very straightforward. One axiom is that the voting system should not be a dictatorship (i.e., one person should not be making the decision for all people). Another is the independence of irrelevant alternatives (i.e., if a losing candidate/choice is

removed or added from the race/consideration, the outcome of the election/decision should not change). A third requires that not only should a voter's preference by transitive, but aggregated preferences should be transitive as well. Arrow's impossibility theorem demonstrates that no voting system upholds all of these desirable features.

Many theorems have followed in the wake of Arrow, including the Gibbard-Satterthwaite theorem, which builds on Arrow's Impossibility Theorem and introduces the problem of strategic voting (Gibbard, 1973; Satterthwaite, 1975). Strategic voting occurs when a voter casts a vote that deviates from their true preferences (e.g., voting for a candidate they think is likely to win rather than for their favorite candidate if they think that their favorite will lose) in order to alter the outcome of an election. Honesty and deception from individuals add another layer of complication to enacting desirable voting systems, and there are voting systems (e.g., median voting) that are designed to minimize the influence of strategic voting. In the rest of this manuscript, we do not consider strategic voting.

Aside from the theoretical work, there is also empirical work within social choice theory, which involves experiments with groups of people who are making decisions from spatial information in a paradigm referred to as "Spatial Voting Games" (e.g., Bianco, Lynch, Miller, & Sened, 2006). Although these games exist in different forms, they typically assign individuals a fixed preference represented in a hypothetical *n*-dimension space (typically 1-2 dimensions). After each individual is assigned their spatial location, the group must decide where to place their collective decision in a deliberative-tug-of-war style game, where each individual is motivated to pull (or "tug") the decision to be as close to them as possible. Thus, for a consensus to be reached in these games, individuals must take others' preferences into account.

From the social choice theory literature, we see that no one voting system may always work

perfectly and different voting systems prioritize different aspects of fairness, which may lead to individual differences in voting system preferences. The current research contributes to this line of work by trying to add a psychological component: understanding which voting systems people prefer in light of the fact that they offer competing strengths and weaknesses.

1.3 Measuring Voting Preferences & Choice Consistency

In order to understand which voting systems individuals believe to be most fair, one must create a measure of voting system preferences. There are at least two obvious ways to do so. One way is to create a survey. The survey method, however, also requires teaching people about how different voting systems work, and the names for different voting systems, so that they can rate each voting system. A limitation of this approach is that even with clear explanations for how a voting system works, people may still not understand it. This could raise a further issue - if someone thinks that a voting system is fair or unfair, but they do not actually understand it, how meaningful is the fairness rating? Perhaps instead the rating reflects a fairly quick and uninformed impression. For this reason, we created an alternative way to assess how people think that an election should be decided.

We created a task in which participants are shown ballots cast by voters, and they decide which candidate they think won the election; we call this task the "Who Won the Election Task" (WWET). Because there are many ways to make this decision it is not a trivial decision. By choosing who they think won, participants essentially provide their opinion about how the election should be decided fairly. Importantly, this task does not require teaching people about the voting systems, instead it is just how they would decide if faced with the voting data. Their choices can be interpreted as revealing preferences for different voting systems, since different choices are consistent with different voting systems. Thus, this task can be viewed as similar to

work in the revealed preferences literature (c.f. Beshears, Choi, Laibson, & Madrian, 2008). *Revealed preferences* are preferences revealed from an individual's choices. For instance, if an individual is willing to pay more for a product that is healthier and overlook a cheaper, but less healthy, alternative–we could state that this individual has revealed a preference for health.

The Who Won the Election Task is also similar to a classic cognitive science research paradigm that uses "summary tasks." Summary tasks have participants extract summary information from visual imagery (Szafir et al., 2016). For instance, an image may be shown with many circles of different sizes and a participant may be asked to estimate the average circle size. Humans can efficiently average visual information across size, orientation, direction, and position (Szafir, et al., 2016, and citations therewithin). Additionally, people can effectively average visual information even with varied distributions (Dakin, 2001; Szafir, et al., 2016). This line of work provides foundational evidence for which we can reasonably believe that choices in our task reflect differences in preference and not choices in error due to a lack of ability to effectively aggregate individual stimuli.

It is not obvious which method of measuring voting systems preferences would be better, so we investigated both of them. If the two methods are *consistent*, that is good news because it will make the job for future surveys (e.g., nationally representative surveys, ballot proposals) of voting preferences easier. However, if they are not consistent, this raises questions for how to accurately measure voting preferences, and in fact there may be some inherent duality in people's preferences–that individuals think they prefer one system but actually prefer another.

1.4 Teaching People about Voting Systems

An issue raised in the previous section is that a survey in which people report preferences for voting systems requires teaching them the names and procedures for the voting systems. This

raises the issues of whether people can understand the voting systems from a relatively brief tutorial, and what ways of teaching these are better versus worse. Helping people to understand voting systems is important for two other reasons. First, if there is interest in adopting a new voting system and the choice is put forth on a ballot initiative, people need to understand the various systems being put forth in order to vote on a voting system. Second, if a new voting system is adopted, it would be important to mount educational campaigns to teach people not just how to fill out a ballot but also how the votes get tallied.

The ease of teaching the public about ranked-choice voting in particular has been a point of contention. For example, the policy director for Massachusetts' recent ballot proposal on ranked-choice voting, Greg Dennis, said "We've known all along that once we're able to inform someone about how ranked-choice works, they support it" (Boston Globe Editorial Board, 2020). However, Gov. Charlie Baker of Massachusetts has stated that he is opposed to the "very complicated" system. Although many disagreements in politics may stem from differences in values, these are specific empirical claims about how easy different voting systems are to learn and which voting systems people prefer and why. In this research, we studied questions such as how best to teach people about voting systems, how malleable are people's preferences for voting system preferences, and when taught about the voting systems, do their preferences change?

Though we do not know of any research that has studied how to teach people about voting systems, a number of websites teach people about voting systems. These websites often use scenarios in which learners are given hypothetical information about distributions of voter preferences and walked through the counterfactual of how different voting systems would handle the situation (e.g., FairVote, 2022; The Center for Election Science, 2021). The information

about voter preferences is often number/percent of voters with each ballot ranking. Often this information is presented in a tabular format, although lists are common too. Many organizations pair teaching with persuasion by also using scenarios that show an undesirable outcome for one voting system and how another voting system would handle that case better (e.g., FairVote, 2022; Mr. Beat, 2019). For instance, the most notorious example of an undesirable outcome may be spoiler effects within plurality voting. (A spoiler effect is when the presence of an unpopular third candidate can change who wins. It is also possible to have spoiler effects in RCV; although the candidates must be fairly competitive for these outcomes to occur, whereas in plurality spoiler effects can occur with low levels of support among the 'spoiler' candidate.)

In the current research we studied how to teach people about voting systems, which is a primary goal of Study 2 (chapter 3). We will address our motivations for the specific teaching interventions used in the introduction to that chapter. At a high level, understanding how a voting system works may be necessary to believing the resulting decision is fair. Additionally, because different voting systems are presenting competing ideas of fair processes, ideally voters understand how the different options work and believe the one used is the fairest option, which also should increase acceptance of the results as fair.

One other aspect related to people's understanding of voting systems is the extent to which understanding is related to preferences. It is possible that people prefer voting systems that they understand better (or think that they understand better). One factor that influences individuals' preferences is *fluency*. People tend to prefer stimuli that they have seen before in contrast to near-identical, but novel stimuli (Bornstein & D'Agostino, 1992; Zajonc, 1968). Positive feelings are typically felt by people as a result of the ease of processing from repeatedly retrieving some target information (Bornstein & D'Agostino, 1992, 1994). For instance, people

may prefer plurality voting both because they are used to it, which increases fluency, and also because it is an easy voting system. Related to evidence from literature on fluency is simplicity theory which states that people have a bias for experiences that minimize cognitive load (Chater, Vitanyi, 2003). Chater and Vitanyi argue that humans' preference for simplicity is reflected in our general preference for the simplest method or simplest explanations that adequately addresses some phenomena (e.g., Occam's Razor). For these reasons, voting system preferences may be influenced by understanding. If an individual only groks plurality voting, then that may be their preferred voting preference when asked. However, if an individual understands an alternative voting system *and* sees a need addressed by that voting system, they may be more likely to prefer that voting system. Whatever the specific reasons for why an individual prefers a given voting system, it is clearly important to evaluate how well individuals understand different voting systems and to address any deficiencies so that people can effectively vote, understand the results, and believe it is a fair process.

1.5 Individual Difference Measures and Collective Decision-Making

It is unclear exactly why people may prefer one voting system over another. The "Who Won the Election Task" can provide insights into what voting systems individuals believe lead to the fairest outcomes. Additionally, it is possible inter-individual differences may also play an important role. Voters may view certain voting systems as being better aligned with their moral values. We identified several measures of individual differences to help explain voting system preferences: moral foundations, utilitarianism, numeracy, and cognitive reflection.

1.5.1 Morality

It is possible that moral priorities influence preferences for different voting systems. Moral priorities have been shown to predict other political decisions like who someone votes for

president (Franks & Scherr, 2015), for instance, and Republicans and Democrats in the U.S. have contrasting moral profiles (e.g., Franks & Scherr, 2015). Additionally, like the work in motivated reasoning–where the role of emotion and motivation has been shown to reliably alter individuals' reasoning processes–the last two decades of research on morality has shown that emotional responses to moral issues can explain reasoning processes, often with much greater explanatory power than 'cold' rational reasoning theories (see Haidt, 2001; Kunda, 1990).

To measure important aspects of morality we used The Moral Foundations Questionnaire (MFQ), which has 5 sub-scales representing distinct morals: authority/respect ("concerns related to social order and the obligations of hierarchical relationships"), harm/care ("basic concerns for the suffering of others, including virtues of caring and compassion"), in-group/loyalty ("concerns related to obligations of group membership, such as loyalty, self-sacrifice and vigilance against betrayal"), fairness/reciprocity ("concerns about unfair treatment, inequality, and more abstract notions of justice"), and purity/sanctity ("concerns about physical and spiritual contagion, including virtues of chastity, wholesomeness and control of desires"; Haidt, Graham, Joseph, 2009). Liberals tend to have higher levels of fairness/reciprocity and harm/care, whereas conservatives tend to have higher levels of in-group/loyalty, authority/respect, and purity/sanctity.

We predict that a preference for plurality voting will be more likely for conservatives, and those with conservative-learning moral foundation traits (greater in-group loyalty, authority/respect, and purity/sanctity). One reason for this prediction is that conservatives may be more motivated to uphold systems and less willing to make structural changes to existing systems.

Another reason may have to do with how the different voting systems prioritize different

aspects of fairness. For instance, plurality voting systems determine the winners based on the option with the most votes. However, in plurality voting it is possible to win with well below 50% of the vote and even for the majority of the voters to not prefer the winning candidate. In this way, plurality voting values the outcome desired by the single most dominant group even if it is against the wishes of the majority. In contrast, RCV winners must have a majority of voters' support. Similarly, a cardinal voting system which determines a winner by calculating the average rating for each candidate and choosing the one with the best average, prioritize winners that are most representative of voters. For these reasons, conservatives and those with conservative moral values may prefer a voting system that prioritizes the desires of the single-most dominant group—which plurality does.

1.5.1.1 Utilitarianism.

We propose that utilitarian views will predict voting system preferences. Utilitarianism is a moral view that seeks to maximize the well-being of all people. Given that voting systems prioritize different aspects of fairness, one factor that could influence voting system preferences is the extent to which an individual holds utilitarian values. We would expect utilitarianism to be associated with RCV and cardinal preferences instead of plurality voting, given their emphasis on majority approval and representativeness, respectively.

To measure utilitarianism, we used the Oxford Utilitarianism Scale, which has two sub-scales: *impartial beneficence,* which measures an "aspect of utilitarianism that seeks to impartially maximize welfare and the greater good, even at expense to oneself" (p. 143; Kahane et al., 2018) and *instrumental harm,* which measures "allowing harm in the service of the greater good" (p. 143).

1.5.2 Cognitive Reflection

Cognitive reflection, the ability to override intuitive responses, has been shown to predict individual differences in one's reasoning ability and personal beliefs. For instance, individuals higher in cognitive reflection tend to have higher SAT scores (Frederick, 2005), lower levels of conventional religious beliefs (Pennycook et al. 2012), greater ability to discern "pseudo-profound bullshit" (as measured by meaningless, but syntactically correct, sentences filled with buzzwords; Pennycook et al. 2015), and rational thinking (Toplak, West, & Stanovich, 2011). Additionally, Pennycook and Rand (2018) found that Trump voters in the 2016 U.S. election had lower levels of cognitive reflection than Clinton supporters.

We predict that cognitive reflection will relate to voting system preferences. For example, one of the reasons individuals may prefer plurality voting is because it is cognitively easier to understand. Additionally, plurality voting is the norm in the U.S. and cognitive reflection has been shown to be related to rejecting conventional thinking. Thus, individuals higher in cognitive reflection may also be more likely to be attracted to atypical practices. Additionally, individuals high in cognitive reflection may be more likely to expend the cognitive effort needed to understand the merits of alternative systems.

1.6 Summary of Studies and Hypotheses

We conducted four studies to examine how individuals aggregate preferences.

1.6.1 Hypotheses & Predictions

(1) **Choice Consistency.** Are judgments about who won an election consistent with their judgments of which voting system is viewed as most fair when directly asked? [*Study 1, 2, 3, 4*]

(2) **Individual Differences.** Are there individual differences that are associated with different voting system preferences? [*Study 1, 4*]

(3) **Understanding and Preferences**. Do people prefer voting systems that they understand better? [*Study 1, 2, 3, 4*]

(4) **Teaching Interventions.** Can we identify better ways to teach people voting systems and do these different teaching methods lead to different views about which systems are most fair? [*Study 2*]

2.0 Study 1

2.1 Methods

2.1.1 Pre-registration

All four studies were pre-registered on Open Science Framework (OSF) [https://osf.io/n75cz/]. For logistical reasons, data collection for all four studies began prior to posting pre-registration. However, data was never examined prior to posting our hypotheses and analysis plans.

2.1.2 Participants

We recruited 300 participants through the University of Pittsburgh's introduction to psychology subject pool. Participants in all four studies had to be at least 18 years old and fluent in English.

2.1.3 Design, Procedures, and Measures

Figure 1 shows the procedural flow of the study. There were four main tasks: the Who Won the Election Task, the Voting System Comprehension Test, Declared Voting System Preferences, and a series of individual difference measures. These are discussed in turn. First participants were introduced to the graphs that were used in the "Who Won the Election Task" (Appendix A1). Second, participants completed 3 questions to ensure they understood how to read the graphs (Appendix A2). These questions are designed to have obviously correct answers (i.e.,

every voting system agrees who the winner should be). Third, participants completed the "Who Won the Election Task" (described below).

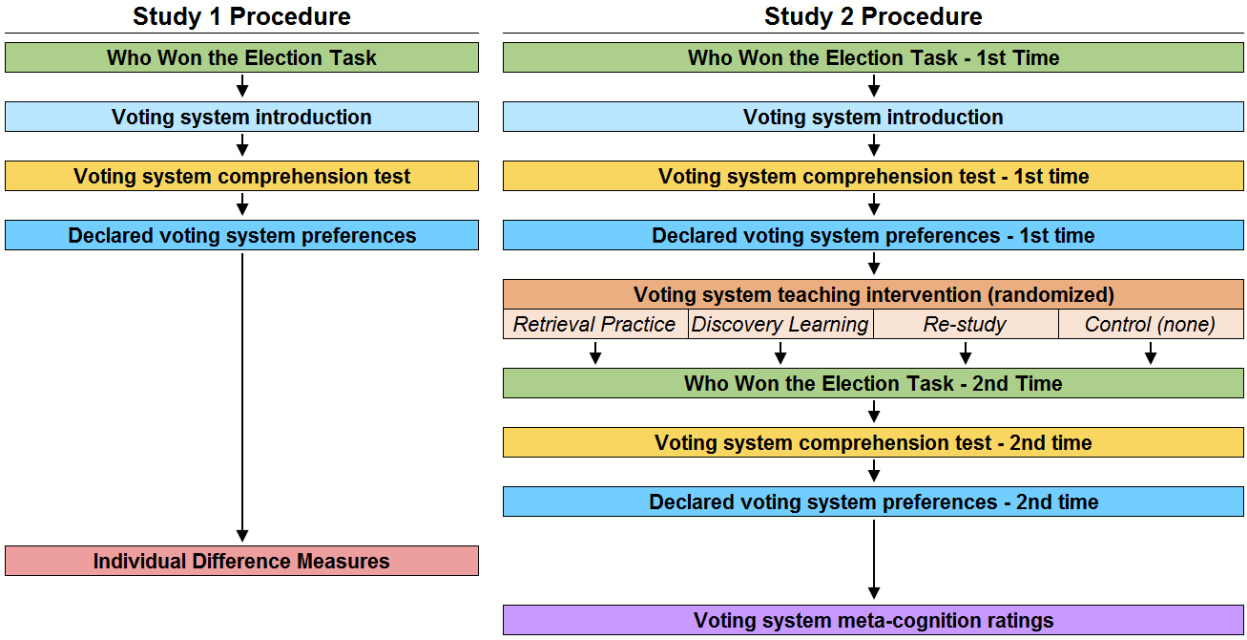


Figure 1. Study 1 and 2 Procedure Overview. Note: There were other blocks of the survey that are not denoted here. The figure only includes analyzed blocks.

2.1.3.1 Who Won the Election Task (WWET).

2.1.3.1.1 The WWET Procedure. Full details for the WWET are available in Appendices A1-A3. In the WWET, participants were asked to imagine that they worked for a company deciding to hire a new employee from three candidates.[1] Each member of the hiring committee voted on whether they wanted a candidate who excelled technically or in terms of leadership on a -100 to +100 scale, and each candidate was assessed through the hiring process at a specific number on the same scale. A graph was provided to participants that represented the preferences of the hiring committee members (vertical black lines) and the candidates preferences (vertical blue

[1] We chose not to use a political voting situation, but an analogous task would be to see three candidates that range from strongly conservative to strongly liberal, and a set of voters who lie on the same dimension, and to need to decide which candidate is best to represent the set of voters. This setup is inspired by the spatial model of voting game literature, in which the theoretical question is how to choose a candidate that mathematically best represents a set of voters. In Studies 3 and 4, a slightly different setup will be used in which each voter rates each candidate rather than being positioned on a single dimension.

lines and letters; e.g., *Figure 2*). Based on the graph, participants judged "Based on the preferences of the hiring committee members in the graph above, which job candidate do you think should be hired?" Every participant saw 10 graphs and thus made 10 judgments.

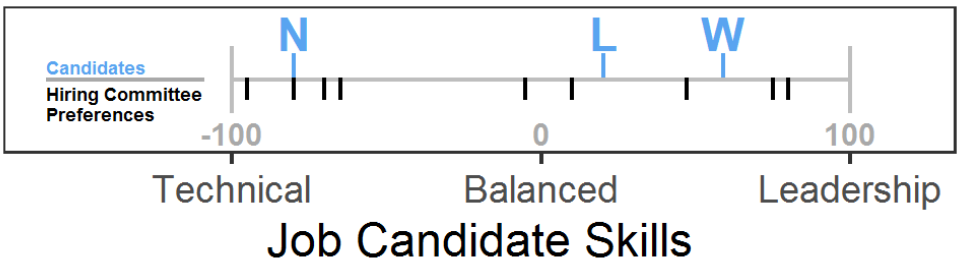


Figure 2. Example stimuli for Who Won the Election Task.

From the figures in the WWET, it is possible to determine which candidate should be hired according to plurality, RCV, and cardinal voting. In *Figure 2*, Candidate N wins according to *plurality* voting, because 4 committee members are closest to N, whereas 3 are closest to W and 2 are closest to L. In *Figure 2*, Candidate W wins according to *RCV*. Since Candidate N had 4 out of 9 votes in the first round, which is less than 50%, an instant-runoff would occur. Candidate L would be eliminated in the instant-runoff because L has the least number of votes. The two voters in the middle of the scale are closer to W than N, so W would win the election with 5 out of 9 votes. In *Figure 2*, Candidate L wins according to *cardinal* voting because L has a lower average distance to all 9 committee members (63.00) than N (71.90) or W (78.70).

2.1.3.1.2 The WWET Stimuli. In most elections, many scoring systems choose the same candidate. Because we wanted the WWET to be able to differentiate participants' preferences for the different voting systems, we chose stimuli for which the different voting systems would elect different candidates. We created 10 figures like *Figure 2*. In 9 of the 10 graphs there were three job candidates and in 1 there were 2 candidates. All graphs had between 7 and 17 voters ('committee members'). We used relatively small numbers of voters and candidates to reduce the

cognitive burden of the task. For each of the 10 figures, we creating a mirrored version flipping the X-axis, and participants were randomly assigned one of the two versions. The locations of the voters were chosen so that it should be relatively easy to determine if they were closer to one candidate or another.

In 4 of the 10 graphs, all three voting systems choose a different winner. (With three candidates, there must be at least 9 voters to get these three voting systems to diverge.) In these four graphs, like in *Figure 2*, the middle candidate was the candidate picked by cardinal voting; since cardinal voting picks the candidate closest to all voters, in contentious elections (with three candidates and in single-dimension voting games) this ends up being the middle candidate.

We did not want to only have graphs in which all three voting systems diverge because this situation is fairly rare in a statistical sense. Thus, we included some other situations. In 2 graphs, cardinal voting and RCV picked the same candidate, and plurality picked a different candidate.

In the remaining 4 graphs, plurality voting and RCV picked the same candidate, and cardinal voting picked a different candidate. In addition, in 3 of the graphs there was a clear loser; none of the three voting systems chose one of the candidates. Since it is hard to justify why the losers would be picked, we used these as a check that participants were putting forth sufficient effort on the task.

2.1.3.1.3 The WWET Scoring System.

We developed two methods for scoring the extent to which participants' choices in the WWET were more or less consistent with each of the three voting systems. Table 1 provides a concrete example of these two methods. Table 1 uses the voting situation from *Figure 2*. The three rows depict the scores that a participant would receive if they chose each of the three candidates as the winner.

The three rows for the Raw Voting Values are the numbers already explained above; L wins for cardinal because L has the lowest average distance to all the committee members (63.00), N wins according to plurality because N has the highest number of committee members that are closest to N (4) , and W wins according to RCV because once L is eliminated W has more committee members that are closest to W than N does (5 vs. 4).

Table 1. Who Won the Election Task Scoring Method for data in Figure 2

	Raw Voting Values			Transformation: Method 1			Transformation: Method 2		
Choice	*Cardinal*	*Plurality*	*RCV*	*Cardinal*	*Plurality*	*RCV*	*Cardinal*	*Plurality*	*RCV*
N	71.90	4	4	0	1	0	.663	.444	.444
L	63.00	2	0	1	0	0	.705	.222	.000
W	78.70	3	5	0	0	1	.632	.333	.556

Note: Higher raw voting values for RCV and plurality reflect better scores, but for cardinal lower raw values are better.

In Scoring Method 1, the choices are scored either a 1 (choice is consistent with a voting system or a 0 (inconsistent with voting system). In *Figure 2*, each of the three voting systems chooses a different candidate, so each choice in Table 1 is associated with a single system. However, for some items in the WWET one choice is consistent with two voting systems. The scores from the 10 items in the WWET are averaged for a given participant, resulting in three numbers representing the consistency between a participant's choices and the three voting systems. Because some choices are consistent with two voting systems, these scores do not sum to 1.

Scoring Method 2 captures the degree to which a given choice agrees with the choice made by each voting system. For plurality and RCV, the scores are calculated by dividing the Raw Voting Values by the total number of voters, so these values sum to 1. For example, if a participant chose W as the winner in *Figure 2*, this choice is 33% consistent with Plurality and 55% consistent with RCV. For Cardinal voting, since the raw voting values are average distances (lower is better), the formula for Scoring Method 2 is (1-(target score/sum of all scores)). Thus,

the score for candidate N is calculated as 1-(71.9/(71.9+63+78.7))=.663. (Note that the three Cardinal scores sum to 2, not 1 like the other systems, but this is inconsequential for subsequent analyses.) The scores for the 10 items in the WWET task are averaged so that each participant has three scores that represent the degree of consistency between their choices each of the three voting systems. A benefit of Scoring Method 2 is that it represents degree of consistency rather than all-or-nothing consistency. For example, in *Figure 2*, even though L is technically the best choice according to cardinal voting, N is also understandable because it is a close second.

Although Scoring Method 2 is useful because it represents the degree of consistency, some of our results are more easily interpreted using a 0-100% agreement scale that Method 1 offers.

2.1.3.2 Voting System Introduction

After the WWET, participants were taught the three voting systems, cardinal, plurality, and ranked-choice voting, and the names for these systems (Appendix A4). After learning a given voting system, participants rated it on a 5-point Likert scale, 1 being "very unfair" and 5 being "very fair."[2]

2.1.3.3 Voting System Comprehension

Next, participants answered 9 questions to test their comprehension of the voting systems (Appendix A5). Participants were shown three graphs, and for each they were asked which candidate won according to the three voting systems, in a counterbalanced order. For these three graphs, the three voting systems picked different candidates. Each question was scored as correct (1) or incorrect (0). For some analyses, the sum of the 3 comprehension questions within a voting

[2] To simplify the results section, we do not report these ratings. Because these ratings are made sequentially after introducing each voting system, rather than at once after all three have been introduced, in hindsight we feel that they are less reliable. We found that the ratings changed more dramatically from the first ratings to the second ratings across our studies, but then became more stabilized at later time points. We believe that not understanding all of the available alternatives can change the calculus of any one choice's worth and thus the later ratings are more reliable.

system were used on a scale of 0-3; and for other analyses, the sum of the 9 comprehension questions were used.

2.1.3.4 Declared voting system preferences

Participants were asked "Which of the three voting systems below do you think is best?" [options: ranked-choice, plurality, cardinal voting]. Participants also rated the fairness of each voting system (Likert-scale 1-5 ratings: 1=very unfair, 5=very fair). See Appendix A6 for question text.

2.1.3.5 Individual Difference Measures

Participants first completed the Cognitive Reflection Test (CRT); it was always included first to control for cognitive fatigue. Then, participants completed measures of numeracy, utilitarianism, and moral foundations, in a random order. After, participants answered some questions about political orientation and demographics.[3]

2.2 Results

2.2.1 Participants

Five participants were removed who did not originally submit complete data. Eleven participants were removed because they did not correctly answer the questions to ensure they could read the graphs (4% of the remaining sample). The criterion for inclusion was not missing more than one question on the graph questions; this criterion was set after examining the distribution of responses with the goal of eliminating participants who were not reasonably attending to the materials. (The same approach is used in all four studies.) In all, 284 participants submitted valid data for analysis.

[3] We also asked six questions about participants' self-reported political knowledge, behavior, and beliefs, but these results are not included for concision; see Appendix A7 for items and descriptive statistics.

2.2.2 Validity Check

As a measure of attention/comprehension, we examined how often the "loser" candidates were selected. Participants selected the loser candidate 6.81% of the time. The majority of "loser" candidate choices occurred for item 5 (~95%), which may be more attributed to the characteristics of that item (e.g., it is more difficult) than a lack of attention or effort on the participants' part. In general, the low frequency of choosing candidates that were not supported by any of the voting systems suggests that participants took the WWET seriously.

2.2.3 Voting System Choices

2.2.3.1 WWET Choices

Descriptive statistics for both WWET choices and declared voting system preferences for all studies can be seen in *Figure 3*. Within the WWET (Method 1), participants chose the plurality option most frequently, followed by the cardinal option and RCV; a one-way repeated measures ANOVA found a significant difference among these choices, $F(2, 566) = 71.93$, $p < .001$, $\eta_G^2 = .198$.[4]

2.2.3.2 Declared Voting System Preferences

In the declared voting system preferences, the majority of participants preferred RCV (*Figure 3*). A Chi-square goodness of fit was significant meaning that the distribution was different from chance, $\chi^2(2) = 47.04$, $p < .001$.

2.2.3.3 Choice Consistency

From the descriptive statistics just discussed, it is already apparent that there is some discrepancy between the choices in the WWET and the declared voting system preferences because plurality is the most common for the WWET and RCV was typically viewed as the best

[4] To aid in interpreting ANOVA effect sizes in this paper, we include Bakeman's (2005) rules of thumb for generalized eta scores: small (.02), medium (.13), large (.26).

for the declared preferences. Here, we examine the consistency between these two different ways of eliciting preferences more formally (*Figure 4*). A Linear Discriminant Analysis (LDA) was conducted to examine if choices in the WWET (Method 2) can successfully predict participants' declared voting system preference. Overall, choices in the WWET were able to predict participants' declared voting system preferences for 41.50% of participants. The hit rate (probability of correctly classifying participants) was 56.30% for cardinal voting, which is above chance, $z = 3.96$, $p < .001$; people who picked cardinal voting also used cardinal voting the most in the WWET. The hit rate for plurality was 52.10%, $z = 3.42$, $p < .001$; people who picked plurality also used plurality the most in the WWET. However, the model could not predict who would choose RCV above chance; the hit rate was 30.2%, $z = -.73$, $p = .465$. The participants who chose RCV in their declared preference used plurality the most in the WWET, similar to those who chose plurality in their declared preference. Summarized another way, participants who declared cardinal the best, tended to make cardinal WWET choices, but both participants who declared plurality and RCV the best tended to make plurality WWET choices.

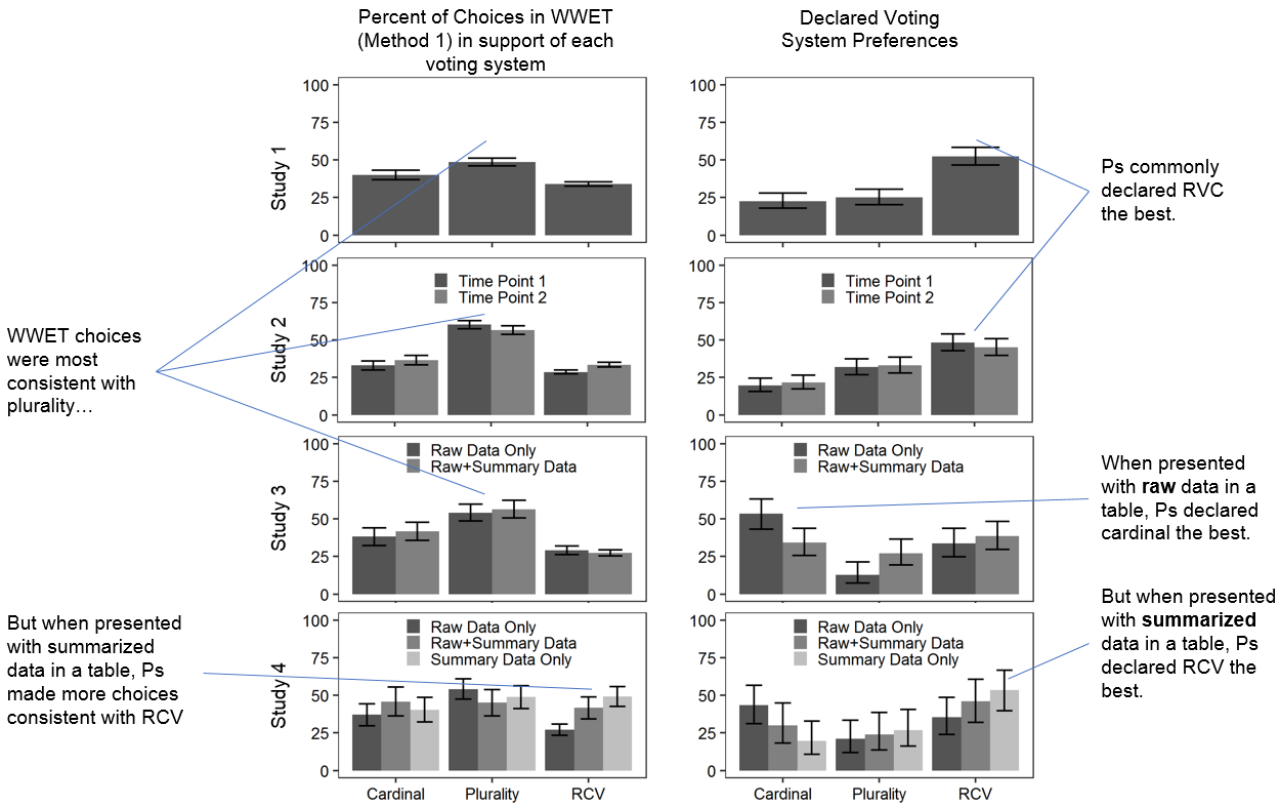


Figure 3. Percent of choices in WWET in support of each voting system and declared voting system preferences and 95% CIs. Note: Some choices in the WWET support two voting systems, so the bars add up to over 100%. For studies in which the tasks were performed more than once, the first is shown.

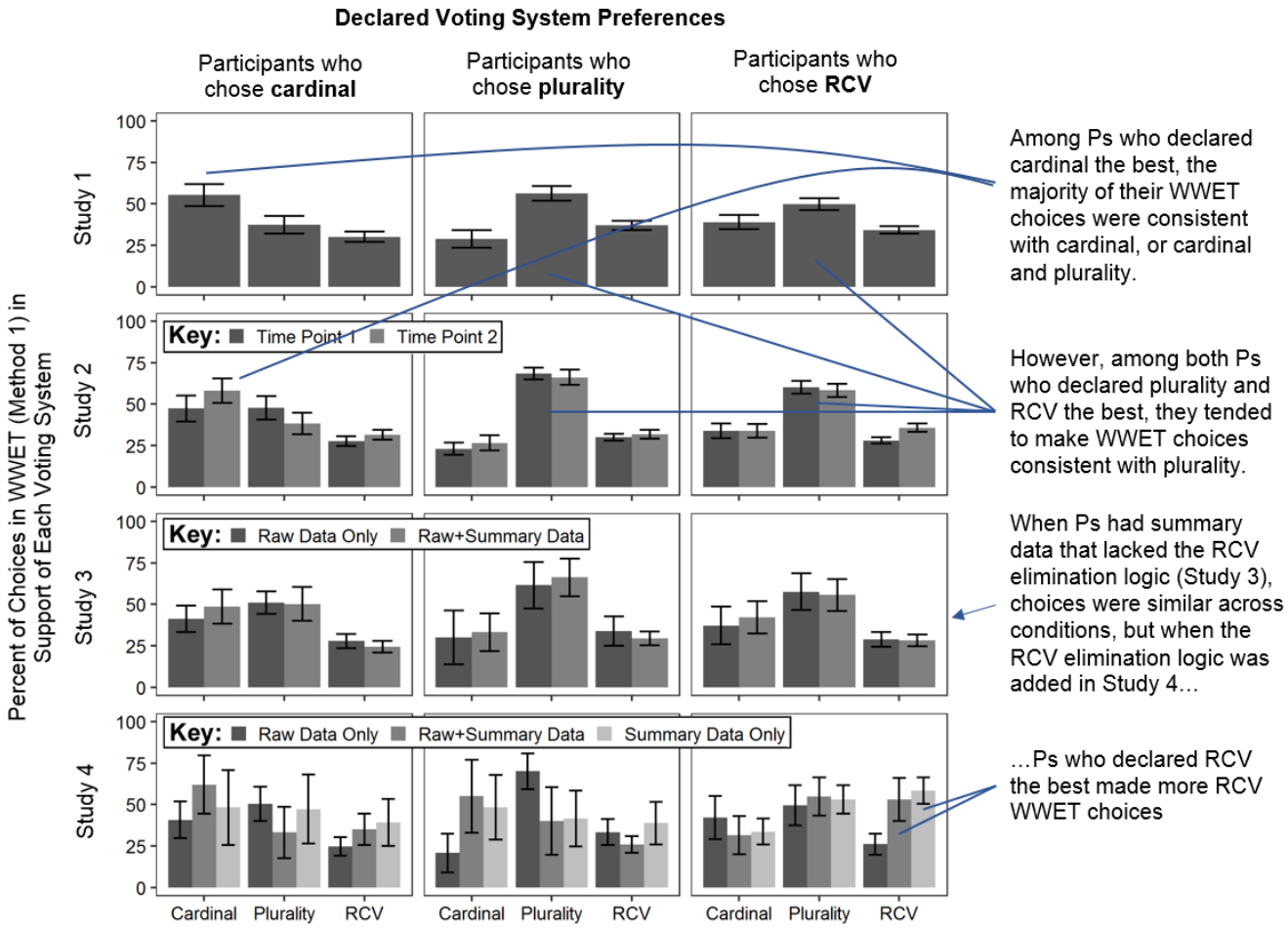


Figure 4. Consistency between declared voting system preferences and choices in WWET. Note: For studies in which the WWET was performed more than once, the first is shown (with the exception of Study 2).

2.2.3 Individual Differences

2.2.3.1 Predicting Declared Voting System Preferences by WWET and Individual Difference Measures

We first assessed whether the individual difference measures could, in aggregate, explain participants' declared voting system preferences above and beyond the WWET. Adding the individual difference measures to the LDA model had a negligible increase in the ability to correctly classify voting preferences choices from 41.50% to 41.90%. From the LDA, we found that participants who tended to make fewer WWET plurality choices tended to make more

WWET cardinal choices (and vice versa). Cardinal WWET choices were associated with higher scores in utilitarianism, cognitive reflection, numeracy, and learning comprehension scores. Plurality WWET choices were associated with higher levels of fairness/reciprocity and conservatism, and lower levels of, authority/respect and harm/care (see Appendix E for complete LDA results).

A series of one-way ANOVAs were conducted to test if each of the three WWET (Method 2) scores and the individual difference measures differed by participants' declared voting system preference (Table 2). There were significant differences by declared voting system preference for all three of the WWET choices. Participants who declared cardinal to be their preferred voting system made WWET that were most in support of cardinal voting. However, both participants who declared plurality and RCV to be their preferred voting system made WWET choices that were mostly in favor of plurality.

With regards to the individual difference measures, there were significant differences for three measures. Participants who declared cardinal and RCV to be their preferred voting system had higher Cognitive Reflection Task scores than plurality. Participants who preferred plurality voting had higher levels of Authority/Respect.

2.2.3.2 Predicting Choices in the WWET by Individual Difference Measures

We also examined how well the individual difference measures predicted choices in the Who Won the Election Task. Three linear regression models were conducted to predict each of the WWET choices separately (cardinal, plurality, & RCV choices) using the Method 1 scoring system (Table 3). The individual difference measures explained 10.59% of the variance in the extent to which participants' WWET choices supported cardinal voting, 10.08% for plurality, and 6.95% for RCV.

Similar to the declared voting system preferences, higher CRT scores were associated with making fewer plurality choices. Unlike for the declared voting system preferences, the WWET choices were not predicted by Authority/Respect, and higher Purity/Sanctity scores were associated with more RCV choices in the WWET.

Table 2. Mean (SD) of Measures by Declared Voting System Preference, and ANOVA Results

	Declared Voting System Preference			ANOVA	
Measures	Cardinal	Plurality	RCV	*F*	*p*
WWET Method 2 Scores					
Cardinal	.68 (.02)	.66 (.02)	.67 (.02)	18.95	< .001
Plurality	.33 (.06)	.39 (.05)	.37 (.06)	16.19	< .001
RCV	.31 (.10)	.40 (.09)	.37 (.10)	13.62	< .001
Individual Difference Measures					
Authority/Respect	14.56 (5.05)	16.13 (5.12)	13.42 (5.39)	6.48	.002
Fairness/Reciprocity	21.84 (4.55)	22.38 (3.84)	22.82 (4.04)	1.29	.276
Harm/Care	21.63 (4.58)	21.68 (4.19)	21.40 (4.38)	.12	.884
In-group/Loyalty	13.48 (5.03)	14.27 (4.84)	12.63 (5.30)	2.55	.080
Purity/Sanctity	13.17 (5.42)	14.23 (5.25)	12.42 (5.76)	2.54	.080
Conservatism	3.06 (1.58)	3.08 (1.40)	2.89 (1.43)	.55	.576
Utilitarianism	3.82 (.65)	3.57 (.81)	3.64 (.82)	1.87	.156
Cognitive Reflection Test	1.52 (1.13)	1.00 (1.03)	1.40 (1.14)	4.30	.014
Numeracy	4.61 (.88)	4.37 (.80)	4.49 (.85)	1.39	.252

Note: WWET choices are using the method 2 scoring system (weighted).

Table 3. Three Regressions Predicting WWET Choices from Individual Difference Measures

	WWET Method 1 Scores					
	Cardinal		Plurality		RCV	
Measures	*β*	*p*	*β*	*p*	*β*	*p*
Individual Difference Measures						
Authority/Respect	-.01	.119	.00	.183	.00	.779
Fairness/Reciprocity	.00	.368	-.01	.246	.00	.417
Harm/Care	.00	.631	.00	.489	.00	.657
In-group/Loyalty	.01	.175	-.01	.129	.00	.277
Purity/Sanctity	.00	.227	.00	.252	.00	.011*
Conservatism	-.01	.646	.01	.515	.01	.348
Utilitarianism	.00	.879	.01	.719	.01	.179
CRT	.03	.069†	-.03	.035*	.00	.657
Numeracy	.03	.107	-.03	.077†	-.01	.230

Note: † = $p < .10$, * = $p < .05$.

2.2.4 Comprehension and Declared Voting System Preferences

A 3x3 Factorial ANOVA was conducted to examine if participants' comprehension of the

voting systems differed for the three voting systems (within-subject) and across declared voting system preferences (between-subjects); see *Figure 5* for descriptive statistics. There were significant differences in the comprehension of the three voting systems, $F(2, 562) = 148.78$, $p < .001$, $\eta_G^2 = .22$. Participants comprehended plurality better than the cardinal, $t(283) = 12.37$, $p < .001$, $d = .62$, and RCV, $t(283) = 16.49$, $p < .001$, $d = .98$. Additionally, participants comprehended cardinal better than RCV, $t(283) = 7.68$, $p < .001$, $d = .46$.

There was also a main effect of declared voting system preference, $F(2, 281) = 4.01$, $p = .019$, $\eta_G^2 = .01$. Participants who preferred plurality voting had lower comprehension scores than participants who preferred cardinal, $t(132.47) = 2.58$, $p = .011$, $d = .46$, and RCV $t(143.63) = 2.53$, $p = .013$, $d = .45$. There was not a significant between participants who declared RCV vs. cardinal voting the best, $t(129.49) = .44$, $p = .660$, $d = .08$.

Additionally, there was a significant interaction between a participant's declared voting system preference and their voting system comprehension, $F(4, 562) = 2.38$, $p = .005$, $\eta_G^2 = .01$. Comprehension scores for plurality and RCV did not differ much based on declared voting system preference. In contrast, there were differences in comprehension of cardinal voting; participants who declared cardinal to be the best system had the best comprehension and those who declared plurality had the lowest. In sum, the main effect of declared voting system preference appears to be mainly driven by the comprehension of cardinal voting.

In sum, plurality voting was the easiest for participants to comprehend, followed by cardinal, and lastly RCV. Participants who declared cardinal voting the best had the highest comprehension scores, followed by RCV and then plurality, and this finding was mainly driven by the comprehension of cardinal voting.

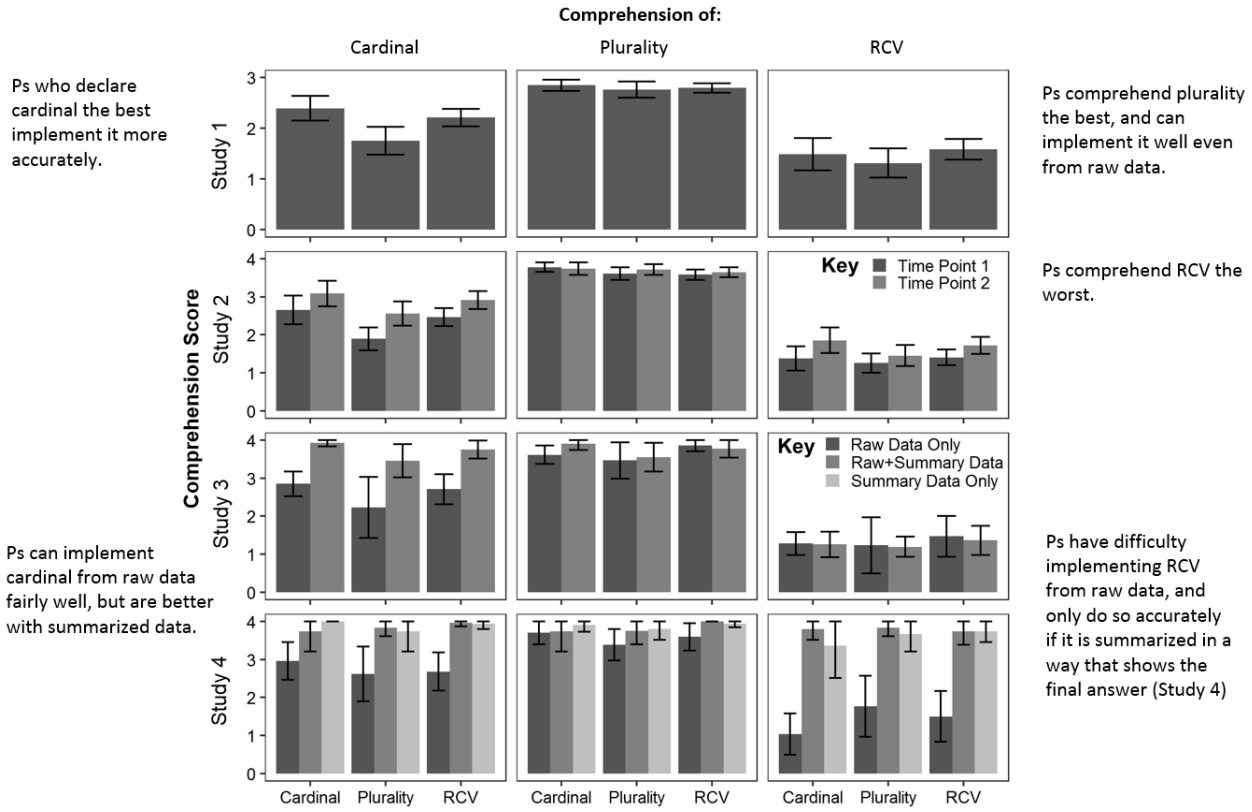


Figure 5. Comprehension scores for the three voting by preferred voting system. Note: For studies in which the WWET was performed more than once, the first is shown (with the exception of Study 2). Study 3 and 4 use the same Key.

2.3 Discussion

Our new WWET was designed to put participants in the position of deciding 'who won the election,' which requires participants to decide how to aggregate raw data to make a decision. We scored their choices in the WWET by how much they support plurality, cardinal, and RCV. We compared these choices to the participants' declared voting system preferences.

Participants' choices in the WWET supported plurality the most, but they declared to prefer RCV the most. When the WWET choices were compared with the declared preferences, participants who declared preferring cardinal made more WWET choices that supported cardinal than the other voting systems. However, both participants who declared preferring plurality and RCV made WWET choices that supported plurality more than the other voting systems.

This suggests that few participants spontaneously use a mental process akin to RCV, which

involves redistributing votes for the least viable candidates, and they had the lowest comprehension scores for RCV, even though it just requires one step beyond plurality. However, this did not seem to stop participants from stating a preference for RCV.

We also found a number of individual difference measures that were associated with participants' WWET choices and their declared voting system preferences. Preferring cardinal voting (especially in comparison to plurality), in both the WWET and in declared preferences, was associated with greater CRT scores and greater comprehension of the voting systems. These results suggest that plurality voting is a cognitively simpler approach. Additionally, plurality-consistent choices in the WWET were associated with higher levels of fairness/reciprocity and lower levels of authority/respect, which went against our predictions. However, plurality-consistent choices in the WWET were associated with lower harm/care and generally had a poorer understanding of the voting systems, especially cardinal, as expected.

3.0 Study 2

One of the primary findings in Study 1 was that participants often had inconsistent voting system preferences, depending on how they were asked. Specifically, participants' revealed preferences in the WWET task tended to favor plurality, but their declared preferences tended to favor RCV. One reason for this inconsistency may be due to a poor understanding of how the voting systems work. In some cases, a participant may even think they are making a decision consistent with a given voting system in the WWET task, but instead their choice is consistent with another voting system.

In Study 2 we added a more extensive teaching phase to see if participants' preferences become more consistent when they understand the voting systems better. It is unclear from Study 1 why a gap in voting system preferences emerged when they were measured in two different

ways. One possibility is that our participants did not understand how to effectively implement the alternative voting systems. It could even be the case that participants believed they were making choices consistent with the voting system they declared as preferring but were mistaken. One reason why it is so important that individuals understand how to implement a voting system is so that they have trust in the eventual outcome. In Study 2, we also compared three teaching interventions to test if they produce any differences in understanding and differences in declared preferences.

There are countless different approaches that one could apply in teaching people about voting systems; we tested three. The first approach we tested was *retrieval practice*, also known as the *testing effect*, which involves having learners retrieve information from their memory in a testing environment as a method to boost recall on subsequent tests (Adesope, Trevisan, & Sundararajan, 2017; Rowland, 2014; Yang, Luo, Vadillo, Yu, & Shanks, 2021). Importantly, the act of retrieving information itself is responsible for the increased ability to recall that information in the future (Adesope, Trevisan, & Sundararajan, 2017; Rowland, 2014; Yang, Luo, Vadillo, Yu, & Shanks, 2021). Furthermore, the learning benefits of retrieval practice can be boosted by providing feedback on test questions (Rowland, 2014).

The benefits of retrieval practice are often compared to a re-study learning intervention, in which a learner re-reads information instead of being prompted to recall the information. Meta-analyses have found that retrieval practice typically outperforms re-study with medium-to-large effect size differences (Adesope, Trevisan, & Sundararajan, 2017; Rowland, 2014; Yang, Luo, Vadillo, Yu, & Shanks, 2021). Re-study will be our second teaching method, because 1) it is a useful comparison group and 2) it is similar to stimuli that has been previously used by political scientists to teach people about voting systems (e.g., Kimball, & Anthony, 2021). In our design,

re-study will provide evidence as to whether it is enough to simply remind people with short text-descriptions of how each voting system works.

The third approach is discovery learning. Discovery learning is very similar to inquiry-based learning and we use the term *discovery learning* to refer to both. In discovery learning, there is "a strong emphasis placed on the learner as an active agent in the process of knowledge acquisition" (de Jong & van Joolingen, 1998, p. 179). Discovery learning creates opportunities for target experiences where learners discover causal relations through problem-solving in a manner similar to a scientist (Pedaste et al., 2015; Svinicki, 1998).

Given the emphasis on learning the mechanisms behind phenomena both in science and in our materials teaching different voting systems, we believe that discovery learning may be an effective approach for learning how voting systems work. Although discovery learning has been especially popular in science domains, it has also been found to be beneficial for learning mathematics (Kamaluddin & Widjajanti, 2019). Given that voting systems operate under logical and mathematical principles, this provides additional support for the idea that discovery learning may be an especially beneficial teaching strategy for voting systems.

We had two main predictions. First, we predict that retrieval practice will lead to the greatest increase in comprehension about voting systems. As previously mentioned, retrieval practice has been robustly demonstrated as an effective teaching intervention. By focusing the testing items on implementing the mechanism behind each voting system, we believe this condition should be especially effective at increasing comprehension. Second, past research has found that teacher-lead learning activities lead to greater student knowledge, but that student-lead activities lead to greater student belief change (Alexander et al., 2002). We predicted that our discovery learning intervention will lead to greater change in participants' preferences than the other interventions

that are less self-directed; individuals may be more persuaded by information that is not provided by an authority figure and rather learned through a peer or self-discovery.

Beyond adding the teaching interventions, we also sought to understand how metacognition may relate comprehension of the voting systems. Metacognition is an individual's knowledge about their own thinking or one's "knowledge and cognition about cognitive phenomena" (Flavell, 1979). In Study 2, we did not have specific *a priori* predictions related to metacognition. However, given the poor understanding of RCV in Study 1, coupled with the fact that most participants declared a preference for RCV, we wanted to investigate the role metacognition may be playing.

3.1 Method

3.1.1 Participants

We recruited 400 participants through the University of Pittsburgh's introduction to psychology subject pool.

3.1.2 Design, Procedures, and Measures

The design of Study 2 was similar to Study 1 with the following two key differences. First, participants were randomly assigned to one of four teaching interventions (retrieval practice, re-study, discovery, and a control condition). Second, the WWET, comprehension questions, and declared voting system preferences were measured twice, once before and once after the teaching intervention. *Figure 1* provides a summary of the procedural flow for Study 2. Appendices B1-B3 present the full details of the introduction, questions related to whether individuals can effectively read the graphs, and WWET. The differences compared to Study 1 are described below.

3.1.2.1 Who Won the Election Task.

Several important changes were made to the WWET. The most important change was to the dependent measure. In Study 1, participants answered a forced-choice question about which candidate should be hired. This single choice of question matches the format of ballots for plurality voting systems, which may have been viewed as an endorsement of plurality voting. To reduce this potential bias, participants answered three questions, which used formats similar to the ballots used for the three types of voting systems. First in alignment with plurality voting, participants answered "Based on the above graph, which job candidate do you think should be hired?" Second, in alignment with RCV, participants were asked to "Rank order the candidates based on who you think should be hired." Third, in alignment with cardinality voting, participants were asked to "Rate the candidates based on who should be hired. (100=most fair to hire; 0=least fair to hire)." For simplicity, if participants' top choice agreed across all three questions, or across two out of three questions, that choice was coded as the winner. If all three questions diverged, or if the plurality and RCV ballot diverged and the cardinal ballot was tied, the participant's response was omitted for that item.

Two other smaller changes were made to the stimuli. In Study 1, the stimuli displayed 'Leadership' and 'Technical' as opposite anchors on a single dimension, but they are not necessarily opposites; a candidate could be high on both qualities. In Study 2, we changed the anchors to "Specialist" and "Generalist" with no middle label (*Figure 6*). Additionally, because the WWET was completed at two time points in Study 2, a second mirrored version of the stimuli with different candidate labels was created.

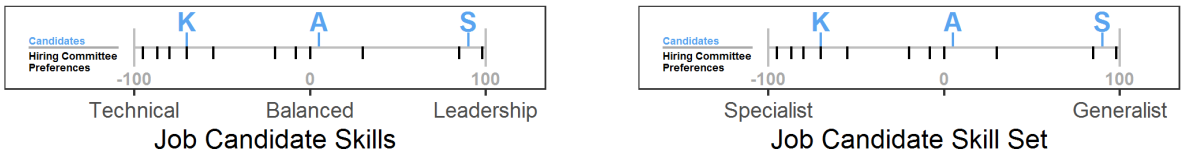


Figure 6. Stimuli change in the "Who Won the Election Task." Note: The X-axis labels were the primary alteration.

3.1.2.2 Voting System Introduction.

This section of the survey was near-identical to Study 1, the only change being that the ratings were answered on a 0-100 scale in Study 2 (instead of a 1-5 scale in Study 1). This was changed so that all cardinal ballot style questions used a similar number range (Appendix B4). However, these data are not analyzed for reasons discussed in Study 1.

3.1.2.3 Voting System Comprehension.

The voting system comprehension check was almost identical to that of Study 1 (see Appendix B5 for complete details). A fourth graph was added to increase measurement sensitivity; thus, the possible correct score in Study 2 is 12 (instead of 9 in Study 1). For the second round of answering the comprehension question, a mirrored version was created.

3.1.2.4 Declared Voting System Preferences

Similar to the change to the WWET dependent measure, participants answered three questions (see Appendix B6 for complete details). First, in alignment with plurality voting, participants answered "Which of the three voting systems below do you think is most fair?" Second, in alignment with RCV, participants were asked to "Order the voting systems below (drag and drop list items), from most fair (top) to least fair (bottom)." Third, in alignment with cardinality voting, participants were asked to "Rate the voting systems below, from very fair (100) to not very fair (0)." Declared voting system preferences were measured after both of the comprehension question blocks.

3.1.2.5 Teaching Intervention

Participants were randomly assigned to one of three teaching interventions (retrieval practice, re-study, or discovery learning), or a control condition which did not involve any additional teaching (see Appendix B7 for complete teaching intervention details). We first discuss the core similarity of the three teaching interventions before explaining how they differed.

In all three teaching conditions, participants learned about an undesirable aspect of each of the three voting systems–we call these undesirable aspects 'quirks' of the voting systems. The two goals of teaching about quirks were to encourage participants to think about what is a fair aggregation processes and to highlight the mechanism of how each system operates. These quirks are well-known phenomena in the literature on voting systems that arise due to the mathematics of the voting systems. We did not include any quirks that arise with strategic voting in which participants vote for candidates in a way that deviates from their true preferences. (Insincere or strategic voting complicates the costs and benefits of each voting system more than we felt could address in these initial studies, but is an important topic for future research into voting system preferences.) The presentation order of the three quirks was randomized.

The first quirk was the "center squeeze phenomenon" which is a weakness of ranked-choice and plurality voting. In the center squeeze phenomenon moderate candidates can lose to extreme candidates because they are "squeezed" by candidates on either side of them taking away voters. The second quirk was the "majority criterion violation," which is a quirk of cardinal voting. In this quirk, the winning candidate according to cardinal voting is not the first choice of a majority of voters according to plurality or RCV. The third quirk was the "spoiler effect," which plurality voting is especially susceptible to. The spoiler effect is when the presence of an unpopular third candidate can change who wins. The spoiler effect is likely the most famous voting quirk with

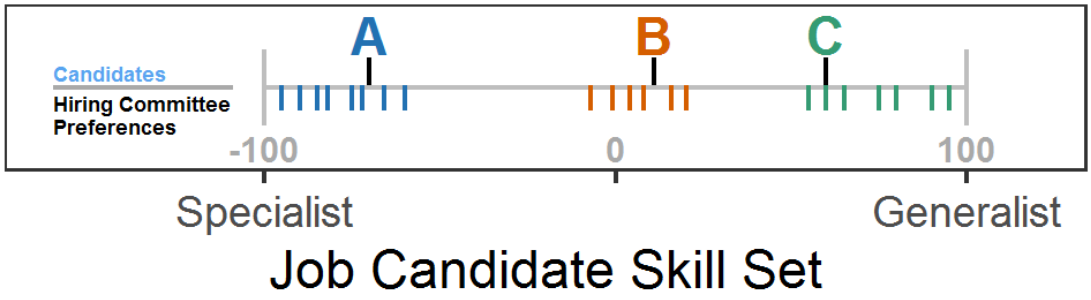
the public.

On the next few pages you will be walked through a specific voting example called the "Center Squeeze Phenomenon," which shows how ranked-choice voting can hurt moderate or "middle-of-the-road" candidates.

[The re-study and retrieval practice interventions occurred here.

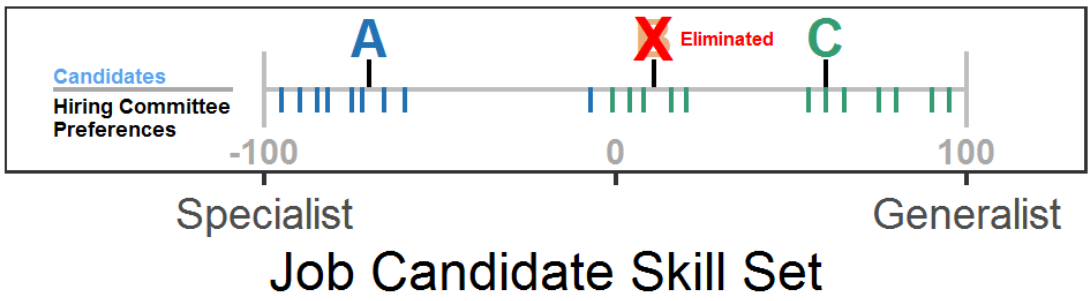
- *For re-study, participants were briefly reminded how each voting system works.*
- *For retrieval practice, it said"...but first we will ask you to determine the results of an election using different voting systems…." Participants were shown the figure below, and asked who won according to each voting system.*

Then, in all conditions participants were shown the following graphs and told who won according to each voting system.]



Who wins according to plurality and ranked-choice voting?

- 8 voters have candidate A as their favorite (8/21 voters or 38%). Therefore, **A wins according to plurality**.
- 6 voters have candidate B as their favorite (6/21 voters or 29%), and
- 7 voters have candidate C as their favorite (7/21 voters or 33%). **C wins according to ranked-choice voting**, because we would eliminate B, since no candidate has over 50%, giving C 12 of the 21 votes (57%).



Who wins according to cardinal voting?

- Candidate A is the furthest from the voters with an average distance of 74.4,
- Candidate B is the closest to the voters with an average distance of 57.6. Therefore, **B wins according to cardinal voting**.
- Candidate C is the second closest to the voters with an average distance of 73.1.

[The Discovery Learning Intervention Occurred Here]

Why is this "center squeeze" example problematic for ranked-choice and plurality voting?

In this "center squeeze" example, there is a fairly good middle-of-the-road candidate, but ranked-choice and plurality voting instead choose extreme candidates. Think about it this way - imagine that the voters were the same but instead of having three candidates there were only two. If this was just an

election with A vs B, B would win. If this was just an election of B vs C, B would win. If there was an election of A vs C, C would win. So in the head-to-head comparisons B wins two of the three elections.

Figure 7. Summary of the Teaching Interventions for the Center Squeeze Phenomenon in Study 2. Note: Text in black was shown to all participants. Text in grey explains how the three teaching interventions were integrated into the explanation of the quirk.

Figure 7 provides an example of how the center squeeze quirk is explained, and how the three interventions were added to the main text. In the re-study condition, participants were provided text descriptions of how each of the three voting systems works prior to learning about each of the three quirks. The reminders were the same as those provided before the comprehension questions. For example, cardinal voting was explained as "Cardinal voting focuses on the distances between each candidate and the voters. The candidate that is closest on average to the voters wins."

In the retrieval practice condition, before learning about each of the three quirks, participants were given three test questions in a random order, giving them practice identifying which candidate won the election according to each voting system, and they were given feedback afterwards. The stimuli were the same as those used later to explain the quirk, so the retrieval practice not only gave them practice implementing the voting systems but also reinforced the idea that different voting systems can lead to different winners, and participants might think about why this is the case.

In the discovery learning condition, while learning about each quirk, participants interacted with a JavaScript applet that allowed them to change the data in the graphs to see how changes in the voters or the candidates would affect the outcomes according to each voting system. The graphs in the applets matched the graphs used in the rest of the teaching intervention. The goal of the applet was to help participant understand how each voting system worked and the quirks behind each voting system. To facilitate self-directed learning, participants were given short-

answer questions meant to guide their "discoveries." The text at the bottom of *Figure 7* explains why the example is problematic occurred on the next page after the discovery learning, and ideally participants in the discovery learning condition already figured out the answer on their own.

The applet differed for each quirk (*Figure 8*). The applet for the center squeeze example could show or hide one of the candidates, which allowed for pairwise comparisons to be made (i.e., what if there were only two instead of three candidates?). When a candidate is removed the hiring committee member preferences are re-colored by which (present) candidate they are closest to. In this example, Candidate B wins out in pairwise comparisons. However, neither plurality nor RCV would choose B as the winner, only cardinal voting would. Participants were given two prompts to answer to help them discover this quirk: 1) "One way to assess the strength of competing candidate is to examine the outcomes from head-to-head matchups (i.e., limiting the options to two candidates, in a 1 vs. 1 competition). Using the interactive graph above, examine the different 1 vs. 1 comparisons. Are these comparisons useful? And if so, what do they reveal about problems with ranked-choice and plurality voting?" and 2) "The name center squeeze phenomenon refers to conditions where middle-of-the-road candidates are disadvantaged in ranked-choice voting (by middle, we mean candidates somewhere between two other candidates). What conditions do you think leads to middle candidates being disadvantaged in ranked-choice voting?"

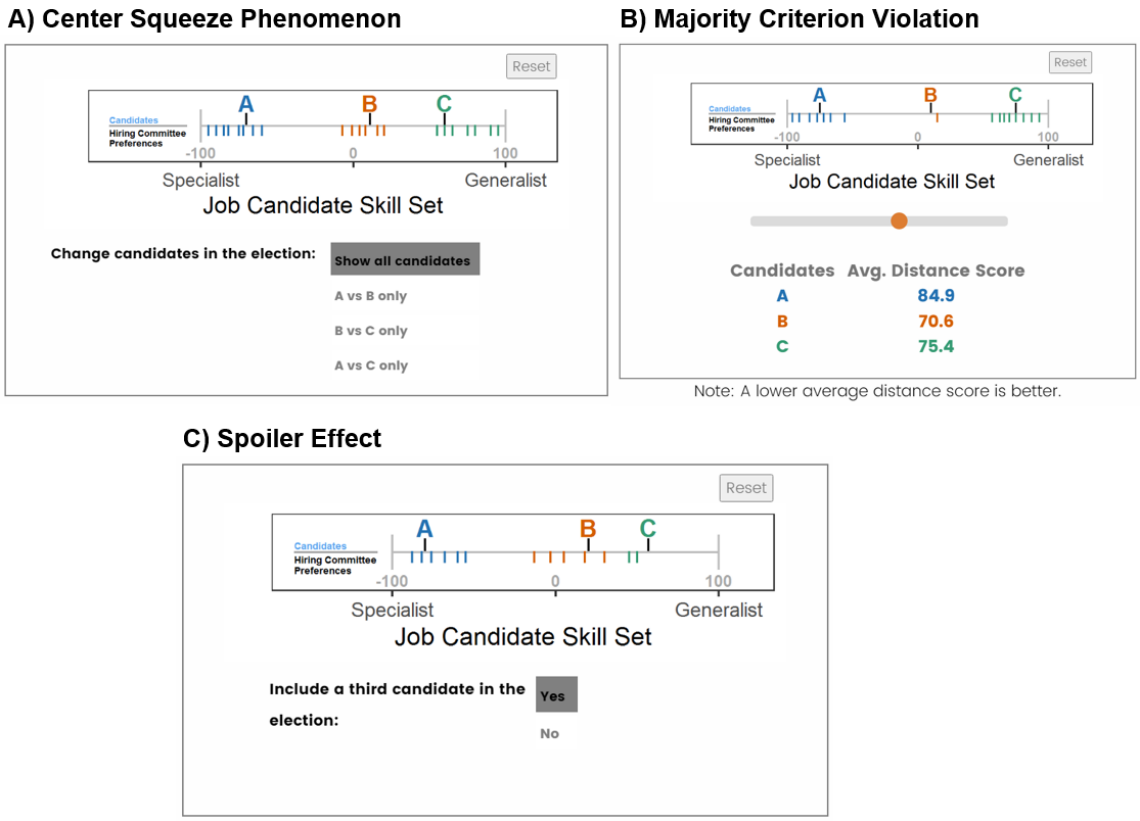


Figure 8. Screenshots of applet for "center squeeze phenomenon" (A), "majority criterion violation" (B), and "spoiler effect" (C).

The majority criterion violation tool had a slider that moved Candidate B's position when the slider was moved left or right. At the far-left position, B's average score is 79.8 and at the far right it is 70.3. This tool demonstrates that the best candidate position for cardinal voting (lowest average distance between each candidate and the voters) is near the median voter. This is problematic because cardinal voting can choose a candidate even in situations in which very few voters like the candidate (e.g., *Figure 8b*). The questions posed to participants were "B wins according to cardinal. Do you see any problems with that outcome?" and "Move around the slider and see how it affects the average distance score for candidate B. What position on the slider is best for candidate B (i.e., has the lowest score) and what does that tell us about cardinal voting?"

The spoiler effect tool was similar to the one used in the center squeeze example, but instead

of allowing participants to show or hide each candidate, only one third 'spoiler' candidate could be shown or hidden (Candidate C in *Figure 8c*). In the figure, even though Candidate B would win when pitted against A, when the spoiler C is added, A wins according to plurality voting. This reveals a weakness of plurality voting: that elections can be dramatically affected by spoiler candidates who win only a small number of votes. The changing states highlight the fact that very few voters (in this case, just two) can change the outcome of the election just by voting for their first or second favorite candidate. The participants were asked the question "What is one possible consequence of adding a candidate to an election using a plurality voting system?"

3.1.2.6 Voting System Metacognition

Participants were asked to self-assess their understanding of the different voting systems on a 0-100 rating scale.

3.1.2.7 Political orientation and demographic questions

Participants did not complete the same individual difference measures as in Study 1. Instead, most individual difference measures were dropped and only the political orientation and demographic questions were asked before participants were debriefed and thanked for their participation.

3.2 Results

3.2.1 Participants

We removed 13 participants who completed the study on their phones instead of a computer, as requested, because some parts of the survey were not mobile compatible. Additionally, 67 participants were removed because they did not correctly answer the questions ensuring they could read the graphs. In all, 321 participants submitted valid data for analysis.

3.2.2 Validity Check and Ballot Agreement

For three of the WWET items, there was a candidate who lost according to all three voting systems. As a measure of attention and comprehension of the figures, we examined how often the "loser" candidates were selected. Participants selected the loser 9.76% of the time at Time 1 and 7.89% at Time 2. Like in Study 1, the majority of the loser candidate choices came from item 5 (Time 1: 93%; Time 2: 92%). The low rates of choosing the 'loser' candidates suggests that participants took the task seriously and understood the stimuli.

Participants completed three ballots for all items in the WWET as well as for their declared voting system preferences. The participants' choices perfectly matched across all three ballots the vast majority of the time (WWET Time 1: 93%; WWET Time 2: 94%; Declared Time 1: 93%; Declared Time 2: 94%). When the three ballots did not match, then typically two did. In less than .5% of cases did all three choices diverge, so few choices had to be removed from analysis. This agreement again means that participants took the task seriously and also that the ballots did not dramatically change their declared preferences.

3.2.3 Voting System Choices

3.2.3.1 WWET Choices

Descriptive statistics for both WWET choices and declared voting system preferences for all studies can be seen in *Figure 3*. Similar to Study 1, at both time points (before and after the teaching intervention), participants chose the plurality option most frequently, then cardinal, and then RCV.

3.2.3.2 Declared Voting System Preferences

A chi-square goodness of fit analysis found that participants' preferences, assessed with the forced-choice measure, were not equal for the three voting systems, $\chi^2(2) = 39.98$, $p < .001$. Like

Study 1, participants most commonly declared a preference for RCV, followed by plurality and then lastly cardinal voting, at both Time 1 and 2 (*Figure 3*).

We had hypothesized that discovery learning may lead to the greatest change in declared voting system preferences. To test this, we ran a 4 (teaching interventions; between subjects) × 3 (voting systems; within subjects) ANOVA on the absolute value of the change in participants' fairness ratings from before versus after the intervention (Figure 9, Table G1 and Table G2). This analysis used the 0-100 fairness rating scale. There was a significant main effect of intervention, $F(3, 317) = 4.94$, $p = .002$, $\eta_G^2 = .021$. The control condition had less change in fairness ratings than discovery learning, $t(149.62) = 4.07$, $p < .001$, $d = .67$, retrieval practice, $t(165.99) = 2.59$, $p = .011$, $d = .39$, and re-study, $t(141.86) = 2.51$, $p = .013$, $d = .40$. There were no significant differences among the three teaching interventions. There was not a main effect of voting system, $F(2, 634) = 1.76$, $p = .173$, $\eta_G^2 = .003$ nor was a significant interaction, $F(6, 634) = 1.21$, $p = .302$, $\eta_G^2 = .006$. In sum, all three teaching interventions led to more preference change than the control condition, and though the discovery learning condition did have the largest effect size, it was not significantly larger than the other interventions.

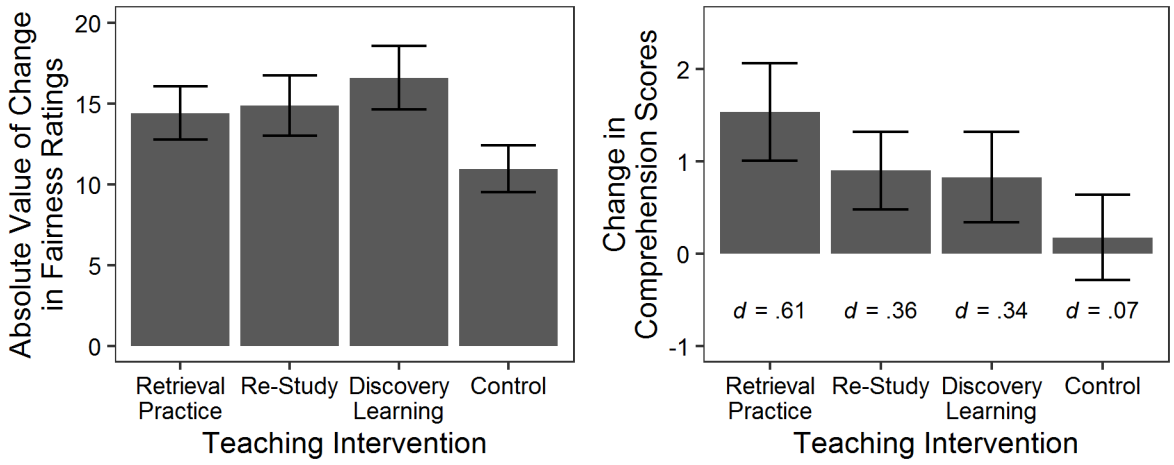


Figure 9. Belief Change in Fairness Ratings Across Teaching Intervention (left panel) & Mean (95% CI) Change in Comprehension Scores by Teaching Intervention in Study 2 (right panel).

3.2.3.3 Choice Consistency

In Study 2, there are two aspects with regards to the consistency of the choices in the WWET and the declared voting system preferences. First, to what extent are they consistent, both before and after the teaching intervention? Second, it would be desirable for any teaching intervention to increase the consistency between declared and revealed (WWET) voting system preferences. Thus, we examined changes in consistency before versus after the teaching intervention. *Figure 4* shows these results for both Time Point 1 and 2.

The choices in the WWET were able to predict participants' voting system preferences for 41.90% of participants at Time 1 and 47.50% at Time 2. The degree of consistency at Time 1 was very similar to that in Study 1, and there appears to be an increase of consistency due to the intervention.[5] Seventy-three percent of participants declared having the same favorite voting system at both time points, implying that, though there is a fair degree of consistency, there also was some change.

The hit rate for plurality choices was 59.80%, $z = 5.76$, $p < .001$ at Time 1 and 61.32%, $z = 6.20$, $p < .001$ at Time 2. The hit rate for cardinal choices was 49.21%, $z = 2.74$, $p < .01$ at Time 1, and 62.32%, $z = 5.18$, $p < .001$ at Time 2. The hit rate for RCV choices was 27.10%, $z = -1.56$, $p = .119$ at Time 1, and 30.34%, $z = -.68$, $p = .497$ at Time 2. In sum, similar to Study 1, people who used plurality and cardinal voting in the WWET tended to declare plurality or cardinal voting, respectively, as their preference. However, the participants' choices in the WWET were not predictive of who declared RCV as the fairest.

3.2.4 Comprehension and Declared Voting System Preferences

See *Figure 5* for results prior to the teaching intervention and for results after the intervention. Similar to Study 1, ANOVAs found that there were significant differences in the

[5] Unfortunately, we are unaware of a way to statistically compare these two.

comprehension of the three voting systems at Time 1 $F(2, 634) = 269.71$, $p < .001$, $\eta_G^2 = .335$, and Time 2 $F(2, 634) = 236.63$, $p < .001$, $\eta_G^2 = .279$. Participants understood plurality better than cardinal at both Time 1, $t(320) = 14.78$, $p < .001$, $d = 1.06$, and Time 2, $t(320) = 10.40$, $p < .001$, $d = .71$, and they understood plurality better than RCV at Time 1, $t(320) = 25.80$, $p < .001$, $d = 2.11$, and at Time 2, $t(320) = 23.83$, $p < .001$, $d = 1.76$. Additionally, participants understood cardinal better than RCV at Time 1, $t(320) = 9.31$, $p < .001$, $d = .67$, and Time 2, $t(320) = 11.76$, $p < .001$, $d = .79$.

There was also a main effect of declared voting system preference on comprehension at both Time 1, $F(2, 317) = 4.36$, $p = .014$, $\eta_G^2 = .011$, and Time 2 $F(2, 317) = 2.82$, $p = .061$, $\eta_G^2 = .009$, though the effect sizes were small. Participants who preferred plurality voting had lower comprehension scores than participants who preferred cardinal at Time 1, $t(126.68) = 2.77$, $p = .007$, $d = .51$, and Time 2, $t(153.86) = 2.30$, $p = .023$, $d = .40$. Participants who preferred plurality also had lower comprehension scores than those who preferred RCV at Time 1, $t(224.86) = 2.31$, $p = .022$, $d = .27$, and Time 2, $t(214.57) = 1.56$, $p = .121$, $d = .19$. There were not significant differences between participants who preferred RCV and cardinal voting at Time 1, $t(116.09) = .99$, $p = .326$, $d = .12$, or Time 2, $t(133.34) = 1.09$, $p = .279$, $d = .14$.

Similar to Study 1, at Time 1 there was an interaction between a participant's declared voting system preference and voting system on their comprehension, $F(4, 634) = 2.75$, $p = .027$, $\eta_G^2 = .010$. Comprehension scores for plurality and RCV did not differ much based on declared voting system preference. In contrast, there were differences in comprehension of cardinal voting; participants who declared cardinal to be the best system had the best comprehension and those who declared plurality had the lowest. There was not a significant interaction at Time 2, $F(4, 634) = 1.61$, $p = .170$, $\eta_G^2 = .005$.

In summary, similar to Study 1, participants understood plurality voting the best overall, and participants who preferred plurality voting understood the voting systems the worst overall. The biggest difference in comprehension was for cardinal voting; participants who preferred plurality understood cardinal voting the worst.

3.2.5 The Impact of Teaching Interventions on Comprehension

To examine learning due to the teaching interventions, we subtracted the comprehension scores of Time 1 from Time 2 (Table F1, Table F2). We ran a 4 (teaching intervention, between subjects) × 3 (voting system, within subject) ANOVA on the change in comprehension scores. We predicted that the retrieval practice teaching intervention would lead to greatest amount of learning, followed by discovery learning, and then re-study, with the control condition having the least learning.

There was a main effect of voting system on improvement in learning scores, $F(2, 634) = 12.39$, $p < .001$, $\eta_G^2 = .024$ (*Figure 5*). The improvement in comprehension was larger for cardinal voting than RCV, $t(320) = 2.03$, $p = .043$, $d = .15$, and plurality, $t(320) = 5.06$, $p < .001$, $d = .38$. Additionally, the improvement in comprehension was larger for RCV than plurality, $t(320) = 3.00$, $p = .003$, $d = .24$.

There was also a main effect of teaching intervention, $F(3, 317) = 5.38$, $p < .01$, $\eta_G^2 = .019$, (*Figure 9*). As expected, the control condition did not exhibit any improvement in learning $t(79) = .75$, $p = .458$. As expected, retrieval practice produced the largest learning gains numerically, and was significantly better than control, $t(164.67) = 3.80$, $p < .001$, $d = .57$, but only marginally better than discovery learning, $t(160.99) = 1.92$, $p = .056$, $d = .31$, and re-study, $t(159.44) = 1.85$, $p = .066$, $d = .28$. The discovery learning and re-study conditions did not significantly differ, $t(146.26) = .22$, $p = .830$, $d = .04$. The re-study intervention significantly outperformed the

control condition $t(154.96) = 2.28$, $p = .024$, $d = .37$, whereas the discovery learning only marginally outperformed the control, $t(151.51) = 1.90$, $p = .060$, $d = .31$.[6]

3.2.6 Voting System Metacognition

Participants self-assessed their understanding of the three voting at the end of the study. We conducted a 3 (voting systems, within-subjects) × 4 (teaching intervention; between-subjects) ANOVA on their judgment of knowledge ratings. There was a significant main effect of voting system, $F(2, 634) = 120.58$, $p < .001$, $\eta_G^2 = .133$ (Table 4). Participants felt that they understood plurality voting better than RCV, $t(320) = 9.08$, $p < .001$, $d = .51$, and cardinal, $t(320) = 14.17$, $p < .001$, $d = .94$, and they felt that they understood RCV better than cardinal, $t(320) = 7.68$, $p < .001$, $d = .47$. Indeed, they did understand plurality the best, but they understood cardinal better than RCV, not the reverse (*Figure 5*). The judgment of knowledge ratings did not differ by teaching intervention, $F(3, 317) = .15$, $p = .093$, $\eta_G^2 = .001$, nor was there an interaction between the teaching interventions and voting system, $F(6 ,634) = .57$, $p < .075$, $\eta_G^2 = .002$.

Table 4. Metacognition judgment of knowledge ratings across voting systems

		Voting System		
Study	*Condition*	Cardinal	Plurality	RCV
2	–	61.87 (28.61)	85.62 (21.70)	74.16 (23.61)
3	Raw+Summary Data	74.50 (24.10)	81.20 (24.50)	74.70 (24.60)
	Raw Data Only	80.20 (22.00)	85.20 (23.30)	73.70 (27.20)
4	Raw+Summary Data	79.30 (23.50)	88.10 (21.90)	77.50 (23.50)
	Raw Data Only	75.00 (26.50)	78.60 (26.90)	75.80 (23.30)
	Summary Data Only	77.90 (25.60)	84.10 (24.60)	76.90 (28.90)

Note: values are means (std. dev.).

The above results examine differences in metacognition across voting systems and

[6] There was a marginal interaction between the learning intervention conditions and the voting system, $F(6, 634) = 1.88$, $p = .083$, $\eta_G^2 = .011$; the most learning occurred in the three teaching interventions conditions (not control), and especially for cardinal and RCV, not plurality which was already near ceiling at Time 1.

conditions, where self-assessments are examined at the aggregate level. Because metacognition is primarily related to how well an individual is assessing their own knowledge, the relationship between judgment of knowledge ratings and comprehension scores for specific voting systems is worth analyzing. A scatterplot showing the relationship between judgment of knowledge ratings and comprehension scores for Studies 2-4 can be seen in Appendix I. An important finding emerges from examining this relationship in Study 2. Although participants on average gave a higher judgment of knowledge rating for plurality, participants with the lowest RCV comprehension scores actually tended to give judgment of knowledge ratings that were higher for RCV than plurality. This provides evidence that those with the poorest understanding of RCV may have the largest gap in accurately assessing their own knowledge, or stated more simply—they are overconfident.

3.3 Discussion

We found a similar pattern of preferences before the teaching intervention as in Study 1. Participants' choices in the WWET favored plurality candidates; however, participants declared that they preferred RCV the most. After the teaching interventions, these patterns persisted, though we found evidence of somewhat greater consistency between revealed and declared preferences, though these patterns persisted.

We also found a similar relationship between voting system understanding and preferences in Study 2 as in Study 1. Participants who preferred alternative voting systems (cardinal, RCV) were more likely to have a better understanding of them at Time 1. However, this relationship became weaker after the teaching intervention.

In line with our hypotheses, retrieval practice showed the most potential for improving understanding of the voting systems, and discovery learning intervention showed the most

potential to change participants' declared preferences. However, the differences between the three interventions were marginal or non-significant. All three interventions produced significantly or marginally more learning than the control condition, and all three produced significantly more change in declared preferences than control condition.

We also found evidence of poor resolution in judgments of knowledge; participants self-assessed themselves as having a better understanding of RCV than cardinal voting, but participants actually understood cardinal voting better than RCV. Additionally, participants who performed the worst on the RCV comprehension tests actually believed they better understood RCV than plurality voting. This provides more evidence that participants' declared preferences are not well aligned with other voting system measures (preferences revelated in the WWET, understanding, and now metacognition).

4.0 Study 3

So far, we relied on the graphical format heavily to assess participants' preferences about voting systems. This format was used extensively for the WWET, but we also used the graphs when explaining the voting systems to participants and asking about their declared voting preferences. These graphs could be viewed as a limitation, or at least an idiosyncrasy of Studies 1 and 2. The main goal of Study 3 is to address some potential concerns about these graphs.

One concern is that our participants may have had difficulty understanding the graphs, or even if the graphs were understandable they may lead people to think about voting in a different way than they normally do since voting data is not typically presented in such graphs. Relatedly, the graphs in Studies 1 and 2 assume that there is just a single dimension on which all candidates and voters' preferences can be described; many real-world group decision making situations are likely more complex. In sum, testing an alternative presentation seems warranted to help provide

convergent evidence of the findings in Studies 1 and 2. To address this concern, in Study 3, we presented participants with the same voting data from Studies 1 and 2 but in a tabular format. This tabular format also ameliorates the single-dimension issue; in the tabular format candidates are rated as higher or lower by voters, and this rating could represent the favorability of a candidate to a voter aggregated across multiple dimensions. Educational materials about voting on the internet also heavily skew towards using a tabular format (e.g., FairVote, 2022; The Center for Election Science, 2021), which may increase the ecological validity of our results.

A second concern with the graphs in Studies 1 and 2 is that participants had to aggregate the data in the graphs to come to a decision. In some real-world situations, such as making a group decision amongst friends and colleagues one would have to aggregate the votes, but in other situations, such as political elections only the people running the election need to do the aggregation. Therefore, it is important to understand users' views about fairness in both situations: when they need to do the aggregation and when they do not.

Aggregation requires different amounts of effort for different voting systems, and the need to implement the aggregation may affect views about fairness differently for different voting systems. For plurality, aggregating involves tallying up how many voters are closest to each candidate, which takes some but not all that much effort. For RCV, aggregating means first doing what is involved for plurality, then deciding whether the instant runoff needs to be enacted (if no candidate has received more than 50% of the vote), and then doing the tally again for the remaining two candidates. Prior to learning about RCV, the idea of doing an instant runoff likely did not occur to the participants, and even after learning about RCV, they may not have realized the situations in which it is necessary or did not feel the motivation to carry it out. For cardinal voting, aggregating means finding the average position of the voters, and then finding which

candidate is closest to the average position, or equivalently, getting a sense of how far away on average the voters are from each candidate. There is considerable research showing that people's visual system can quickly extract summary properties such the centroid of a set of dots in a cluster (e.g., Morgan & Glennerster, 1991; see also Alvarez, 2011; Szafir et al., 2016, for reviews), so perhaps the task is not very hard. But on the other hand, the distributions in the graphs were often bimodal or trimodal, and sometimes the average distances for two candidates were not all that different, so doing the aggregation from the visual stimuli in Experiments 1 and 2 may not have been very easy.

It is possible that participants wanted to implement RCV or cardinal voting, but that they felt like it was too hard from the graphs. Perhaps if the aggregation is provided to participants, they may make different choices in the WWET, and if cardinal voting feels easier, perhaps participants will like it more in their declared preferences. In Study 3, half of the participants were randomized to a condition that displays summary information in addition to raw data. We predicted that if participants have access to the summary information, they will be more likely to use cardinal voting or potentially RCV in the WWET. Additionally, we predicted that the summary may lead participants to view cardinality as fairer. Without the summary participants need to intuitively estimate the averages, which is likely to feel imprecise and therefore unfair, but with the averages given to them it is likely to feel very precise. Furthermore, the participants may view cardinal voting as better able to take into account the voters' preferences because it uses a fine-grained measure of preferences, whereas RCV and plurality use more coarse-grained ordinal and categorical measurement scales. In sum, with the summary, people may prefer cardinal voting, and there may be more agreement between their decisions in the WWET task and their declared voting preferences.

A third concern with the graphs in Studies 1 and 2 is that they do not make the different voting systems very obvious. This could be viewed this as an advantage; the graphs do not make the three voting systems obvious, so if a participant spontaneously tends to use one system presumably this is because that voting system feels natural to them. One concern, though, is that perhaps the graphs may lead people to choose certain voting systems that they would not use in other circumstances without graphs. Another way to get around this issue is to present participants with the three voting systems and let them decide which one to use. This was the strategy we used in Study 3.

Having the voting systems summarized in tables may also influence perceptions of how easy the voting systems are to implement as a result of greater fluency. We predict that participants' judgment of knowledge ratings for voting systems will be greater in the condition where summary tables are present versus when they are absent.

4.1 Method

4.1.1 Participants

We recruited 200 participants through the University of Pittsburgh's introduction to psychology subject pool.

4.1.2 Design, Procedures, and Measures

Participants were randomized to one of two conditions, raw data only, or raw plus summary data. These differences in the stimuli appeared throughout many blocks of the study (WWET, voting system comprehension tests, teaching intervention). *Figure 10* provides a summary of the procedural flow for Study 3. For concision, there are some tasks that we did not analyze here and therefore do not explain in the methods section or in *Figure 10*.[7]

[7] After the voting system comprehension test, participants answered the declared voting system questions again, then learned about the quirks of the voting systems using the material common to all three teaching interventions in

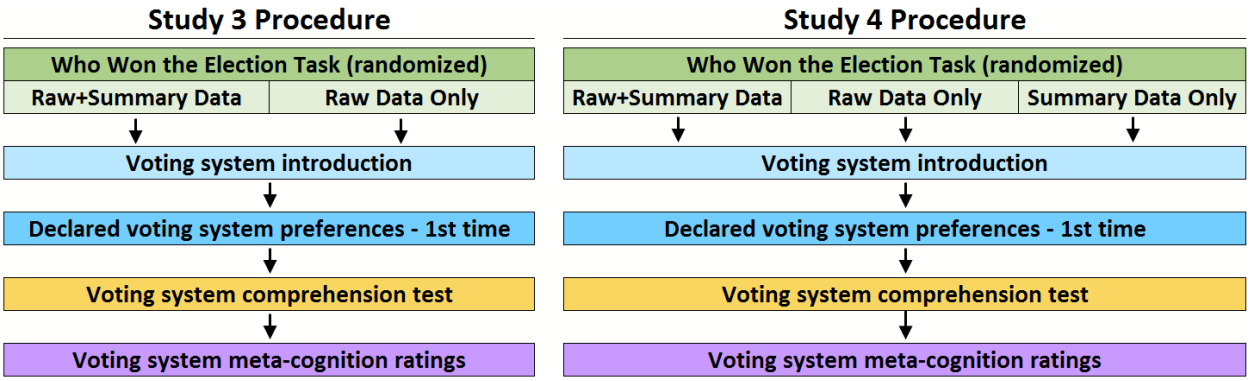


Figure 10. Overview of Procedures for Studies 3 and 4. Note: Blocks that were not analyzed were omitted from this figure.

4.1.2.1 Who Won the Election Task

Leading into the WWET, participants received an introduction to reading the tables (Appendix C1) and were asked 9 questions to verify that they were paying attention and understood how to read the tables.

The Who Won the Election Task worked similarly to the previous two studies, except participants were presented with election results information in a table like the one in *Figure 11* instead of the graph used in Studies 1 and 2 (Appendix C3). Participants in the Raw+Summary condition saw tables just like *Figure 11*, which showed each individual voter's choices at the top of the screen as well as summaries of all the voters at the bottom; participants in the 'Raw Data Only' condition only saw the top half. One important aspect of these stimuli is that they show the voters' choices as if the voters had filled out three separate 'ballots' – one in which they chose their top candidate, one in which they ranked the candidates, and one in which they rated each candidate on a scale of 0-100. So, in essence, the WWET is measuring which of these three ballots participants use to decide which candidate won the election.

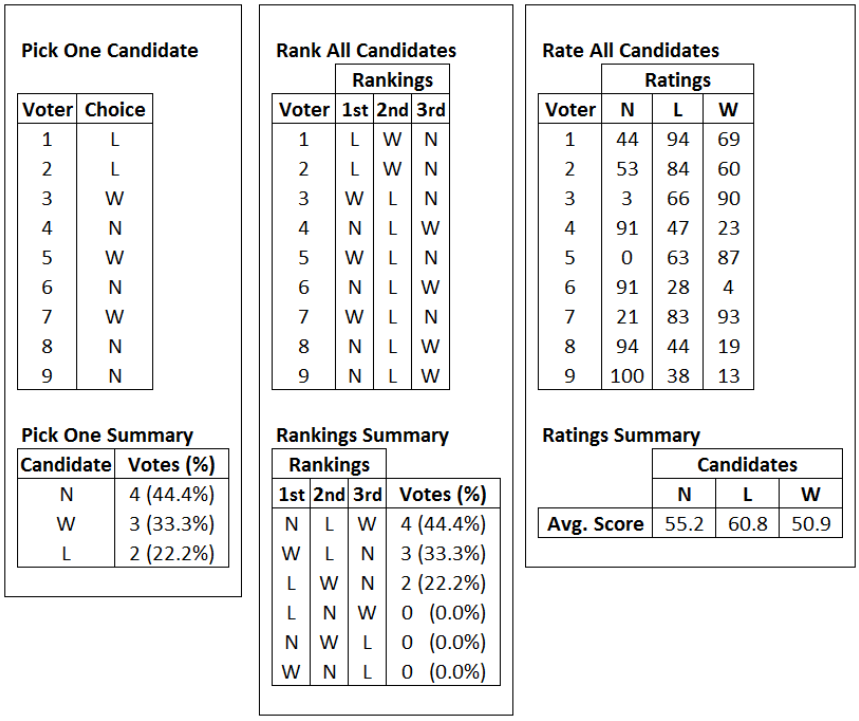
Study 2 (Appendix C7), followed by the declared voting system preferences a third time. The judgment of knowledge ratings occurred afterwards. At the end of the study, participants were asked five additional questions to understand other factors that could contribute to their preferences for different voting systems, such as prior familiarity with the voting systems, and which voting system they perceived would be best for their preferred political party (Appendix C8). These questions were not analyzed in the current report.

The data presented in the tables like in *Figure 11* were derived directly from the graphs in Studies 1 and 2 (e.g., compare *Figure 11* with *Figure 2*). The ratings for each pair of candidates and voters in the "Rate All Candidates" column were derived from first computing the distances between each candidate and voter in the stimuli in Studies 1 and 2, and then linearly rescaling the distance score so that the largest distance was converted into a score of 0, the 'worst' rating from a voter, and the smallest distance was converted into a score of 100, the 'best' rating from a voter. The other two columns were derived from the ratings; thus, the voters' preferences were consistent across all three ballots (e.g., in *Figure 11*, Voter 1 chose K for 'pick one', ranked K the highest, and rated K the highest. Despite the fact that a given voter is consistent across all three ballots, the aggregated decisions for the three voting systems can diverge.

One important aspect of the summary tables is that they clearly show the winner for plurality and cardinal, as these just involve the sums or averages of the scores. The Rankings Summary in *Figure 11* shows how the outcomes for RCV were displayed; it shows the number of voters who selected each of the 6 ranks of candidates. Note that unlike the Pick One Summary for plurality and Ratings Summary for cardinal, the participants still need to decide how to convert this summary into a choice of which candidate won. The reason why we did not clearly show who won in the Rankings Summary is that this would have required explaining rank-choice voting and the instant run-off rule as part of the WWET. In Studies 1 and 2, the three voting systems (including rank-choice voting and the instant run-off logic) were explained after the WWET leading into the declared voting system preferences; participants need to know the names of the three voting systems and how they work to rate them on fairness. So, for consistency with the prior studies, we thought that it made the most sense to let participants decide who won in the WWET and then introduced the instant run-off logic leading into the declared voting system

preferences. (In Study 4, we took the opposite approach.)

With regards to the dependent measures, in addition to answering the three questions for each item like in Study 2 about who should win (rate the three candidates, rate them, and choose one), participants also answered a fourth question which asked "Which table did you primarily use to make your decisions?"



Pick One Candidate

Voter	Choice
1	L
2	L
3	W
4	N
5	W
6	N
7	W
8	N
9	N

Pick One Summary

Candidate	Votes (%)
N	4 (44.4%)
W	3 (33.3%)
L	2 (22.2%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	L	W	N
2	L	W	N
3	W	L	N
4	N	L	W
5	W	L	N
6	N	L	W
7	W	L	N
8	N	L	W
9	N	L	W

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
N	L	W	4 (44.4%)
W	L	N	3 (33.3%)
L	W	N	2 (22.2%)
L	N	W	0 (0.0%)
N	W	L	0 (0.0%)
W	N	L	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	N	L	W
1	44	94	69
2	53	84	60
3	3	66	90
4	91	47	23
5	0	63	87
6	91	28	4
7	21	83	93
8	94	44	19
9	100	38	13

Ratings Summary

	Candidates		
	N	L	W
Avg. Score	55.2	60.8	50.9

Figure 11. Tabular stimuli used in the WWET for Study 3. Participants in the Raw+Summary condition saw figures like this. Participants in the Raw Data Only condition only saw the top half.

The data presented in the three ballots is functionally identical to the data presented in the graph version of the task (*Figure 11*). Each row in the three ballots corresponds to the same individual voter. Voter 1 in *Figure 11* chooses candidate K as their favorite, they also rank candidate K first, followed by A, then S. Additionally, Voter 1 rates K the highest (95), followed by A (54), and then S (8). In the graph version of the WWET, the cardinal system measures closer proximity as better score rating. However, in the tabular version, we inverted this measure, so that higher scores represent a better rating (which we believed is more intuitive for the

participants). This was accomplished by first measuring the distance from each voter to each candidate. Then this array of numbers was run through an inverse scaling function with a range of 0-100, where 100 was the lowest number in the untransformed array, and 0 was previously the largest. The untransformed and transformed data have a near-perfect negative correlation, made imperfect by rounding to whole numbers so that all voter ratings were integers.

In addition to answering the three ballot questions for each WWET item (like in Study 2), participants also answered a fourth question which asked, "Which table did you primarily use to make your decisions?" These questions will be separately examined in the results section.

4.1.2.2 Other Tasks

The other tasks were very similar to those in Study 2, though were updated to accommodate the new tabular stimuli. Furthermore, the two conditions of the tabular stimuli (Raw+Summary vs. Raw Data Only) were carried through the other tasks. Other differences from Study 2 or important clarifications are as follows.

For the declared voting system preferences questions, in addition to asking the same questions in Study 2 about the fairness of the voting systems, participants also answered three questions about which voting system they would prefer to use. The reason we asked about both fairness and prefer to use is because an individual could feel that one voting system is the most fair, but could still prefer to use another voting system (e.g., if they feel that the first is too complicated).

For the comprehension test questions, participants were only shown the relevant part of the tabular stimuli. For example, when participants were asked which candidate would win according to RCV, they were only shown the middle column stimuli like in *Figure 11*; the left and right columns were whited out. Participants in the Raw Data Only condition when asked

about who would win according to RCV only saw the top-middle section of stimuli like in *Figure 11*.

4.2 Results

4.2.1 Participants

We removed 13 participants who completed the study on their phones instead of a computer, and 55 participants were removed because they did not correctly answer the questions ensuring they could read the tables. In all, 217 participants submitted valid data for analysis.

4.2.2 Validity Check and Ballot Agreement

For three of the WWET items there was a candidate option in which each of the three voting systems agreed was not the winner. As a measure of attention/comprehension, we examined how often the "loser" candidates were selected. Participants selected the loser candidate 3.07% of the time (raw+summary data condition: 1%; raw data only: 5%).

Participants completed three ballots for all items in the WWET as well as for declared voting system preferences. We found that the three ballots perfectly matched most of the time for both WWET ballots choices (92.72%) and declared voting system preferences (Time Point 1: 86.18%; Time Point 2: 85.71%; Time Point 3: 86.64%). When the three ballots did not match, then typically two did. In less than 1.30% of cases did the ballots problematically disagree, in which case the response for that item would be removed from analysis.

Together, the high consistency in ballot matches and low frequency of choosing candidates that were not supported by any of the voting systems, suggests that, in general, participants took the study seriously.

4.2.3 Voting System Choices

4.2.3.1 WWET Choices

The descriptive statistics for the choices in the WWET can be seen in *Figure 3*. Like Studies 1 and 2, participants chose the plurality option most frequently, followed by the cardinal option, and then RCV. This pattern of choices was similar across the two conditions; there were no significant differences across the conditions for cardinal [$t(214.22) = .85$, $p = .397$, $d = .11$], plurality [$t(214.85) = .55$, $p = .581$, $d = .07$], or RCV [$t(189.04) = .90$, $p = .370$, $d = .12$] candidates.

In the WWET in Study 3 participants also stated which of the three tables ("rate all candidates"/cardinal, "pick one candidate"/plurality, or "rank all candidates"/RCV) they primarily used to make their decision. Table 5 shows the percent of time that participants said they used each of the three tables in the WWET by condition. Participants in the raw data only condition reported using the 'Rank'/RCV table most frequently (47%), followed by the 'Rate'/cardinal (36%) and 'Pick One'/plurality tables (17%). Contrastingly, participants in the raw and summary data condition reported using the 'Rate'/Cardinal (40%) table most frequently, followed by 'Pick One'/plurality (28%) and 'Rank'/RCV (17%) tables.

Table 5. Percent of Time Participants Reported Using Each of the Three Tables in Studies 3 and 4

	Table That Participants Said They Used					
	Study 3			**Study 4**		
Stimuli Condition	'Rate'/ Cardinal	'Pick one'/ Plurality	'Rank'/ RCV	'Rate'/ Cardinal	'Pick one'/ Plurality	'Rank'/ RCV
Raw Data Only	36%	17%	47%	39%	14%	47%
Raw+Summary Data	40%	28%	17%	45%	17%	37%
Summary Data Only	–	–	–	38%	19%	43%

Table 6. Probability that a Choice in the WWET is Consistent with Each Voting System Given that a Participant Said they Used a Given Table in Studies 3 and 4

	Table That Participants Said They Used					
	Study 3			**Study 4**		
Choice is Consistent with	'Rate'/ Cardinal	'Pick one'/ Plurality	'Rank'/ RCV	'Rate'/ Cardinal	'Pick one'/ Plurality	'Rank'/ RCV
Raw Data Only Condition						
Cardinal	49%	6%	42%	48%	6%	42%
Plurality	45%	83%	51%	46%	84%	58%
RCV	25%	39%	29%	23%	29%	33%
Raw+Summary Data Condition						
Cardinal	77%	9%	27%	81%	3%	21%
Plurality	22%	89%	71%	16%	94%	58%
RCV	22%	34%	28%	24%	31%	69%
Summary Data Only Condition						
Cardinal	–	–	–	83%	3%	19%
Plurality	–	–	–	16%	92%	59%
RCV	–	–	–	25%	30%	62%

Note: Some choices in the WWET support two voting systems, so the columns add up to over 100%.

Table 6 shows the probability that, if a participant said they were using a given table, they made a choice consistent with this that table. The columns sum to more than one because sometimes choices were consistent with two voting systems. Participants who said that they used the 'Pick One'/plurality table plurality tended to make decisions consistent with plurality about 85% of the time. An independent samples t-test comparing the percent of choices for each participant that were consistent with plurality if the participant said that they used the plurality table did not find a significant difference between conditions [$t(127.89) = .29$, $p = .772$, $d = .04$].

Participants who said that they used the 'Rate All Candidates' / cardinal table tended to make decisions consistent with cardinal voting in the Raw+Summary condition (77%), but less so in the Raw Data Only condition (49%), and there was a significant difference in consistency conditions [$t(178.85) = 4.20$, $p < .001$, $d = .59$]. In the Raw+Summary condition, it should have been trivially easy to implement cardinal voting because participants could just look at the

means, but in the Raw Data Only condition participants had to estimate the means, so it is not all that surprising that they were less accurate.

Most importantly, in both conditions, when participants said that they were using the 'Rank All Candidates' / RCV table, their choices were only consistent with RCV about 29% of the time; there was not a difference between conditions [$t(182.99) = .76$, $p = .448$, $d = .11$]. Instead of implementing RCV, most of their choices were consistent with plurality. In sum, even when people think that they are using RCV, they tend to fail to implement the instant runoff logic and resort to making a decision based on the voters' first choices, effectively implementing plurality voting. This finding fits with the findings from Studies 1 and 2 that even when participants declared RCV to be the most fair, they rarely used RCV in the WWET and instead primarily used plurality. The current finding goes even farther and reveals that even when participants say that they are using the rankings *in the WWET*, they still primarily use plurality.

4.2.3.2 Declared Voting System Preferences

We conducted a 2 (condition; between-subjects) × 3 (voting systems; within-subjects) ANOVA on subjects' declared fairness ratings (see Table 7 for 0-100 ratings; see *Figure 3* for the single choice measure). There was no main effect of condition [$F(1, 215) = .14$, $p = .709$, $\eta_G^2 = .000$]. There was a main effect of voting systems [$F(2, 430) = 3.06$, $p = .048$, $\eta_G^2 = .010$], which should be interpreted in light of a significant interaction [$F(2, 430) = 6.52$, $p = .002$, $\eta_G^2 = .021$]. There was no difference in fairness ratings for RCV across the two conditions [$t(208.58) = .88$, $p = .382$, $d = .12$]. Participants in the Raw+Summary condition rated plurality as more fair than those in the Raw Data Only condition [$t(209.40) = 2.45$, $p = .015$, $d = .32$]. In contrast, participants in the Raw+Summary rated cardinal voting as less fair than those in the Raw Data Only condition [$t(214.06) = 2.64$, $p = .009$, $d = -.35$]. This last finding about cardinality runs

directly opposite to our hypothesis that people would view cardinal voting as more fair if they are given a summary and do not have to estimate the means on their own.

Table 7. Mean (SD) Declared 'Fairness' Ratings and 'Would Like to Use' Ratings on 0-100 Scale

	Study 3			Study 4		
Condition	Cardinal	Plurality	RCV	Cardinal	Plurality	RCV
	Declared 'Fairness' Ratings					
Raw Data Only	71 (23)	53 (24)	64 (24)	68 (22)	52 (25)	63 (23)
Raw+Summary Data	62 (25)	61 (23)	67 (23)	59 (30)	59 (26)	69 (18)
Summary Only	–	–	–	63 (21)	57 (28)	71 (20)
	'Would Like to Use' Ratings					
Raw Data Only	67 (25)	55 (30)	65 (27)	59 (28)	49 (27)	64 (26)
Raw+Summary Data	61 (24)	58 (27)	66 (25)	57 (32)	58 (28)	65 (24)
Summary Only	–	–	–	55 (28)	53 (30)	69 (26)

4.2.3.3 Declared Voting System "Fairness" versus "Would Like to Use" Ratings

Participants rated each of the voting systems on both how fair they believed they were and how much they would like to use them (*Figure 7*). Without *a priori* predictions, we were curious if participants perceived a gap between how fair a voting system was and how much they desired to use it. To examine the possibility of a gap in these ratings, we calculated paired t-tests for the three voting systems. On average, participants gave a higher rating for fairness than desire to use for cardinal voting, $t(216) = 2.27$, $p = .024$, $d = .15$, but there was not a difference for the other two voting systems. We also conducted a 2 (condition; between-subjects) × 3 (voting system; within-subjects) on the difference between the fairness ratings minus would like to use ratings; however, there were no significant results. One reason cardinal may be fairer than it is desirable to use is that it offers a fine granularity that may potentially reflect more nuanced and hence accurate preferences. However, when it comes to using cardinal, people may be wary that everyone's votes would be honest and accurate.

4.2.3.4 Choice Consistency

A LDA found that choices in the WWET were only able to predict declared voting system preferences for 34.42% of participants, which is not significantly better than chance (*Figure 4*). The hit rate for three groups was above chance for plurality (65.91% ; $z = 4.67$, $p < .001$), at chance for RCV (31.18%; $z = -2.35$, $p = .019$) and for cardinal voting (20.51% ; $z = -.37$, $p = .711$; see Appendix E for complete analysis results).[8]

4.2.4 Comprehension and Declared Voting System Preferences

See *Figure 5* for descriptive statistics of the comprehension scores of the three graphs. We expected that the comprehension scores in the Raw+Summary condition would be near ceiling for the cardinal and plurality table because both of these clearly identify the winner, and indeed these were near ceiling.

We conducted a one-way ANOVA (voting system; within-subjects) for just the Raw Data Only condition. There was a main effect of voting system, $F(2, 196) = 89.03$, $p < .001$, $\eta_G^2 = .335$. Participants had higher comprehension scores on the plurality test than cardinal [$t(100) = 7.14$, $p < .001$, $d = .92$], and RCV [$t(100) = 16.11$, $p < .001$, $d = 2.15$], and higher on the cardinal test than RCV [$t(100) = 8.67$, $p < .001$, $d = 1.08$]. This pattern of better comprehension for plurality, then cardinal, then RCV is the same pattern that was found in Studies 1 and 2, and suggests that the difficulty understanding cardinal and especially RCV from raw data is not due to just a quirk of the graphical format in Studies 1 and 2.

4.2.5 Voting System Metacognition

Participants rated how well they understand each of the voting systems (Table 4).

A 2 (condition; between-subjects) × 3 (voting systems; within subjects) ANOVA on participants'

[8] We conducted LDA's limited to each condition as well and found similar results. The overall prediction rate for raw+summary data (37.72%) and raw data only (35.64%) were also non-significant. See Appendix E for more details.

judgment of knowledge ratings did not any effect of condition [$F(1, 215) = 1.09$, $p = .297$, $\eta_G^2 = .004$], voting systems [$F(2, 430) = 1.67$, $p = .190$, $\eta_G^2 = .002$], nor was there an interaction [$F(2, 430) = 2.45$, $p = .088$, $\eta_G^2 = .003$]. The lack of differences is striking. Participants' comprehension scores clearly indicate a hierarchy for the difficulty of the three voting systems. However, at the group level, participants felt they were equally proficient in their understanding for each system. This provides further evidence of a poor resolution of how well participants actually understand the different voting systems.

The relationship between judgment of knowledge ratings and comprehension scores for specific voting systems is shown in Appendix I. This figure shows how RCV judgment of knowledge ratings are equal or, in some places, even higher than judgment of knowledge ratings for cardinal. This finding is in stark contrast to evidence presented earlier that participants have a poorer understanding of RCV compared to cardinal voting.

4.3 Discussion

We find a similar pattern of preferences in the WWET; like Studies 1 and 2 participants chose the plurality option most frequently, followed by the cardinal option, and then RCV. In Study 3 participants also reported which table they were primarily using in the WWET. Participants in the raw data only condition reported using the 'Rank'/RCV table most frequently whereas for participants in the raw and summary data condition they most frequently reported using the 'Rate'/Cardinal table.

Counter to our predictions, cardinal voting was perceived as being fairer when summary information was omitted. One reason for this pattern may be due to the awareness that a wide distribution of ratings seen in the raw data are being aggregated into a single number in the summary data (i.e., the average). The question then, is whether this pattern of results hold when

participants only see summary results. In Study 4, we predict that participants will prefer cardinal voting more in the condition with only summary tables, versus when participants see raw data and summary tables.

In Study 3, for the first time, we examined if there is a gap between how fair the voting systems are viewed by participants versus how desirable they are to use. We found some evidence that cardinal voting was viewed as being more fair than it was desirable. It is unclear why this may be the case. One explanation is that the finer granularity of preferences in cardinal can reflect more accurate preferences—which is viewed as being fairer—but when using cardinal voting, participants may be skeptical that others who honestly relay their preferences.

Unlike Studies 1 and 2, when we used a tabular version of the WWET in Study 3, we could not predict declared voting system preferences from decisions in the WWET although we were fairly successful for predicting individuals with a declared preference for plurality from their plurality preferences in the WWET. The other two systems were at chance levels.

Like prior studies, in the Raw Data Only condition we found that participants best understood plurality, followed by cardinal, and the poorest understanding for RCV. Strikingly, participants performed equivalently on the RCV comprehension test in both the Raw Data Only and the Raw+Summary Data condition. This suggests that the issue participants are primarily having with RCV is not from being able to gather proportional data about the order rankings, but rather with the ability to implement the instant-runoff logic specifically. We will examine this issue directly in Study 4.

We found more evidence of poor resolution in judgments of learning; participants self-assessed themselves as having an equivalent understanding of RCV as cardinal and plurality voting, but participants actually understood cardinal and plurality voting better than RCV. We

did not find that judgment of knowledge ratings were affected by whether or not summary tables were present.

5.0 Study 4

In Studies 1 and 2, participants rated cardinal voting the lowest on fairness, and we hypothesized that this might be due to feeling that it was difficult to aggregate the raw data and find the winner according to cardinal voting. To test this hypothesis, in Study 3, we presented election data in a numerical format displayed in tables, and half the participants also received a summary of the election results. Overall, cardinal voting was viewed as fairer in Study 3 with tables than Studies 1 and 2 with graphs, but, contrary to our hypothesis, we found that participants thought that cardinal voting was more fair without the summary rather than with the summary. One reason for this finding is that our participants may have felt that taking an average of the ratings does not do a good job of capturing the raw voting data. Many of the distributions we used involved bimodal or trimodal distributions, so they may have felt that an average does not really represent the voting data well.

To test this hypothesis, in Study 4 we added a third 'Summary Only' condition in which participants only saw the summary results and did not see the raw voting data. We predicted that participants would prefer cardinal voting more in the condition since they cannot see how the average does a poor job compared to the Raw+Summary condition. Alternatively, because the Raw+Summary condition was rated as less fair than the Raw Only condition in Study 3. It is possible that the new Summary Only condition will continue this trend of being rated even worse. One reason could be that when only an average is provided people could feel like they don't understand how it connects to the raw votes – many different distributions of votes could produce the same average. In contrast, the summary for plurality voting is extremely

straightforward – it is just a sum of the votes – and the summary for RCV, which is described next, is also quite straightforward.

Another change we made in Study 4 was that we changed the design of the summary tables for RCV to display the instant-runoff logic so that participants can clearly see which candidate was eliminated and which candidate was the winner. In Study 3, the summary table for RCV showed the proportion of voters who preferred specific orderings of the candidates, but the participants in the WWET still had to think about what to do with the ranks. Even participants who believed that RCV was the fairest and those who said that they were using the RCV tables tended to simply resort to the winner according to the first choice, so in effect they implemented plurality voting from the ranks rather than actually implementing RCV. In Study 4 we explained the instant-runoff logic prior to the WWET, and during the WWET the summary table clearly showed who won if the instant-runoff was implemented.

We hypothesized that in the Summary Only condition, when the RCV winner was clearly labeled, that people would feel that RCV is not very fair, because it sometimes diverges from plurality voting which participants have tended to use in the WWET. In contrast, in the Raw Only condition, participants might rate RCV as fairer because they might not realize how it diverges from plurality voting. Alternatively, people might notice that it diverges from plurality and decide that they like RCV, in which case they might choose to use it in the WWET and might also rate it as fair.

5.1 Method

5.1.1 Participants

We recruited 200 participants through the University of Pittsburgh's introduction to psychology subject pool.

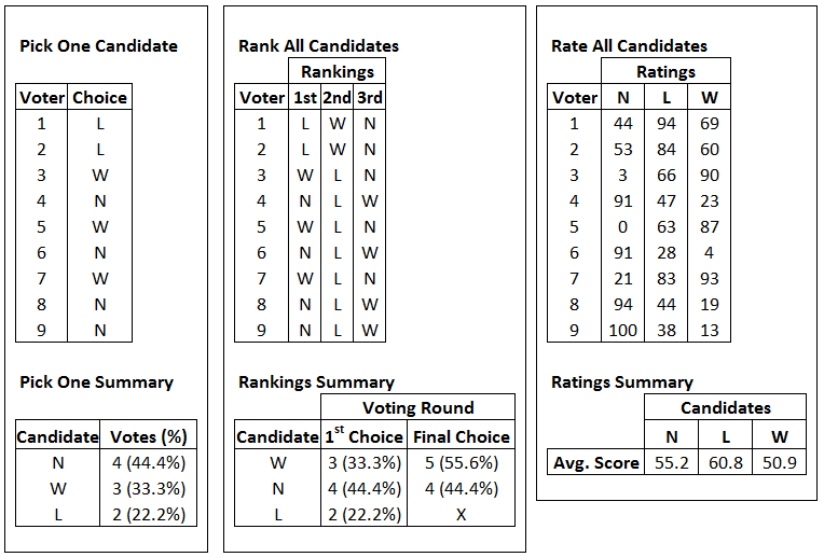
5.1.2 Design, Procedures, and Measures

Figure 10 provides a summary of the procedural flow for Study 4. There were two important changes compared to Study 3. First, in addition to the Raw+Summary Data and Raw Data Only conditions, participants could also be randomized to a Summary Data Only condition; in this condition they only saw the bottom half the stimuli like in *Figure 12*.

Second, the 'Rankings Summary' part of the stimuli in the WWET was changed; compare *Figure 12* vs. *Figure 11*. The Rankings Summary in Study 4 shows both the distributions of the voter's first choices across the three candidates, and the distribution for the final choice after eliminating one of the candidates due to the instant run-off. Participants in the Summary Data condition and the Raw+Summary condition learned about the instant run-off logic as part of the WWET in Study 4, rather than afterwards in Study 3, and these participants were asked an additional question, with feedback upon answering, to make sure they knew how to read the Rankings Summary.

Appendices D1-D5 provide full details of the study including some blocks of the study that are not discussed in this presentation or analyzed.[9]

[9] After judgment of knowledge ratings, participants were asked six open-ended free response questions about the pros and cons for each of the three voting systems. At the end of the study, a number of individual differences were measured (MFQ, CRT, NFCC). We had planned to assess the relations between the individual differences and declared voting system preferences like in Study 1, but failed to realize that since there were three conditions in Study 4, it would make sense to analyze these relations separately, and we feel that the study is underpowered to analyze these separately for the three conditions.



Pick One Candidate

Voter	Choice
1	L
2	L
3	W
4	N
5	W
6	N
7	W
8	N
9	N

Pick One Summary

Candidate	Votes (%)
N	4 (44.4%)
W	3 (33.3%)
L	2 (22.2%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	L	W	N
2	L	W	N
3	W	L	N
4	N	L	W
5	W	L	N
6	N	L	W
7	W	L	N
8	N	L	W
9	N	L	W

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
W	3 (33.3%)	5 (55.6%)
N	4 (44.4%)	4 (44.4%)
L	2 (22.2%)	X

Rate All Candidates

	Ratings		
Voter	N	L	W
1	44	94	69
2	53	84	60
3	3	66	90
4	91	47	23
5	0	63	87
6	91	28	4
7	21	83	93
8	94	44	19
9	100	38	13

Ratings Summary

	Candidates		
	N	L	W
Avg. Score	55.2	60.8	50.9

Figure 12. Tabular stimuli used in the WWET for Study 4. Participants in the Raw+Summary condition saw figures like this. Participants in the Raw Data Only condition only saw the top half, and participants in the Summary Data Only condition only saw the bottom half. The Rankings Summary is different for Study 4 compared to Study 3.

5.2 Results

5.2.1 Participants

We removed 4 participants who completed the study on their phones instead of a computer. Additionally, we removed 30 participants (18% of the remaining sample; 11% from the Raw Data Only condition, 18% from the Raw+Summary condition, and 25% from the Summary Only condition) because they did not correctly answer the questions ensuring they could read the tables. In all, 169 participants submitted valid data for analysis.

5.2.2 Validity Check and Ballot Agreement

It appears that participants took the study seriously. Participants only selected the loser candidate in the WWET in 4.54% of cases (Raw+Summary: 3%; Raw Data Only: 8%; Summary Data Only: 2%). When answering the three questions (ratings, rankings, and pick one), they matched perfectly for 92.25% of cases for the WWET, and 89.94% of cases for the declared preferences. Across the three conditions, the three questions entirely diverged at maximum only

.60% of cases, so few cases had to be dropped from analysis.

5.2.3 Voting System Choices

5.2.3.1 WWET Choices

The descriptive statistics for the choices in the WWET can be seen in *Figure 3*. Like all prior studies, in the Raw Only condition, participants chose the plurality option most frequently, followed by the cardinal option, and then RCV. However, in the Summary Only condition and the Raw+Summary condition, RCV was much higher. This makes sense in that these are the only two conditions across all studies in which participants were taught the instant run-off logic as one option in the WWET. At the same time, participants could have dismissed the RCV winner when they noticed that it diverged from the plurality winner in the Raw+Summary and Summary Only conditions, but nevertheless they often chose the RCV winner. A three-group ANOVA found that there were differences in the extent of making RCV choices in the WWET task based on condition, $F(2, 166) = 15.00$, $p < .001$, $\eta_G^2 = .153$. Participants were less likely to make RCV choices in the Raw Only condition than Raw+Summary [$t(75.34) = 3.51$, $p < .001$, $d = .72$] and Summary Only [$t(87.68) = 5.75$, $p < .001$, $d = 1.05$]; there was no difference between Raw+Summary vs. Summary Only [$t(102.76) = 1.51$, $p = .133$, $d = -.30$]. There was no influence of condition on making cardinal [$F(2, 166) = 1.07$, $p = .345$, $\eta_G^2 = .013$] nor plurality choices [$F(2, 166) = 1.41$, $p = .248$, $\eta_G^2 = .017$].

We then examined the percent of times that participants said that they used each of the three tables (Table 5). Like in Study 3, participants in the raw data only condition reported using the 'Rank'/RCV table most frequently (47%), followed by the 'Rate'/cardinal (39%) and 'Pick One'/plurality tables (14%). Participants in the raw and summary data condition reported using the 'Rate'/Cardinal (45%) table most frequently, followed by 'Rank'/RCV (37%) and 'Pick

One'/plurality (17%) tables. Additionally, participants in the summary only data condition reported using the 'Rank'/RCV (43%) most frequently, followed by the 'Rate'/cardinal (38%) and 'Pick One'/plurality (19%) tables.

We next examine the probability that if participants said they were using a given table, that they made a choice consistent with this that table (see Table 6). Participants who said that they used the 'Pick One'/plurality table plurality tended to make decisions consistent with plurality around 90% of the time; participants in all three conditions were likely to have matching choices if they noted the plurality ballot was primarily used, and choice consistency did not significantly differ across conditions, $F(2, 92) = 1.75$, $p = .180$, $\eta_G^2 = .037$.

Participants who said that they used the 'Rate All Candidates' / cardinal table tended to make decisions consistent with cardinal voting in the both Raw+Summary condition (81%) and the Summary Data Only condition (83%), but less so in the Raw Data Only condition (48%), and this lead to a significant difference in consistency conditions across the groups, $F(2, 142) = 13.79$, $p < .001$, $\eta_G^2 = .163$. As we mentioned in Study 3, this finding is not surprising given that participants that had to estimate the answer in the Raw Data Only conditions, whereas the summary tables presented the correct answer in the other two conditions. The more interesting question is if we find the same pattern in Study 4, in which the summary tables for RCV display the elimination logic (in contrast to Study 3, which did not implement this logic).

Unlike Study 3, we found that participants who said that they used the 'Rank'/RCV table tended to make decisions consistent with RCV in both Raw+Summary condition (69%) and the Summary Data Only condition (62%), but less so in the Raw Data Only condition (33%), and this lead to a significant difference in consistency conditions across the groups, $F(2,137) = 38.91$, $p < .001$, $\eta_G^2 = .372$. By using tables that clearly displayed the RCV winner in Study 4, we were

able to dramatically increase the choice consistency between reported RCV table use and choosing a candidate consistent with RCV.

5.2.3.2 Declared Voting System Preferences

We hypothesized that participants did not think that cardinal voting was fair when they saw the raw and summary data because the summary average does not capture the distribution of voters preferences. Therefore, we predicted that participants would prefer cardinal voting more in the Summary Only condition than the Raw+Summary condition. Alternatively, participants in Study 3 thought the Raw+Summary was less fair than the Raw Only for cardinal voting, so it is possible that this trend would continue – participants might feel that a Summary Only presentation of cardinal voting is even less fair than Raw+Summary. This could happen if they worry that the average is not a good representation of the voters' preferences, despite not actually knowing their preferences in the Summary condition (since the raw data is absent).

In Studies 1 and 2, participants felt that RCV was the fairest voting system despite rarely using it in the WWET and despite not understanding it well in the comprehension questions. It seemed that what was really happening was that they thought RCV was fair but only used the first-choice picks, essentially defaulting back to plurality voting. We hypothesized that in Study 4 when the summary clearly showed the RCV winner after implementing the instant runoff logic, that participants would realize that RCV can lead to a different winner than plurality, and that this realization may make them feel that RCV is less fair. Alternatively, it could be that when participants are helped to understand RCV even more by clearly showing the winner, they feel that it is fairer, despite the fact that it could disagree with plurality. However, an independent samples *t*-test revealed no difference in fairness ratings for cardinal voting between the Summary Data Only and the Raw+Summary data conditions, $t(88.69) = .44$, $p = .440$, $d = .16$.

We also conducted a 2 (condition, between-subjects) x 3 (voting systems; within-subjects) ANOVA to test for differences in fairness ratings (see in Table 7). The fairness ratings differed across voting systems, $F(2, 332) = 9.53$, $p < .001$, $\eta_G^2 = .038$. Participants rated plurality as less fair than both cardinal [$t(168) = 2.65$, $p = .009$, $d = .29$] and RCV [$t(168) = 4.19$, $p < .001$, $d = .48$]; cardinal was rated as marginally less fair than RCV [$t(168) = 1.68$, $p = .095$, $d = .18$]. We did not find a significant difference in fairness ratings across condition [$F(2, 166) = .41$, $p = .667$, $\eta_G^2 = .002$]. Additionally, we found a marginal interaction between condition and voting systems [$F(4, 332) = 2.31$, $p = .058$, $\eta_G^2 = .019$].

5.2.3.3 Declared Voting System "Fairness" versus "Would Like to Use" Ratings

We follow-up the analysis from Study 3 that examined if participants perceived a gap between how fair a voting system was and how much they desired to use it. Table 7 shows the mean "would like to use" ratings as well as the "fairness" ratings. Similar to Study 3, on average, participants had a higher rating for how fair they perceived cardinal voting relative to how much they would like to use it ($t(168) = 3.57$, $p < .001$, $d = .27$). A similar marginal trend was found for plurality ($t(168) = 1.88$, $p = .062$, $d = .14$), where participants thought plurality was more fair than it was desirable. There was no difference for RCV [$t(168) = .78$, $p = .436$, $d = .06$]. Replicating the cardinal result, lends credence to the idea that using a 0-100 rating scale is viewed as more fair than it is desirable to use.

5.2.3.3 Choice Consistency

We conducted separate LDAs for each condition using the choices in the WWET to predict participants' declared fairness ratings (Table 8). In the Raw Data Only condition, the hit rate for choosing Plurality as the fairest was significant; as can be seen in *Figure 4*, participants who chose plurality as the fairest tended to make WWET choices in line with plurality voting, which

was seen in previous studies. A new finding was that in the Raw+Summary and Summary Only conditions, the hit rate for choosing RCV was significant; participants who declared RCV as the fairest tended to make WWET choices in line with RCV. This has never happened in prior studies and presumably the reason for the difference is that this is the only study in which the instant run-off logic was explained to participants prior to the WWET and the RCV winner was clearly displayed in the WWET. In sum, this finding shows that if participants are taught about the instant-runoff logic of RCV and if the logic is implemented for participants, then there is in fact consistency between people who think that RCV is fair and those that choose to use it when deciding who won an election. In prior studies even when participants thought that RCV was the fairest, they did not implement it spontaneously in the WWET.

Table 8. LDA Hit Rate by condition for Study 4

	Voting_System		
Condition	Cardinal	Plurality	RCV
Raw+Summary Data	.33	.42	.65**
Raw Data Only	.07	.62*	.23
Summary Data Only	.36	.40	.57**

Note: *$p < .05$; **$p < .01$,

5.2.4 Comprehension and Declared Voting System Preferences

See *Figure 5* for descriptive statistics of the comprehension scores of the three graphs. We expected that the comprehension scores to be near-ceiling level in the Raw+Summary and Summary Only conditions because the summary tables clearly labeled the winning candidate, and indeed they were near ceiling. We conducted a one-way ANOVA (voting system; within-subjects) for just the Raw Data Only condition. The comprehension scores differed across the three voting system tests, $F(2, 118) = 50.73$, $p < .001$, $\eta_G^2 = .325$. Like all the prior studies, when presented with only raw data, participants understood plurality better than cardinal [$t(61) = 5.24$,

$p < .001$, $d = .76$], and RCV [$t(61) = 10.53$, $p < .001$, $d = 1.85$], and they understood cardinal better than the RCV test [$t(61) = 6.02$, $p < .001$, $d = 1.03$]. There was no main effect of declared preferences $F(2, 59) = .01$, $p = .994$, $\eta_G^2 = .000$ nor an interaction [$F(4, 118) = 1.52$, $p = .202$, $\eta_G^2 = .028$].

5.2.5 Voting System Metacognition

We conducted a 3 (condition; between-subjects) x 3 (voting system; within-subjects) ANOVA on participants' judgment of knowledge ratings (Table 4). The judgment of knowledge ratings differed across the three voting systems, $F(2, 332) = 7.83$, $p < .001$, $\eta_G^2 = .015$. Participants felt that they understood Plurality ($M = 83.27$, $SD = 24.90$) better than cardinal ($M = 77.27$, $SD = 25.24$) [$t(168) = 3.19$, $p = .002$, $d = .24$] and RCV ($M = 76.64$, $SD = 25.21$) [$t(168) = 3.29$, $p = .001$, $d = .26$]. There was no difference between RCV and cardinal voting, $t(168) = .34$, $p = .731$, $d = .03$. The judgment of knowledge ratings did not differ across the three conditions, $F(2, 166) = .91$, $p = .405$, $\eta_G^2 = .007$ and there was no interaction, $F(4 ,332) = .73$, $p = .568$, $\eta_G^2 = .003$. Similar to Study 3, we found that participants believed they understood cardinal and RCV at equivalent levels, despite evidence to the contrary.

Examining the relationship between comprehension scores and judgment of knowledge ratings (Appendix I), it is clear that participants in the Raw Data Only condition falsely believed they had a better understanding of RCV than cardinal, especially at the lower levels of RCV comprehension. This evidence is in line with the findings from Studies 2 and 3; participants are not accurately self-assessing their understanding of RCV.

5.3 Discussion

Like the three prior studies, we found that participants most frequently made plurality-consistent decision in the WWET. However, unlike the prior three studies, we found evidence

that people with declared preferences for RCV, indeed made RCV-consistent choices in the WWET much of the time—as long as summary information for voting systems was present.

Counter to our first prediction that participants would give the highest fairness ratings for cardinal voting in the Summary Only condition, participants gave the highest fairness ratings for cardinal voting in the Raw Only condition, then Raw+Summary, and lowest in the Summary Only condition. It is possible that by only having raw data present, participants were especially attuned to the complexity in preferences and that cardinal voting was the only system to capture those preferences. However, when participants begin examining how these complex preferences are aggregated into a single output, cardinal voting becomes less desirable—especially when only summary information is seen and not the raw preference data itself.

Our second prediction was that participants would give higher fairness ratings for RCV in the Raw Only condition than the Summary Only condition. Counter to this prediction, we found that participants gave the highest fairness ratings for RCV in the Summary Only condition, and lowest in the Raw Only condition. We suspect that two factors may have been in play. First, instead of being troubled that RCV diverges from plurality, participants may have decided that indeed they like how RCV chooses a winner; rank choice voting prioritizes winners who are majority supported, which is not a requirement of plurality. Second, participants may have had an illusory feeling that they understood RCV well since the summary table provided the clear winner after the instant-runoff was implemented. In Study 4 participants rated their meta-cognitive understanding for RCV as similarly high as cardinal voting, despite the fact that Studies 1-3 studies show that our participants struggled much more to understand RCV than cardinal.

Like all the prior studies, when presented with only raw data, participants understood

plurality best, followed by cardinal, and RCV was the least understood. Participants in the conditions with summary data were finally able to do well on the RCV test, once the logic of the system was clearly spelled out. However, these results highlight how much participants have struggled with implementing RCV on simple test items.

The metacognition results were congruent with those in Studies 2 and 3; we found that participants believed they understood RCV at levels that did not match their comprehension scores.

6.0 General Discussion

In four studies, we evaluated how participants' preferences about voting systems, the consistency in these preferences when measured in different ways, and whether their understanding of the voting systems and individual differences predicted their voting system preferences. Additionally, we tested educational interventions, which improved participants' understanding of the voting systems.

6.1 Summary of Results

6.1.1 Choices and Choice Consistency

We measured participants' preferences about voting systems two ways. First, we developed the 'Who Won the Election Task' (WWET), in which participants were shown the results of a hypothetical election in which a group of people were voting on which of two or three candidates to hire. In most of the WWET cases participants were shown 'raw data' about each voter's preferences regarding each candidate. Different implementations of the WWET accomplished this in different ways (ratings, rankings, choices of a single candidate, or preferences on a one-dimensional scale), but the core task of the participants was the same – to decide which candidate should be selected based on the voters' preferences. This allowed us to calculate the

degree to which participants' choices agreed with the three voting systems. Second, we also asked participants their 'declared preferences' – how fair they thought each of the three voting systems were after formally being taught the voting systems.

Across all the studies, participants' choices in the WWET were most consistent with plurality voting. This is perhaps unsurprising given that plurality voting is the simplest form of voting and what is used by far the most in political elections in the United States. However, participants tended to view RCV as fairer than plurality. In Studies 3 and 4 participants even sometimes viewed cardinal voting as fairer than plurality. These studies did not include nationally representative samples, and samples from different groups may exhibit different preferences. For example, perhaps a less liberal sample may not be as favorable to RCV as this sample. Still, these studies suggest that at least in certain groups there is considerable disagreement between people's declared preferences and how they would actually decide who won an election if put in charge of doing so.

However, when statistically analyzing the consistency between the WWET and the declared preferences, the amount of consistency was fairly low. In Studies 1 and 2, which used the graphical format, there was above-chance consistency, though the WWET only explained about 41% of participants' choices when making declared preferences (chance is 33%). In Studies 3 and 4, which used a tabular format there was not even above-chance consistency. Of course, the obvious question then is why is there such a gap when preferences are measured in two ways? We found that one contributing factor was because participants do not understand how to implement the voting systems, which relates to the next question.

6.1.2 Comprehension and Declared Preferences

The second main question we asked was if people prefer voting systems that they understand

better. When participants' comprehension of the three voting systems was tested by asking them which candidate won according to a given voting system in a task very similar to the WWET, their understanding of plurality was near ceiling, followed by cardinal voting, and lastly RCV. Something remarkable about this pattern is that it cannot be explained by lack of ability to derive the necessary information from the graphs (Studies 1 and 2) or tables (Studies 3 and 4). If, for example, participants could visually see which candidate was closest to each voter in Studies 1 and 2, but had difficulty calculating the average distances between each candidate and all the voters, then they should have done well on the comprehension tests for both plurality and RCV, but poorly for cardinal voting; RCV simply requires participants to derive the same information from the graphs as plurality voting.

The reason that participants understood RCV so poorly is that they often failed to implement the instant-runoff logic and instead just focused on each voter's first choice in the rankings, so they ended up choosing the plurality winner rather than the RCV winner. This point was made very clearly in Studies 3 and 4. In Study 3, in both the "Raw" and "Raw+Summary" conditions participants had to implement the instant run-off, and participants' comprehension was equally poor in both. In Study 4, in the "Raw+Summary" and "Summary" conditions the RCV winner after doing the instant-runoff was shown to participants, and their comprehension was near ceiling. In sum, the WWET shows that participants rarely spontaneously implement the instant run-off of RCV on their own, and the comprehension tests demonstrated that even after being taught about the instant run-off logic, participants had difficulty implementing it (unless it was essentially done for them).

But again, the analyses above involved examining comprehension and declared preferences separately; it is possible that participants who have higher comprehension of a given voting

system would rate it higher in the declared preferences. We found that participants with a declared preference for cardinal voting indeed understood cardinal voting better than participants who preferred plurality or RCV, and participants with a declared preference for plurality tended to have a poor understanding of cardinal voting in Studies 1 and 2. But aside from this relation between understanding and declared preferences for cardinal voting, there was no relation between understanding and preferences for RCV and plurality; participants who preferred RCV did not understand it better, and almost all participants understood plurality well even if they did not prefer it.

In summary, there were only occasional relations between comprehension and declared preferences; there is some evidence that participants who understood cardinal better declared it fairer, but there were no relations between comprehension and declared preferences for plurality and RCV.

6.1.2 Metacognition

In addition to the objective comprehension measures, we also assessed participants' judgments of knowledge for how well they understood each voting system. In Studies 2 and 4, participants gave higher meta-cognitive ratings for RCV than cardinal, and in Study 3, there was not a difference; participants failed to recognize that they actually understood RCV worse than cardinal and plurality. We also examined the relationship between metacognition and comprehension, which told a complimentary story. We found clear evidence that participants were not accurately self-assessing their understanding of the voting systems, particularly RCV. In Studies 2 and 4, we found some evidence suggesting that individuals with the poorest comprehension for RCV were especially overconfident and believed they understood RCV much better than they really did. These findings are related to the literature on "illusion of explanatory

depth" as well as the Dunning-Kruger effect–where individuals with the lease domain-specific knowledge are most overconfident of their understanding (Dunning, 2011).

People often have an "illusion of explanatory depth" when trying to explain complex phenomena (Rozenblit & Keil, 2002). This "illusion" refers to people thinking that they know much more about something works than they do. For example, an individual may self-assess a high level of understanding for how a sewing machine works, but may struggle to offer a concrete explanation if asked to do so. In our studies, we found that people' self-assessments for their knowledge about RCV works was not well-aligned with their actual understanding when formally evaluated. It may be that RCV brings more complexity than first meets the eye and people underestimate their understanding in response. Additionally, the process of completing a RCV ballot (i.e., ranking preferences) is a much simpler than the instant-runoff logic used by RCV. Perhaps people tend to focus more on the process of voting than the mechanism(s) for aggregating the ballots. In this way, the fluency (or ease) experienced when thinking about the ballot process influences feelings of understanding about the voting system more generally.

6.1.3 Individual Differences

A third main question was whether there were individual differences that lead participants to prefer one voting system over another. The individual difference measures had limited success in explaining our participants voting system preferences. The Cognitive Reflection Test was a notable stand out, however. One reason CRT may have been an effective predictor of preferring cardinal voting over plurality voting is that participants may be higher in deliberative thinking and more willing to expend the effort to understand the voting systems beyond the simple method of plurality voting.

The morality-focused measures were less predictive than we had hoped. One reason this may

have been the case is because participants did not understand that the three competing voting systems operationalize fairness in different ways. Indeed, if participants did, it is possible that different moral priorities would share a stronger relationship to voting system choices. For instance, the hypothesis that more utilitarian values would be associated with cardinal instead of plurality voting seems quite straightforward. But if participants do not understand that cardinal voting best fits the average voter preferences, then this connection would not be clear to them. One question that would be interesting for future research is to longitudinally measure understanding and preferences, in addition to a baseline measure of moral beliefs, and examine if, as individuals grow in their understanding of different voting systems, there is greater matching between morality and voting system preferences. In our studies, the interventions were very short (~15 minutes) and we did not train for advanced understanding. However, examining the transition from novice understanding to more advanced understanding, in particular, would provide a better understanding of the relationship between voting system preference and moral beliefs.

6.1.4 Teaching Interventions

Our fourth main question concerned the ability to successfully design a teaching intervention to improve participants' understanding of the different voting systems, which was addressed in Study 2. We found that all three interventions affected both how well participants understood the voting systems as well as their ratings about their fairness. In particular, retrieval practice was most effective at increasing participant understanding, and discovery learning showed the most promise for impacting how fair participants viewed the competing voting systems. Given that all three interventions had an impact on both comprehension and preferences for voting systems, relative to the control condition, this suggests that pointing out strengths and weaknesses of

competing voting systems may in itself be a fruitful approach to teaching people about voting. And in fact, after our short intervention participants' choices in the WWET were somewhat more consistent with their declared preferences.

6.2 Who Won the Election Task

A notable contribution from this dissertation is the creation of the Who Won the Election Task. We used the WWET in three ways, all of which were novel. First, asking participants who they think won an election it serves as a revealed preference measure; their preferences about voting systems could be assessed without them even knowing about voting systems. Second, we used the WWET to teach participants about the voting systems. Third, after teaching participants about the voting systems, we asked them to judge who won an election according to each voting system. Thus, it can serve as a measure of voting system comprehension. Each one of these three aspects provided valuable insights into the psychology of voting and speaks more generally to the role cognitive scientists can play in creating a more informed body politic and these three are addressed in turn.

First, the way that voting system preferences were captured in the WWET is a notable contribution; in a similar fashion as literature on revealed preferences, we examined how behaviors can reveal preferences that may differ from declared preferences. Existing measures of voting system preferences really only capture preferences about the ballot types, not the aggregation method. For instance, Blais, Plescia, and Semi (2021) asked participants to rate their satisfaction with single-choice, rankings, ratings, or approval voting (e.g., yes/no for each option) ballots. A problem with this approach is that the same voting system can still be used with different ballots; we found that when shown rankings people still often implement plurality instead of RCV. Other approaches that have been used involve showing people official ballots

for different voting methods and asking which method they prefer (e.g., Donovan, Tolbert, & Gracey, 2019), or providing a text explanation of how a voting system works to people and then having people rate how much they like it (e.g., Kimball & Anthony, 2021). The WWET captures voting system preferences when people are tasked with actually wrestling with how the system works relative to competing methods. Our approach of showing different granularities of data allows people to consider counterfactual outcomes that differ by which voting method is used to determine the winner, which we believe is a richer and more informative metric then the other approaches mentioned here.

Second, we believe that the WWET does a better job of assessing comprehension than existing ways to assess voters' understanding of different voting systems. For instance, a recent study presented RCV ballots to voters with instructions on how to complete the ballot and then asked them: "How well do you think you understand ranked-choice voting?" (Donovan, Tolbert, & Gracey, 2019). They found that only 13% of voters reported having a poor understanding of the system and 54% stated that they understood the ballot "very well" or "extremely well." These results echo others, in which individuals also self-report understanding RCV (see Kimball & Anthony, 2021 for references). For instance, 85% of voters surveyed in Alaska deemed RCV "simple" (Moser, 2022). The problem with these measures is that they only examine if individuals can correctly fill out an RCV ballot, not whether they understand how RCV's aggregation method works. The former is a necessity for voters' to effectively cast a RCV ballot, but the latter is necessary for them to understand the results–and possibly more importantly to buy into the system altogether. Furthermore, these measures are only self-reports and do not assess understanding against an objective standard.

An additional contribution of this project has been connecting our novel approach for

capturing voting systems preferences to formal teaching methods which can improve people's understanding of voting systems. Although the positive benefits of retrieval practice were not surprising, the fact that the re-study condition positively benefitted comprehension simply by repeating the explanations of how the voting systems worked was surprising. One can imagine that a combination of the three methods may be especially beneficial. Participants may begin with reading a description of how a voting system works. Next, they are tested on implementing the mechanism to find the winning candidate, which could be repeated a few times as in our intervention. Next, participants are walked through scenarios showing the strengths and weaknesses of the voting system and asked thoughtful questions which enhance the learners' understanding. We found this approach can be effectively employed in just 10-15 minutes.

6.3 Limitations

One notable limitation of this line of work is that undergraduate students are not representative of the American voting electorate. Indeed, supporters of RCV have been found to be younger, more educated, and Democratic-leaning (Anthony & Kimball, 2021), which also demographically overlaps with our convenience sample of students taking Intro to Psychology at the University of Pittsburgh. Although we do not necessarily think that the general public has a better understanding of voting systems than our predominately 18-year-old convenience sample, it is likely that the general public has different preferences about voting systems. In order to understand if these findings generalize to the US population it would be important to obtain a nationally-representative sample.

Another limitation to the external validity of our paper is that in the WWET we used a scenario about hypothetical hiring decisions, so we were studying 'cool cognition' rather than 'hot' (Redlawsk, 2002). In contrast, in real-world voting situations, many voters would have

strong preferences for which candidate they want to win, and they may change their preferences about the voting systems depending on which voting system they think is most likely to lead to their preferred candidate winning. In this initial investigation of voting system preferences, we decided to focus on the simpler 'cold cognition' and not also add in politically-motivated decision making, but an important direction for future research is to study how political bias can change perceptions about fairness of outcomes and voting system preferences.

Similar to the point above, we purposely did not examine strategic voting. Strategic voting occurs when voters submit ballots for candidates that deviate from their sincere preferences because they believe doing so is more likely to lead to desirable outcome. This helped simplify the problem space of the current work, but because in real-world elections, individuals can benefit from misrepresenting their sincere preferences, strategic voting is an important aspect to incorporate in future research. Furthermore, even though we tried to emphasize sincere preferences in the voting data presented in the WWET and other survey areas, it is clear from open-ended responses from participants that strategic voting was still on some of their minds. It was not uncommon for participants to say that they believed a disadvantage of cardinal voting was that voters would not use the middle of the 0-100 scale, but rather would only give very low or very high scores. Despite our precautions and attempts to eliminate any influence from strategic voting, at least some participants were still thinking about how this system would play out in real-world cases.

We made a number of decisions to limit the situations that we tested in the WWET in order to narrow the scope of this project. First, we limited the number of candidates to a maximum of 3. Second, we limited the voting systems under investigation to only plurality, RCV, and cardinal voting. However, there are many more systems that could have been studied. Originally, we also

wanted to include approval voting (which is another type of cardinal voting in which options are given either a 1 or a 0), but for a number of reasons did not fit well into the visual design of the WWET. However, other voting systems like Borda (and likely many others) could fit nicely with the current visual design, but were not investigated. Third, we only investigated cases with a fairly small number of voters, up to 18. However, political elections typically involve many thousands of voters or potentially many millions. This dramatically alters how the raw voting data could be displayed to participants. For example, in the graphical format in Studies 1 and 2 it could be displayed like a histogram. For the tables used in Studies 3 and 4, we do not see a way to display the raw data, so presumably the 'summary only' display would be used.

Whenever participants learned about quirks in the voting systems, there was never information presented about how often each quirk occurs. In free-response reports, participants found the violation of the majority criterion, a weakness of cardinal voting, to be especially aversive; however, violations of the majority criterion may be far less likely to occur compared to a spoiler effect in plurality voting. Participant ratings of the voting systems may have looked different if participants had this information. More broadly, only a limited number of quirks were presented to participants. In sum, future research could investigate other quirks, how problematic people view each quirks, and how best to explain quirks to voters.

This dissertation presents a small subset of the possible analyses. The data collected across the four studies is ripe for further analysis. One direction is to use longitudinal SEM which could offer valuable insights, particularly within the studies where participants completed the WWET several times with comprehensions tests and preference measurements in between. These analyses could shed light on how understanding and preferences interact in a more nuanced manner.

Although we did not consider additional voting systems, the choices in the WWET could also be scored for their consistency to other voting systems (e.g., Borda voting). Participant data from our four studies can be scored for congruence with additional voting systems, which may reveal interesting insights into how choices in the WWET map onto other systems.

6.4 Conclusion

The goal of this paper was not to seek a definitive answer to the question of what voting system people prefer, but to develop methods to understand how to go about studying this question. We found that people had different voting system preferences, depending on how the preference was measured. We also found that although many people preferred alternative voting systems, they struggled to understand them and do not always accurately self-assess their knowledge. This raises important questions about the accuracy of reported voting system preferences.

We found robust evidence that individuals especially struggle to understand how RCV really works. This difficulty appears to be specifically related to understanding how to implement the formal logic of the system, where one must first check if a candidate has over 50% of the vote, prior to eliminating the weakest candidate and reassigning their votes to the next candidate, before tallying the votes again. Proponents of RCV should consider educational campaigns to address how votes get tallied and not just how to correctly fill out a ballot. We found that our approach to teaching voting systems was effective and may be improved upon with refinement of the methods used. Based on other reports, people appear to be able to fill out a RCV ballot appropriately (i.e., they understand how to rank their preferences), but there is not a similarly high level of understanding for how RCV really works. It was beyond the scope of the current work to address the question “how important is it for voters to comprehend the voting system

they use?"–but in the wake of this dissertation that question seems important, especially with the growing interest in using voter systems beyond plurality voting.

We found a strong relationship between what voting systems participants thought were fair and what voting systems they want to use. People seeking to reform how we vote would benefit from explaining to people the strengths and weaknesses of different voting systems and how these specifically relate to fairness. Voting systems are of unique interest for many reasons, but maybe most of all because they attempt to put formal mathematical operations for how fairness can be achieved. Humanity would undoubtedly benefit from spending more time reflecting on how fairness is operationalized, and which methods lead to the most favorable outcomes.

We believe this research serves as an important contribution to a timely issue. Given the prominent conversations on fairness in voting as well as concerns about trust in our government institutions, more generally, we sought to add valuable insights into how we can address fairness at a structural level that is consistent with individuals' values and preferences while also strengthening our democracy in the process. The benefits of collective decision-making can be experienced when people come together with an agreed upon system for aggregating preferences. If individuals understand the process for reaching an outcome and think that process is fair, then they should accept the outcome as fair—whether it is their preferred outcome or not. For important collective decisions like electing the president of the United States of America, these conditions may even be necessary for a peaceful transition of power.

Importantly, having an agreed upon democratic process that people believe is fair and in which they accept the outcome are necessary but insufficient conditions for a functioning democracy. As Nobel laureate and social choice theorist Amartya Sen points out, democracies are governance by discussion (Sen, 2018). For citizens to have buy-in for their democracy, they

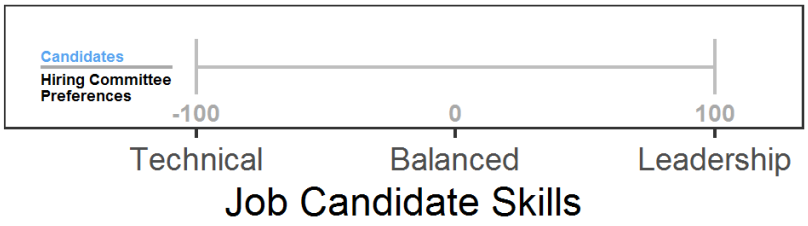
also must believe that their votes have an intended impact on later decision-making. If voters elect specific candidates or political parties with the goal of enacting specific policies—and these policies do not come to pass—voters may begin to believe that their democracy is not serving them and in turn has diminishing value.

Appendix A. Study 1 Survey Materials

A1. How to read graph introduction

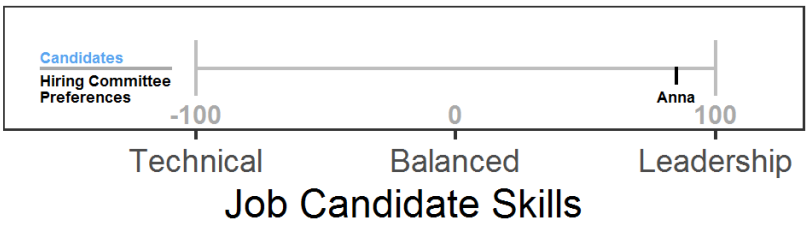
This study is about how groups of people make collective decisions. We will be using an example of a company hiring a job candidate. There is a hiring committee of people who work for the company who have different preferences of the types of candidates they want to hire. The question is how to decide which candidate to hire based on the preferences of the committee members.

We will be showing you figures like the one below. This figure shows the skillset of job candidates in terms of having very good technical skills (-100), very good leadership skills (100), or being balanced (0).



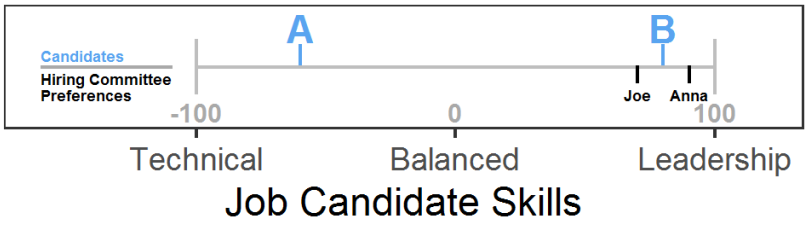
Black lines represent individual preferences of members on the hiring committee.

For example, Anna prefers to hire someone who is high in leadership skills.



Blue lines represent job candidates. Here there are two job candidates. Candidate "A" has more technical skills. Candidate "B" has more leadership skills.

The figure below shows that the two members of the **hiring committee, Joe and Anna**, both want a candidate high on leadership. In this case, both of their preferences match well with the **skills of candidate "B"**. In this example, candidate "B" should be hired.

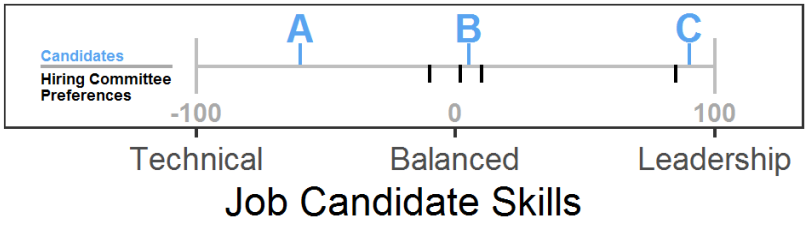


The figure below shows four black lines. These lines represent the preferences of the four members of the hiring committee. From now on we will hide the names of members on the

hiring committee just to make the diagram less cluttered.

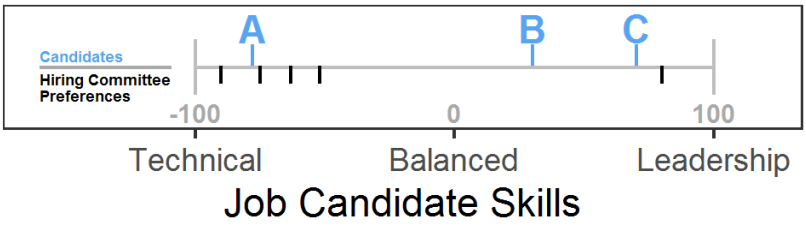
Three members of the hiring committee want a candidate with a balanced skillset, and one member wants to hire a candidate with strong leadership skills.

In this example, **candidate "B"** is a good match for 3 of the 4 members of the hiring committee. Only one member of the hiring committee is aligned with candidate "C", who has more leadership skills. Candidate "A", who has a more technical skillset, is not close to any members of the hiring committee. In this example, Candidate "B" would be hired.



A2. Graph comprehension

Comprehension item 1:
The figure below shows the skillsets of three job candidates ("A", "B", and "C"). In addition, this figure shows five black lines, which represent the preferences of the five members of the hiring committee.

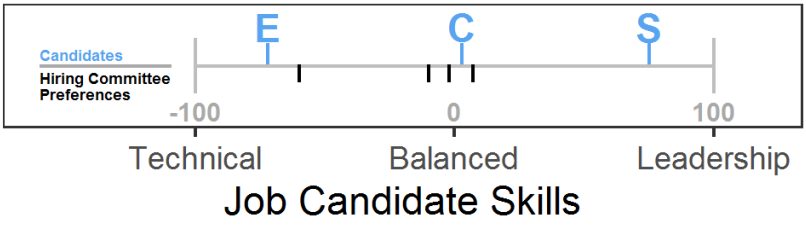


Based on the preferences of the hiring committee members in the graph above, **which job candidate do you think should be hired**?

- Candidate A (*correct answer*)
- Candidate B
- Candidate C

If participant correctly chose "Candidate A", they would see the message:
"Correct! Candidate A is the best match based on the hiring committee's preferences."
If participant chose Candidate B or C, they would see the message:
"Incorrect. Candidate A is the best match based on the hiring committee's preferences. Notice that four members of the hiring committee have a preference for a technical candidate and that Candidate A most closely fits their preferences."

Comprehension item 2:
The figure below shows the skillsets of three job candidates ("E", "C", and "S"). In addition, this figure shows four black lines, which represent the preferences of the four members of the hiring committee.



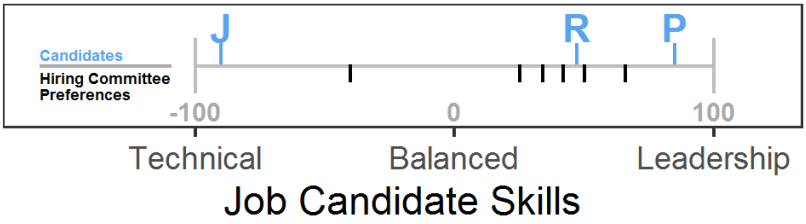
Based on the preferences of the hiring committee members in the graph above, **which job candidate do you think should be hired**?

- Candidate E
- Candidate C (*correct answer*)

- Candidate S

If participant correctly chose “Candidate C”, they would see the message:
“Correct! Candidate C is the best match based on the hiring committee's preferences.”
If participant chose Candidate E or S, they would see the message:
“Incorrect. Candidate C is the best match based on the hiring committee's preferences. Notice that three members of the hiring committee have a preference for a balanced candidate and that Candidate C most closely fits their preferences.”

Comprehension item 3:
The figure below shows the skillsets of three job candidates ("J", "R", and "P"). In addition, this figure shows six black lines, which represent the preferences of the six members of the hiring committee.



Based on the preferences of the hiring committee members in the graph above, **which job candidate do you think should be hired**?

- Candidate J
- Candidate R (*correct answer*)
- Candidate P

If participant correctly chose “Candidate R”, they would see the message:
“Correct! Candidate R is the best match based on the hiring committee's preferences.”
If participant chose Candidate J or P, they would see the message:
“Incorrect. Candidate R is the best match based on the hiring committee's preferences. Notice that five members of the hiring committee have a preference for a leadership leaning candidate and that Candidate R most closely fits their preferences.”

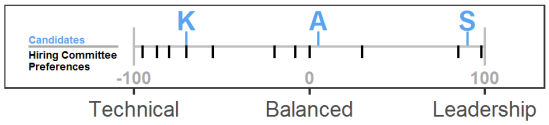
A3. Who Won the Election Task

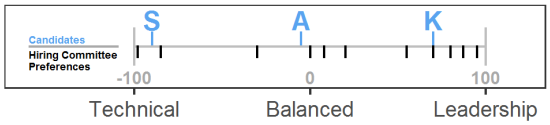
[Note: All item choices had randomized presentation order. Participants were randomized to 1 of 2 stimuli sets. The left-side images comprise set 1, and the right-side images comprise set 2. The voting system consistent with choosing a particular candidate is noted in the items below in italics.]

Introduction text:

> In the following pages you will be shown figures like the ones you just saw and asked which job candidate you think should be hired.

Item 1:

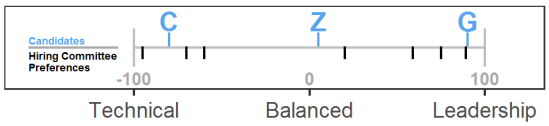


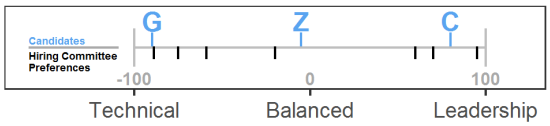


Based on the preferences of the hiring committee members in the graph above, **which job candidate do you think should be hired**?

- Candidate K (*Plurality*)
- Candidate A (*Ranked-Choice / Cardinal*)
- Candidate S (*None*)

Item 2:

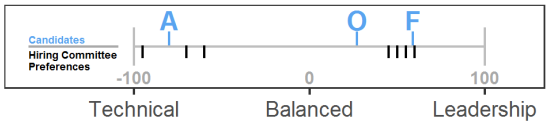


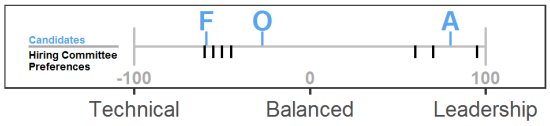


Based on the preferences of the hiring committee members in the graph above, **which job candidate do you think should be hired**?

- Candidate C (*Plurality*)
- Candidate Z (*Cardinal*)
- Candidate G (*Plurality / Ranked-Choice*)

Item 3:



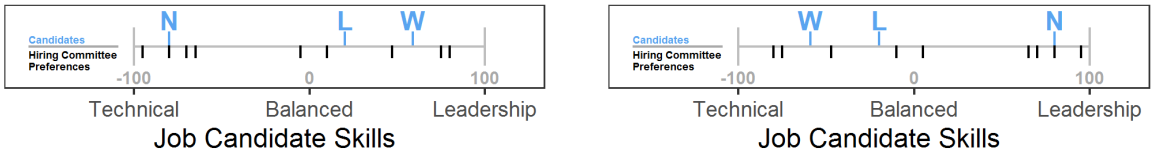


Based on the preferences of the hiring committee members in the graph above, **which job candidate do you think should be hired**?

- Candidate A (*None*)

- Candidate O (*Cardinal*)
- Candidate F (*Plurality / Ranked-Choice*)

Item 4:



Based on the preferences of the hiring committee members in the graph above, **which job candidate do you think should be hired**?

- Candidate N (*Plurality*)
- Candidate L (*Cardinal*)
- Candidate W (*Ranked-Choice*)

Item 5:

Based on the preferences of the hiring committee members in the graph above, **which job candidate do you think should be hired**?

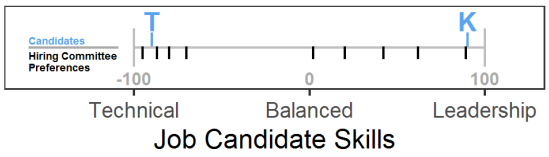
- Candidate I (*Plurality*)
- Candidate Q (*Cardinal / Ranked-Choice*)
- Candidate Z (*None*)

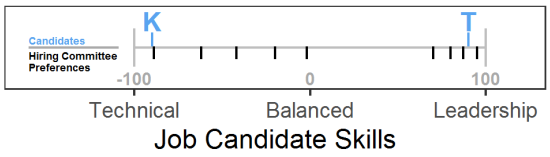
Item 6:

Based on the preferences of the hiring committee members in the graph above, **which job candidate do you think should be hired**?

- Candidate U (*Ranked-Choice*)
- Candidate S (*Cardinal*)
- Candidate J (*Plurality*)

Item 7:

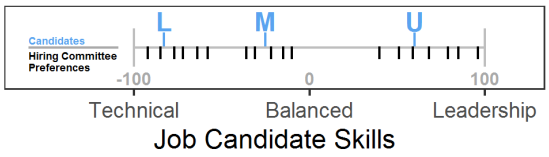


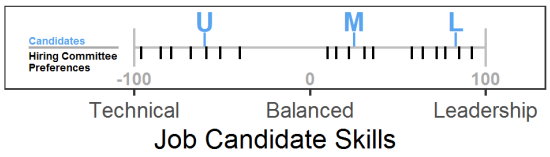


Based on the preferences of the hiring committee members in the graph above, **which job candidate do you think should be hired**?

- Candidate T (*Cardinal*)
- Candidate K (*Plurality / Ranked-Choice*)

Item 8:

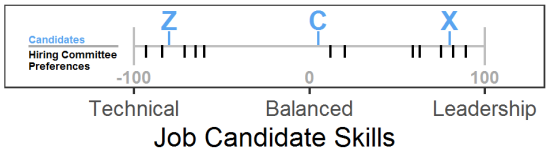


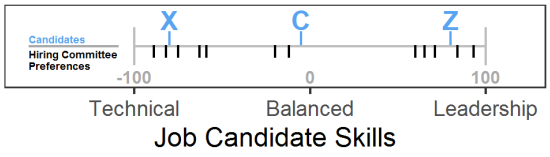


Based on the preferences of the hiring committee members in the graph above, **which job candidate do you think should be hired**?

- Candidate L (*Ranked-Choice*)
- Candidate M (*Cardinal*)
- Candidate U (*Plurality*)

Item 9:

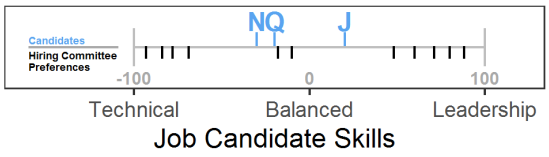


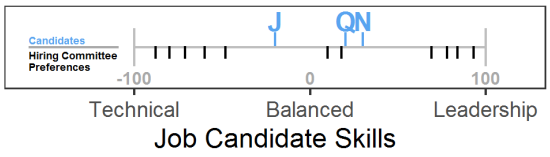


Based on the preferences of the hiring committee members in the graph above, **which job candidate do you think should be hired**?

- Candidate Z (*Plurality*)
- Candidate C (*Cardinal*)
- Candidate X (*Plurality / Ranked-Choice*)

Item 10:





Based on the preferences of the hiring committee members in the graph above, **which job candidate do you think should be hired**?

- Candidate N (*Ranked-Choice*)
- Candidate Q (*Cardinal*)

- Candidate J (*Plurality*)

Table Stimuli Item:
[Note: Participant's were randomized to 1 of the 6 different counter-balanced formats below.]

Instructions:
Now we are going to ask you a similar question, however, this time we are going to show you committee members' preferences in a different format.

Nine hiring committee members were asked to provide their hiring preference for three job candidates (candidates: "J", "M", and "E"). The table below shows the nine hiring committee members' 1st, 2nd, and 3rd preference.

Version 1

	Preference		
Member	*1st*	*2nd*	*3rd*
1	J	M	E
2	J	M	E
3	J	M	E
4	J	M	E
5	M	E	J
6	M	E	J
7	E	M	J
8	E	M	J
9	E	M	J

Version 2

	Preference		
Member	*1st*	*2nd*	*3rd*
1	E	M	J
2	E	M	J
3	E	M	J
4	J	M	E
5	J	M	E
6	J	M	E
7	J	M	E
8	M	E	J
9	M	E	J

Version 3

	Preference		
Member	*1st*	*2nd*	*3rd*
1	M	E	J
2	M	E	J
3	J	M	E
4	J	M	E
5	J	M	E
6	J	M	E
7	E	M	J
8	E	M	J
9	E	M	J

Version 4

	Preference		
Member	*1st*	*2nd*	*3rd*
1	J	M	E
2	J	M	E
3	J	M	E
4	J	M	E
5	E	M	J
6	E	M	J
7	E	M	J
8	M	E	J
9	M	E	J

Version 5

	Preference		
Member	*1st*	*2nd*	*3rd*
1	E	M	J
2	E	M	J
3	E	M	J
4	M	E	J
5	M	E	J

Version 6

	Preference		
Member	*1st*	*2nd*	*3rd*
1	M	E	J
2	M	E	J
3	E	M	J
4	E	M	J
5	E	M	J

6	J	M	E
7	J	M	E
8	J	M	E
9	J	M	E

6	J	M	E
7	J	M	E
8	J	M	E
9	J	M	E

Based on the preferences of the hiring committee members in the table above, which job candidate do you think should be hired?

- Candidate **J** (*Plurality*)
- Candidate **M** (*Cardinal*)
- Candidate **E** (*Ranked-Choice*)

A4. Voting system introduction

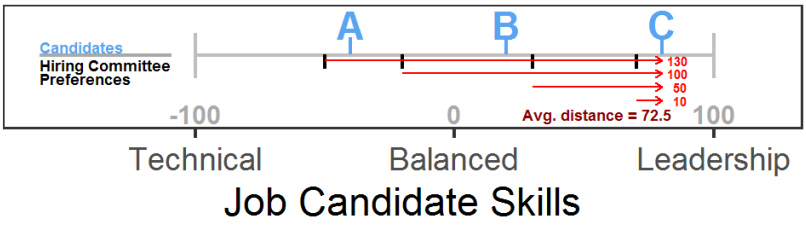
Note: Participants were formally introduced to cardinal, plurality, and ranked-choice voting in this block. The voting system order was randomized for each participant.

Cardinal Voting

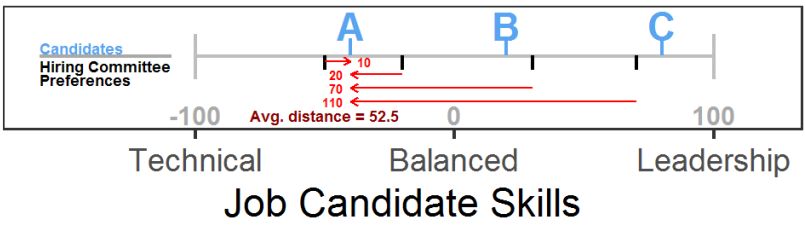
One method to determine election outcomes is to focus on the distances between each candidate and the voters. This is called **Cardinal voting**. In Cardinal voting the candidate that is closest on average to the voters wins. In the figures below we show the distances between the four voters and the three candidates.

As you can see, C is the farthest away from the voters on average - the arrows are the longest and the average distance is 72.5. A and B are both closer to the voters on average, though the average distances for B (42.5) is somewhat shorter than for A (52.5). Thus, according to cardinal voting, **Candidate B would win**.

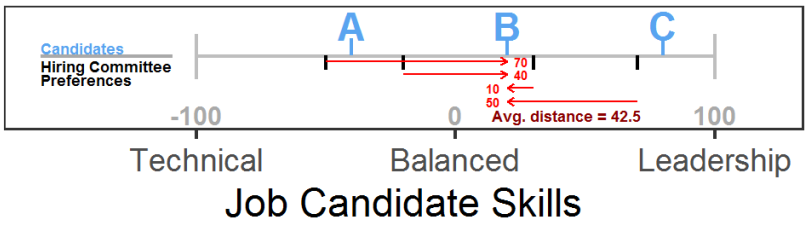
Distances to Candidate C



Distances to Candidate A



Distances to Candidate B



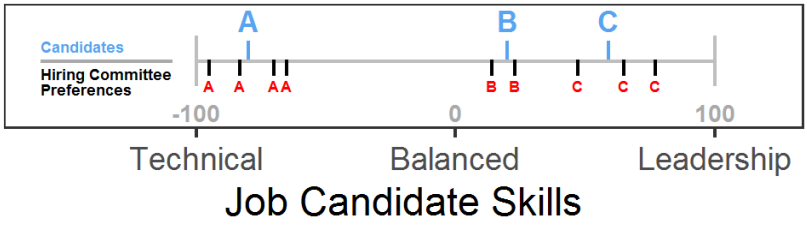
How fair is Cardinal Voting? [Likert-scale 1-5 response]

1. completely unfair	2.	3.	4.	5. completely fair

Plurality Voting

One method to determine election outcomes is to allow all voters to choose only one candidate and the candidate with the most votes wins. This is called **Plurality voting**. Plurality voting prioritizes selecting a winner who has the single most support. A winning candidate in a plurality system needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.

For example, in the graph below "Candidate A," has 4 out of 9 hiring committee members support (as noted by the red letters below the hiring preferences). "Candidate B" has 2 out of the 9 hiring committee members support. "Candidate C" has 3 out of the 9 hiring committee members support. Using a Plurality voting system, Candidate A would win, even though they did not receive more than 50% of the votes.



How fair is Plurality Voting? [Likert-scale 1-5 response]

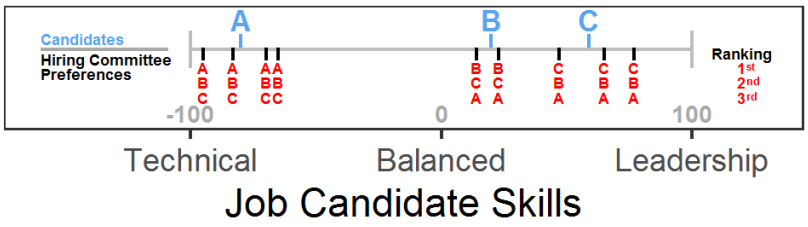
1. completely unfair	2.	3.	4.	5. completely fair

Ranked-Choice Voting

One method to determine election outcomes is to focus on how each voter ranks the candidates (e.g., first, second, third preference). This is called **Ranked-Choice voting**. Ranked-Choice voting prioritizes majority approval and a winning candidate must have over 50% of the total votes. This works by first counting all voters' first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support. If so, this candidate wins. If not, the process is repeated.

For example, in the graph below, there are four voters who are very close to Candidate A, so A gets first place for these four. There are two voters who are closest to B, and there are three voters who are closest to C. Although A has the most first place votes (4 out of 9), A does not have over 50% of the votes. Because B has the fewest first place votes, B would be removed as a candidate, and B's votes would be reassigned. In this case, the two voters closest to B would be reassigned to C, because they are closer to C (their 2nd choice) than to A (their 3rd choice). Once B is

removed, there are 5 votes for C, and 4 votes for A, so Candidate C would win.



How fair is Ranked-Choice Voting? [Likert-scale 1-5 response]

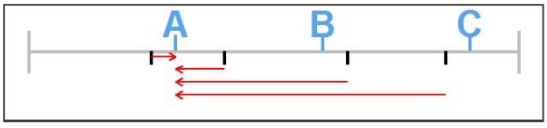
1. completely unfair	2.	3.	4.	5. completely fair

A5. Voting system comprehension

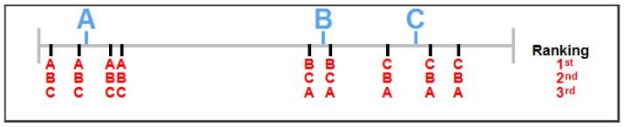
Instructions:
The next few pages will ask you to identify which job candidate should be hired based on the different ways to determine elections that you just learned about.

We will present the logos below to help you remember how the different methods work.

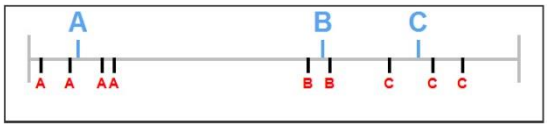
- **Cardinal voting** focuses on the distances between each candidate and the voters. The candidate that is closest on average to the voters wins.



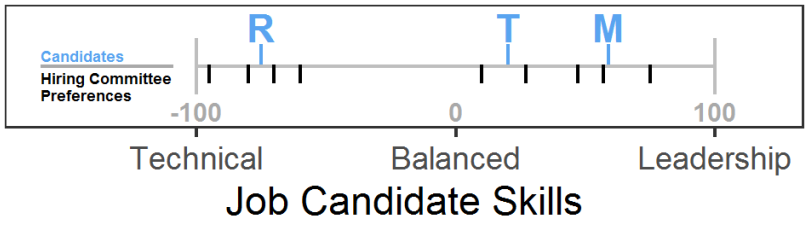
- **Ranked-Choice voting** first counts all voters' first preference and sees if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support.



- **Plurality voting** has each voter pick one candidate and the winning candidate is the one with the most votes, even if they do not have over 50% of the total votes.



Learning item 1a:
This question is about **Cardinal Voting**

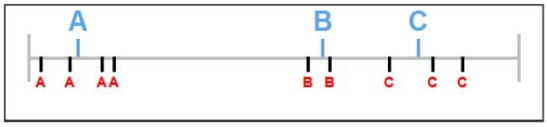


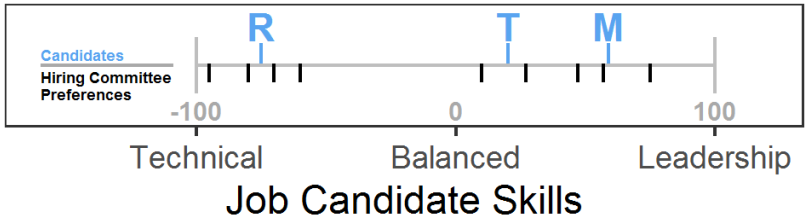
In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Cardinal voting system**?

[*Remember: In Cardinal voting a winning candidate is the candidate who is closest on average to all of the voters.*]

- Candidate **R**
- Candidate **T** (*correct answer*)
- Candidate **M**

Learning item 1b:
This question is about **Plurality Voting**



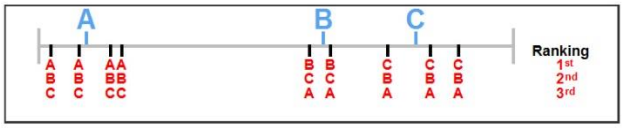


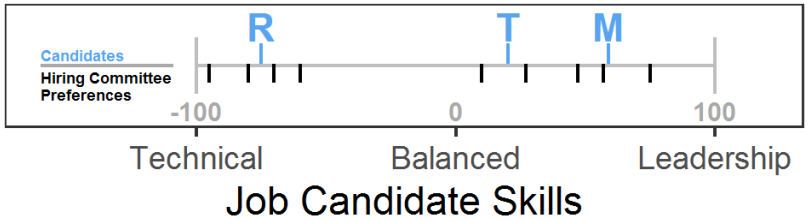
In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Plurality voting system**?

[*Remember: In plurality voting a winning candidate needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.*]

- Candidate **R** (*correct answer*)
- Candidate **T**
- Candidate **M**

Learning item 1c:
This question is about **Ranked-Choice Voting**



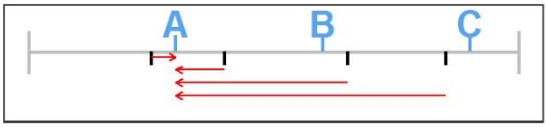


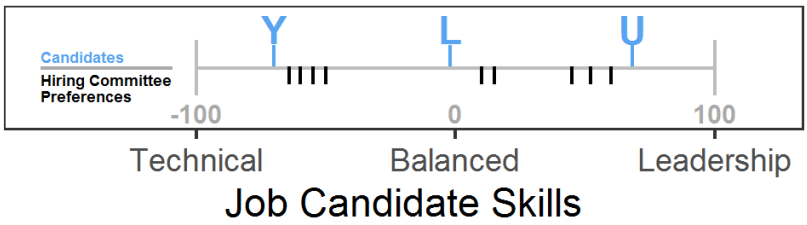
In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Ranked-Choice voting system**?

[*Remember: In Ranked-Choice voting a winning candidate must have over 50% of the total votes. This works by first counting all voters first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support.*]

- Candidate **R**
- Candidate **T**
- Candidate **M** (*correct answer*)

Learning item 2a:
This question is about **Cardinal Voting**



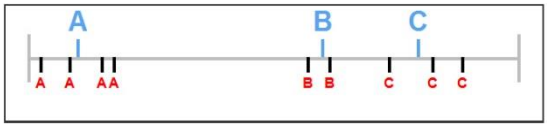


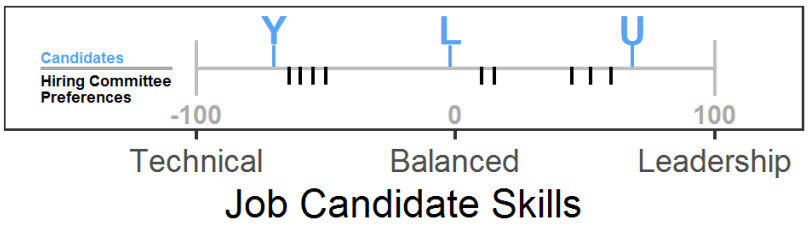
In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Cardinal voting system**?

[*Remember: In Cardinal voting a winning candidate is the candidate who is closest on average to all of the voters.*]

- Candidate **Y**
- Candidate **L** (*correct answer*)
- Candidate **U**

Learning item 2b:
This question is about **Plurality Voting**



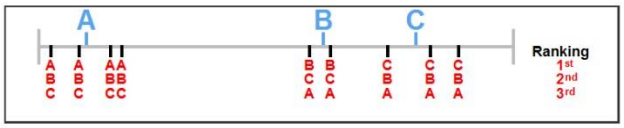


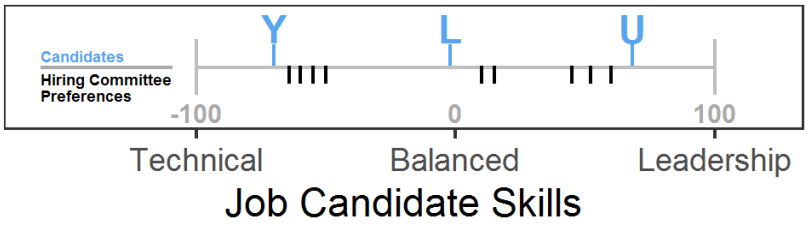
In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Plurality voting system**?

[*Remember: In plurality voting a winning candidate needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.*]

- Candidate **Y** (*correct answer*)
- Candidate **L**
- Candidate **U**

Learning item 2c:
This question is about **Ranked-Choice Voting**





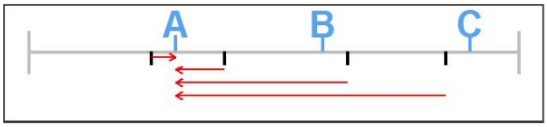
In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Ranked-Choice voting system**?

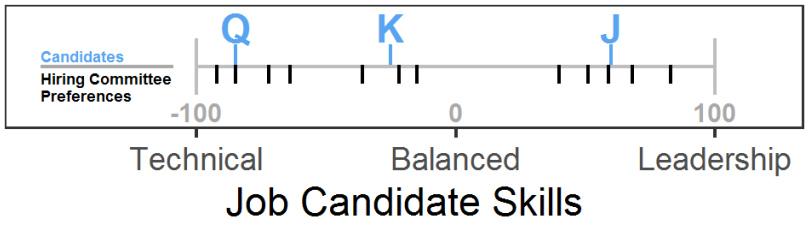
[*Remember: In Ranked-Choice voting a winning candidate must have over 50% of the total votes. This works by first counting all voters first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support.*]

- Candidate **Y**
- Candidate **L**
- Candidate **U** (*correct answer*)

Learning item 3a:

This question is about **Cardinal Voting**



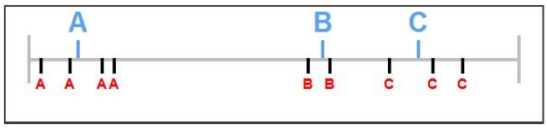


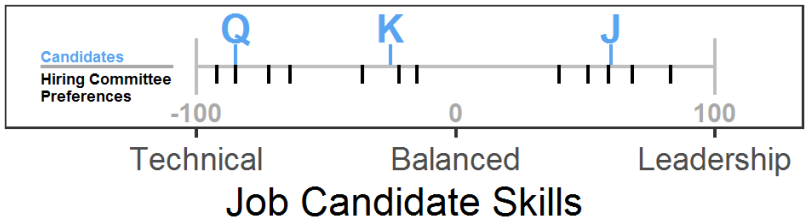
In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Cardinal voting system**?

[*Remember: In Cardinal voting a winning candidate is the candidate who is closest on average to all of the voters.*]

- Candidate **Q**
- Candidate **K** (*correct answer*)
- Candidate **J**

Learning item 3b:
This question is about **Plurality Voting**



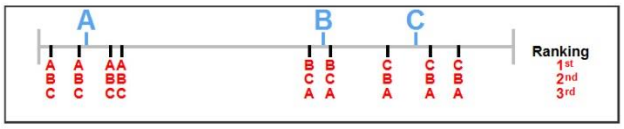


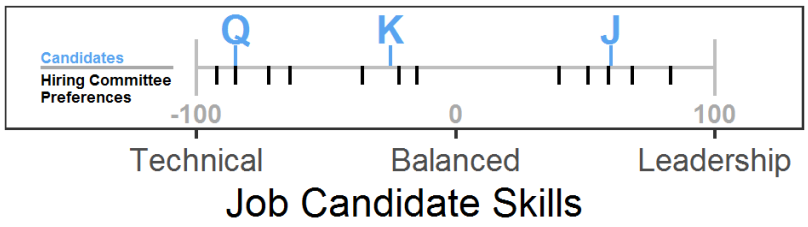
In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Plurality voting system**?

[*Remember: In plurality voting a winning candidate needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.*]

- Candidate **Q**
- Candidate **K**
- Candidate **J** (*correct answer*)

Learning item 3c:
This question is about **Ranked-Choice Voting**





In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Ranked-Choice voting system**?

[*Remember: In Ranked-Choice voting a winning candidate must have over 50% of the total votes. This works by first counting all voters first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support.*]

- Candidate **Q** (*correct answer*)
- Candidate **K**
- Candidate **J**

A6. Declared voting system preferences

1. Which of the three voting systems below do you think is best?
   (see figures and descriptions below if you need a reminder)
   A. Ranked-Choice
   B. Plurality
   C. Cardinal

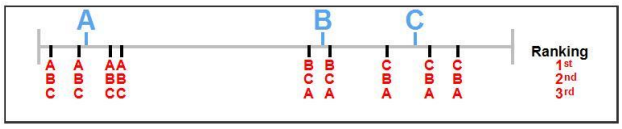
[the figures and explanations below were included as a reminder to participants]

Which of the three voting systems below do you think is best?

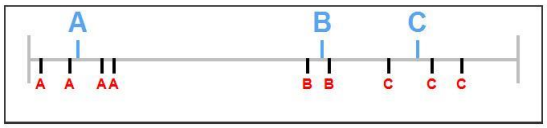
Remember:

- *In **Cardinal** voting a winning candidate is the candidate who is closest on average to the all of the voters. The figure below provides an example of how cardinal voting works for calculating the score of a Candidate by measuring the distance from each member of the hiring committee.*

- *In **Ranked-Choice** voting a winning candidate must have over 50% of the total votes. This works by first counting all voters first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support. The figure below shows how ranked-choice voting works. The red letters represent preference rankings for each member of the hiring committee. The top row is the first pick, the middle row is the second favorite, and the bottom row is the least* favorite.



- *In **plurality** voting a winning candidate needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win. The figure below shows how plurality voting works.*



2. How fair do you think each of the following voting systems are? [Likert-scale 1-5 ratings: 1=very unfair, 5=very fair]
   2a. Ranked-Choice
   2b. Plurality
   2c. Cardinal

A7. Self-reported political knowledge, behavior, and beliefs

Note: Descriptive statistics are included from Study 1 along with Question text.

Question Text	
n (%)	Answer Text
Q1. "*Right now most elections in the United States, for local, state, and federal office, use the Plurality voting system. Do you feel like laws should be changed to switch to either a Ranked-Choice or Cardinal voting system instead?*"	
96 (34%)	Keep using the Plurality voting system
48 (17%)	No preference
140 (49%)	Change the laws to switch to using a Cardinal or Ranked Choice voting system
Q2. "*Is it possible in the United States for Candidate A to get more total votes than Candidate B, but for Candidate B to win the electoral college, so that Candidate B becomes president instead of Candidate A?*"	
240 (85%)	Yes, this is possible and has happened
10 (4%)	Yes, it is possible but it has not happened in modern history
5 (2%)	No, this is not possible
29 (10%)	I'm not sure
Q3. "*Thinking for a moment about the way in which the president is elected in this country, which would you prefer?*"	
54 (19%)	Keep the current system, in which the candidate who wins the most votes in the Electoral College wins the election
33 (12%)	No preference
197 (69%)	Amend the Constitution or change the current system so the candidate who receives the most total votes nationwide wins the election.
Q4. "*Were you eligible to vote in the 2020 United States Presidential election?*"	
256 (90%)	Yes
28 (10%)	No
Q5. "*Did you vote in the 2020 United States Presidential election?*"	
234 (82%)	Yes
50 (18%)	No
Q6. "*How important is voting?*" [Likert-scale 1-5 response]	
2 (1%)	1-not very important
1 (0%)	2-
13 (5%)	3-
41 (14%)	4-
227 (80%)	5-very important

Note: *n* refers to participant count responses. *N* = 284. Question #3 is from the 2020 Pew Research Center's American Trends Panel [https://www.pewresearch.org/wp-content/uploads/2020/03/Electoral-College-topline.pdf]

A8. Demographic Questions

1. What is your age? [fill in the blank]

2. What is your gender?
   - Male
   - Female
   - Non-binary
   - Other / prefer not to say

3. What ethnicity do you consider yourself to be?
   - American Indian / Alaska Native
   - Asian
   - Black / African American
   - Hispanic / Latino
   - Native Hawaiian / Other Pacific Islander
   - White / Caucasian
   - Other

4. Do you live in the United States of America?
   - Yes
   - No

5. What is your highest level of education you have attained?

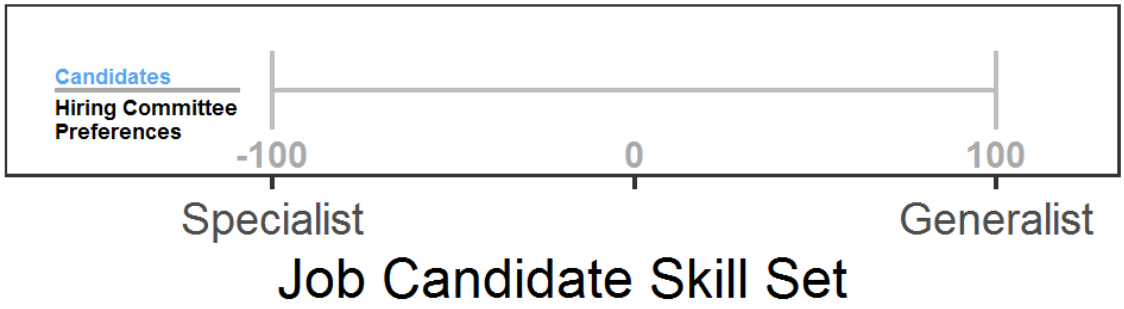
   - Not applicable
   - We did not attend high school
   - We did not complete high school
   - We completed high or got a G.E.D.
   - We completed some college classes but We did not receive a degree
   - We received an associate's degree
   - We received a bachelor's degree
   - We attended graduate school but did not receive a degree
   - We received a graduate degree (master's MBA, Ph.D., etc)

Appendix B. Study 2 Survey Materials

B1. How to read graph introduction

This study is about how groups of people make collective decisions. We will be using an example of a company hiring a job candidate.

There is a hiring committee of people who work for the company who have different preferences of the types of candidates they want to hire. The question is how to decide which candidate to hire based on the preferences of the committee members. We will be showing you figures like the one below. This figure shows the skillset of job candidates in terms of having either a specialist skillset (-100), a generalist skillset (100), or a balanced skillset somewhere between these two ends (0).

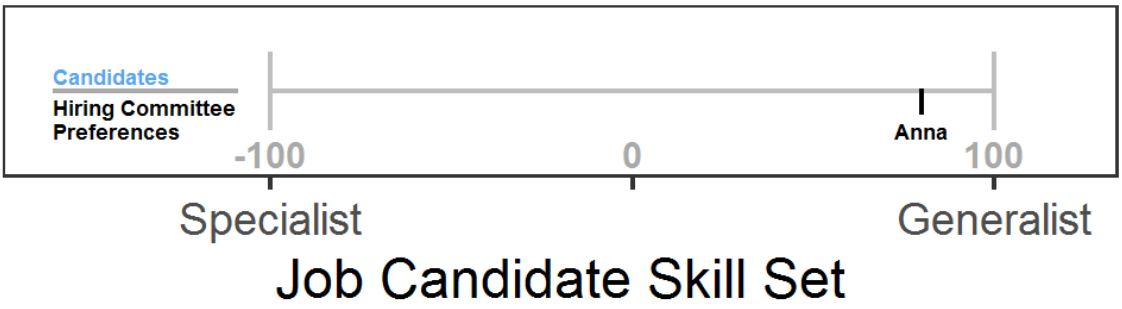


A specialist is someone who has a very targeted skillset (i.e., being extremely skilled at a narrowly defined ability, but not being able to do as many things as a generalist). A generalist is someone who has a very broad skillset (i.e., being able to do many things, although not as well as a specialist at any one task).

Both specialists and generalists can be equally valuable to companies. Whether it is better to hire more of a specialist or a generalist, depends on the preferences of the hiring committee members.

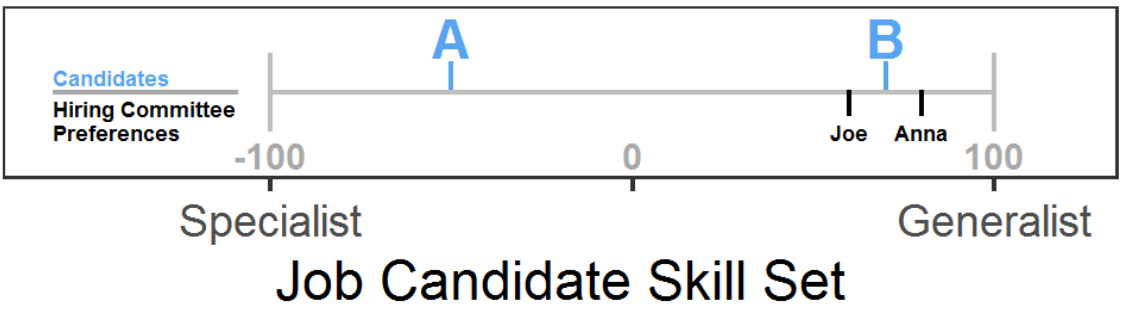
Black lines represent individual preferences of members on the hiring committee.

For example, Anna prefers to hire someone who has a generalist skillset.



Blue lines represent job candidates. Here there are two job candidates. Candidate "A" has more of a specialist skillset. Candidate "B" has more of a generalist skillset.

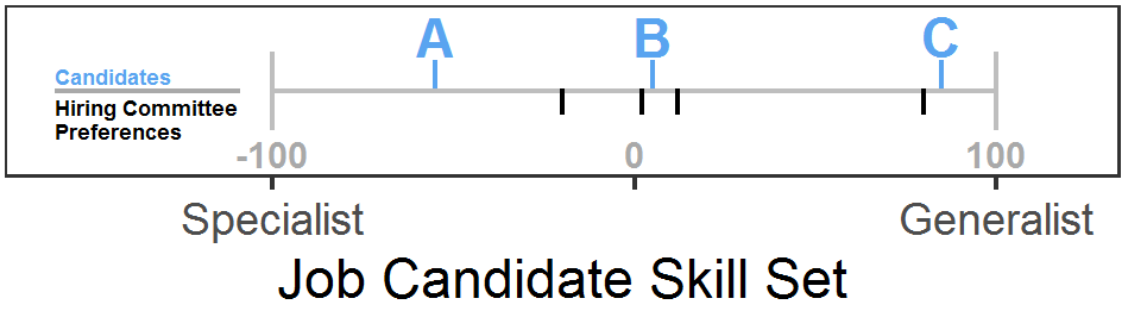
The figure below shows that the two members of the **hiring committee**, **Joe and Anna**, both want a candidate with a high level of generalist skillset. In this case, both of their preferences match well with the **skills of candidate "B"**. In this example, candidate "B" should be hired.



The figure below shows four black lines. These lines represent the preferences of the four members of the hiring committee. From now on we will hide the names of members on the hiring committee just to make the diagram less cluttered.

Three members of the hiring committee want a candidate with a balanced skillset (somewhere between a specialist and a generalist), and one member wants to hire a candidate with a strong generalist skillset.

In this example, **candidate "B"** is a good match for 3 of the 4 members of the hiring committee. Only one member of the hiring committee is aligned with candidate "C", who has a more generalist skillset. Candidate "A", who has a more specialist skillset, is not close to any members of the hiring committee.



Job Candidate Skill Set

B2. Graph comprehension

Q1. For the purposes of this study, is a Generalist or a Specialist skill set better for a job candidate?

- Specialist
- Generalist
- It depends on the preferences of the hiring committee members. (*Correct Answer*)

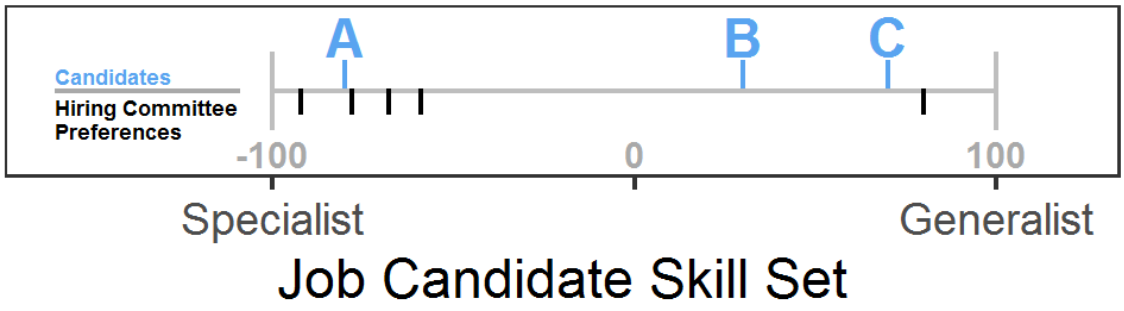
If incorrect answer is chosen, following message was displayed:

Incorrect! For the purposes of this study, both specialists and generalists can be equally valuable to companies. **Whether it is better to hire more of a specialist or a generalist, depends on the preferences of the hiring committee members**.

If correct answer is chosen, following message was displayed:

Correct! For the purposes of this study, both specialists and generalists can be equally valuable to companies. **Whether it is better to hire more of a specialist or a generalist, depends on the preferences of the hiring committee members**.

Q2. The figure below shows the skillsets of three job candidates ("A", "B", and "C"). In addition, this figure shows five black lines, which represent the preferences of the five members of the hiring committee.



Q2a. Based on the graph above, **which job candidate do you think should be hired**?
[website allowed single choice)

- Candidate **A** (*Correct Answer*)
- Candidate **B**
- Candidate **C**

Q2b. **Rank order the candidates based on who you think should be hired**.
(Top = first to hire; bottom = last to hire)
[website allowed click-and-drag functionality for rankings]

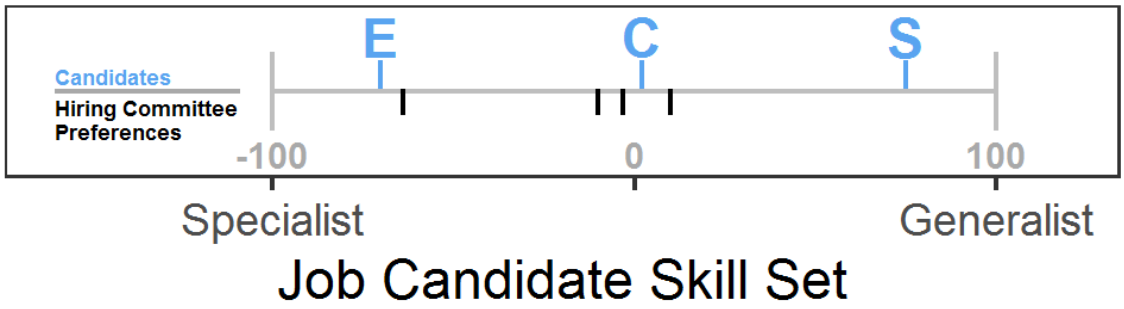
- Candidate **A** (*Correct Answer*)

- Candidate **B**
- Candidate **C**

Q2c. **Rate the candidates based on who should be hired**.
(100=most fair to hire; 0=least fair to hire)
[website had a range slider]

- Candidate **A** (*Correct Answer*)
- Candidate **B**
- Candidate **C**

Q3. The figure below shows the skillsets of three job candidates ("E", "C", and "S"). In addition, this figure shows four black lines, which represent the preferences of the four members of the hiring committee.



Q3a. Based on the graph above, **which job candidate do you think should be hired**?
[website allowed single choice)

- Candidate **E**
- Candidate **C** (*Correct Answer*)
- Candidate **S**

Q3b. **Rank order the candidates based on who you think should be hired**.
(Top = first to hire; bottom = last to hire)
[website allowed click-and-drag functionality for rankings]

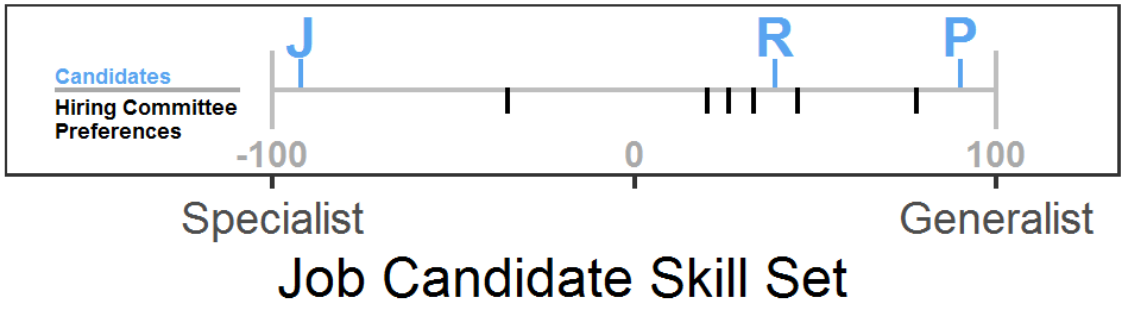
- Candidate **E**
- Candidate **C** (*Correct Answer*)
- Candidate **S**

Q3c. **Rate the candidates based on who should be hired**.
(100=most fair to hire; 0=least fair to hire)
[website had a range slider]

- Candidate **E**
- Candidate **C** (*Correct Answer*)
- Candidate **S**

Q4. The figure below shows the skillsets of three job candidates ("J", "R", and "P"). In addition,

this figure shows six black lines, which represent the preferences of the six members of the hiring committee.



Q4a. Based on the graph above, **which job candidate do you think should be hired**?
[website allowed single choice)

- Candidate **J**
- Candidate **R** (*Correct Answer*)
- Candidate **P**

Q4b. **Rank order the candidates based on who you think should be hired**.
(Top = first to hire; bottom = last to hire)
[website allowed click-and-drag functionality for rankings]

- Candidate **J**
- Candidate **R** (*Correct Answer*)
- Candidate **P**

Q4c. **Rate the candidates based on who should be hired**.
(100=most fair to hire; 0=least fair to hire)
[website had a range slider]

- Candidate **J**
- Candidate **R** (*Correct Answer*)
- Candidate **P**

Note: The order of each response choice was randomized for each item.

B3. Who Won the Election Task

[Note: All items and item choices were presented in random order. Version A & B refer to counterbalancing stimuli. Version B is a "mirror reflection" of Version A's data. Participants completed both versions, at different time points. Participants were randomly assigned to complete either Version A or Version B first. The left-side images comprise Version A, and the right-side images comprise Version B. The voting system consistent with choosing a particular candidate is noted in the items below in italics.]

For each item participants answered three questions which appeared below each graph.

1. Based on the above graph, which job candidate do you think should be hired?
- Candidate A
- Candidate B
- Candidate C

2. Rank order the candidates based on who you think should be hired.
(Top = first to hire; bottom = last to hire)
- Candidate A
- Candidate B
- Candidate C

3. Rate the candidates based on who should be hired.
(100=most fair to hire; 0=least fair to hire)
- Candidate A
- Candidate B
- Candidate C

The stimuli items are presented below.

Item 1:

- Candidate K (*Plurality*)
- Candidate A (*Ranked-Choice / Cardinal*)
- Candidate S (*None*)

Item 2:

- Candidate C (*Plurality*)
- Candidate Z (*Cardinal*)
- Candidate G (*Plurality / Ranked-Choice*)

Item 3:

- Candidate A (*None*)
- Candidate O (*Cardinal*)
- Candidate F (*Plurality / Ranked-Choice*)

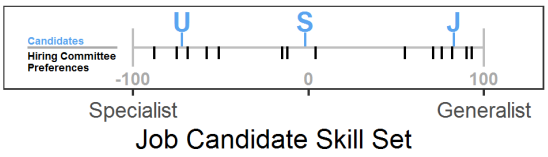
Item 4:

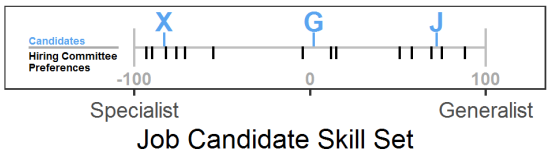
- Candidate N (*Plurality*)
- Candidate L (*Cardinal*)
- Candidate W (*Ranked-Choice*)

Item 5:

- Candidate I (*Plurality*)
- Candidate Q (*Cardinal / Ranked-Choice*)
- Candidate Z (*None*)

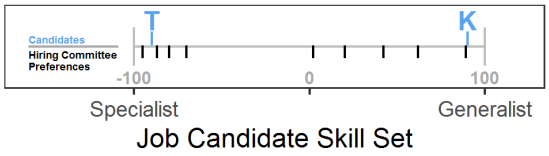
Item 6:

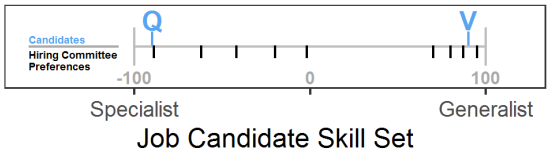




- Candidate U (*Ranked-Choice*)
- Candidate S (*Cardinal*)
- Candidate J (*Plurality*)

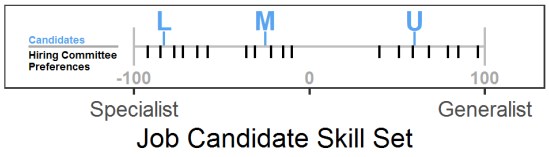
Item 7:

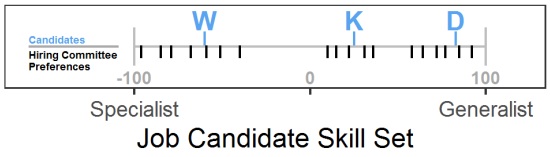




- Candidate T (*Cardinal*)
- Candidate K (*Plurality / Ranked-Choice*)

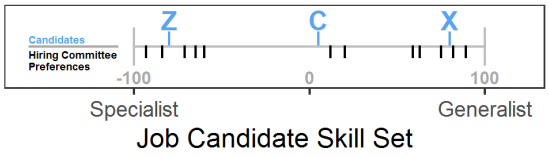
Item 8:

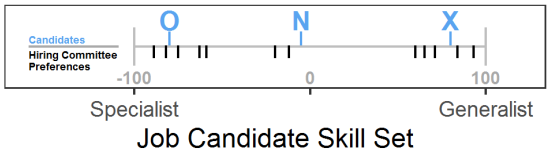




- Candidate L (*Ranked-Choice*)
- Candidate M (*Cardinal*)
- Candidate U (*Plurality*)

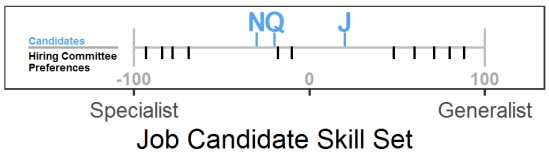
Item 9:

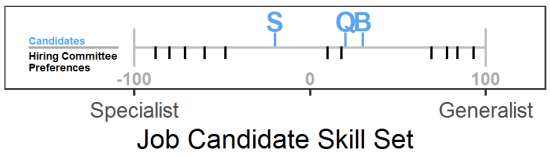




- Candidate Z (*Plurality*)
- Candidate C (*Cardinal*)
- Candidate X (*Plurality / Ranked-Choice*)

Item 10:





- Candidate N (*Ranked-Choice*)
- Candidate Q (*Cardinal*)
- Candidate J (*Plurality*)

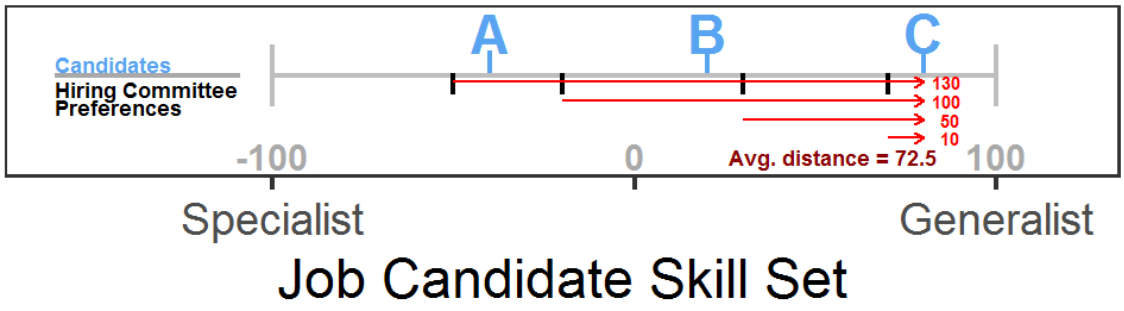
B4. Voting system introduction

[Note: Participants were formally introduced to cardinal, plurality, and ranked-choice voting in this block. The voting system order was randomized for each participant.]

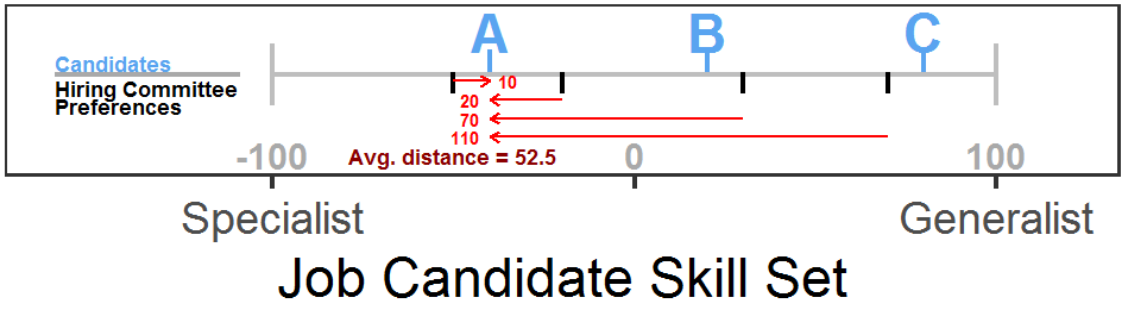
[Cardinal]
One method to determine election outcomes is to focus on the distances between each candidate and the voters. This is called **cardinal voting**. In cardinal voting the candidate that is closest on average to the voters wins. In the figures below we show the distances between the four voters and the three candidates.

As you can see, C is the farthest away from the voters on average - the arrows are the longest and the average distance is 72.5. A and B are both closer to the voters on average, though the average distances for B (42.5) is somewhat shorter than for A (52.5). Thus, according to cardinal voting, **candidate B would win**.

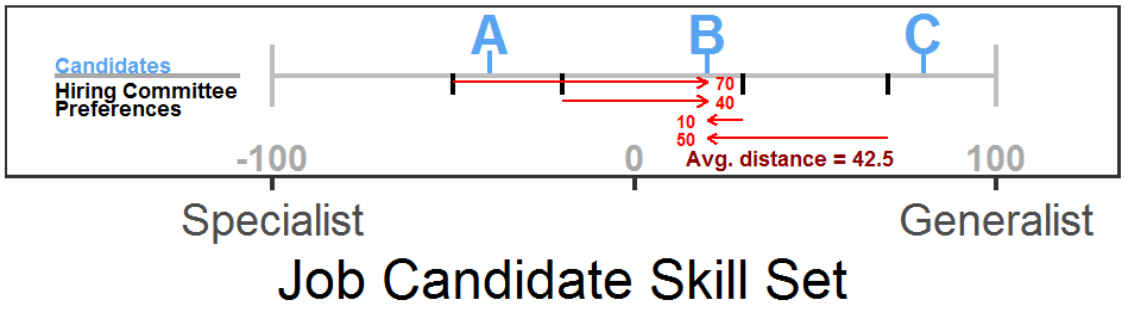
Distances to candidate C



Distances to candidate A



Distances to candidate B

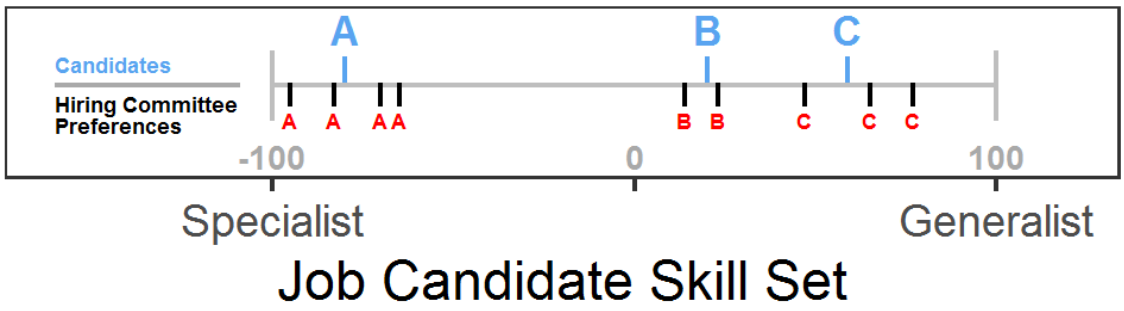


How fair is cardinal voting? Rate from very fair (100) to not very fair (0). [slider question format]

[Plurality]

One method to determine election outcomes is to allow all voters to choose only one candidate and the candidate with the most votes wins. This is called **plurality voting**. Plurality voting prioritizes selecting a winner who has the single most support. A winning candidate in a plurality system needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.

For example, in the graph below "Candidate A" has 4 out of 9 hiring committee members support (as noted by the red letters below the hiring preferences). "Candidate B" has 2 out of the 9 hiring committee members support. "Candidate C" has 3 out of the 9 hiring committee members support. Using a Plurality voting system, Candidate A would win, even though they did not receive more than 50% of the votes.



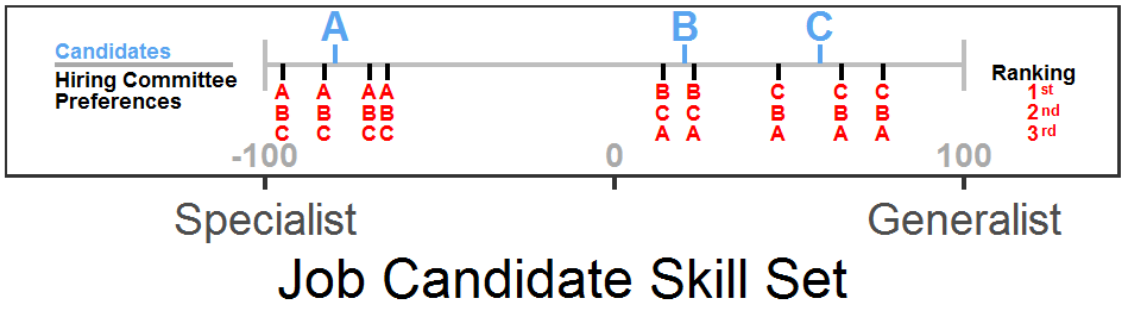
How fair is plurality voting? Rate from very fair (100) to not very fair (0). [slider question format]

[Ranked-Choice Voting]

One method to determine election outcomes is to focus on how each voter ranks the candidates (e.g., first, second, third preference). This is called **ranked-choice voting**. Ranked-choice voting prioritizes majority approval and a winning candidate must have over 50% of the total votes. This works by first counting all voters' first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is

"eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support. If so, this candidate wins. If not, the process is repeated.

For example, in the graph below, there are four voters who are very close to Candidate A, so A gets first place for these four. There are two voters who are closest to B, and there are three voters who are closest to C. Although A has the most first place votes (4 out of 9), A does not have over 50% of the votes. Because B has the fewest first place votes, B would be removed as a candidate, and B's votes would be reassigned. In this case, the two voters closest to B would be reassigned to C, because they are closer to C (their 2nd choice) than to A (their 3rd choice). Once B is removed, there are 5 votes for C, and 4 votes for A, so Candidate C would win.



How fair is ranked-choice voting? Rate from very fair (100) to not very fair (0). [slider question format]

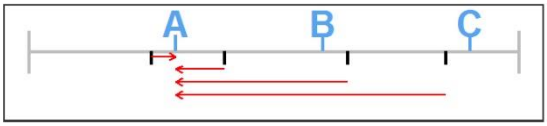
B5. Voting system comprehension

[Note: Participants completed the voting system comprehension twice in study 2. The two tests were mirrored versions of each other, any differences are superficial. Participants were randomized to complete either version A or version B first.]

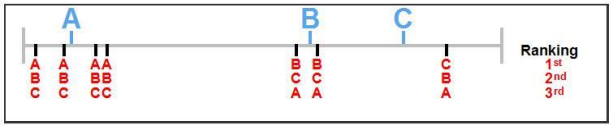
Instructions:
The next few pages will ask you to identify which job candidate should be hired based on the different ways to determine elections that you just learned about.

We will present the logos below to help you remember how the different methods work.

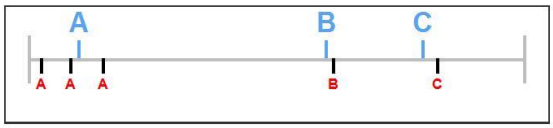
- **Cardinal voting** focuses on the distances between each candidate and the voters. The candidate that is closest on average to the voters wins.



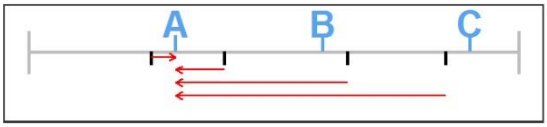
- **Ranked-Choice voting** first counts all voters' first preference and sees if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support.



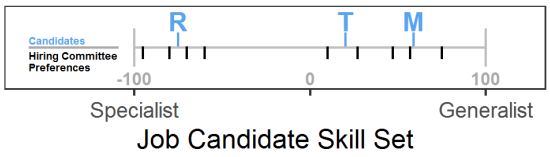
- **Plurality voting** has each voter pick one candidate and the winning candidate is the one with the most votes, even if they do not have over 50% of the total votes.



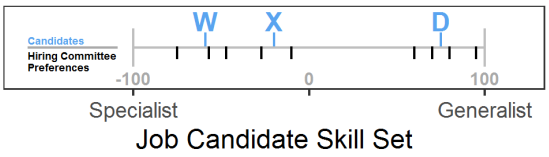
Learning item 1a:
This question is about **Cardinal Voting**



Version A



Version B

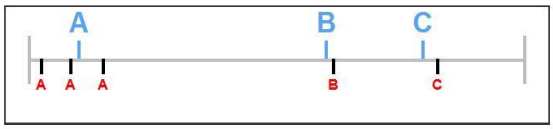


In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Cardinal voting system**?

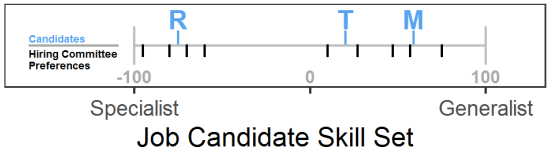
[*Remember: In Cardinal voting a winning candidate is the candidate who is closest on average to all of the voters.*]

- Candidate **R**
- Candidate **T** (*correct answer*)
- Candidate **M**

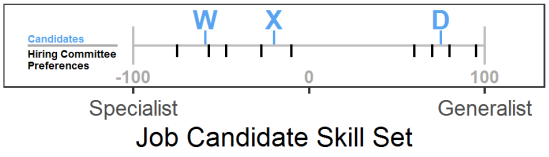
Learning item 1b:
This question is about **Plurality Voting**



Version A



Version B

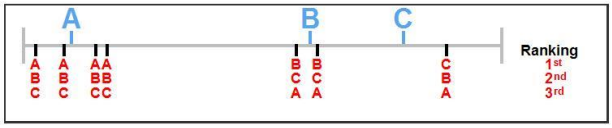


In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Plurality voting system**?

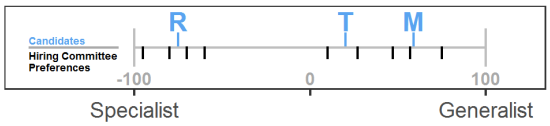
[*Remember: In plurality voting a winning candidate needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.*]

- Candidate **R** (*correct answer*)
- Candidate **T**
- Candidate **M**

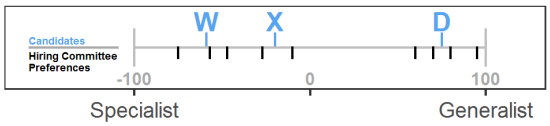
Learning item 1c:
This question is about **Ranked-Choice Voting**



Version A



Version B



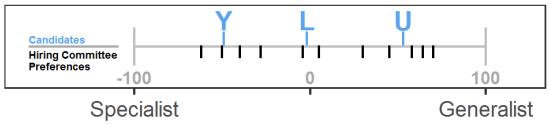
In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Ranked-Choice voting system**?

[*Remember: In Ranked-Choice voting a winning candidate must have over 50% of the total votes. This works by first counting all voters first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support.*]

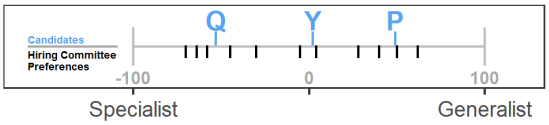
- Candidate **R**
- Candidate **T**
- Candidate **M** (*correct answer*)

Learning item 2a:
This question is about **Cardinal Voting**

Version A



Version B

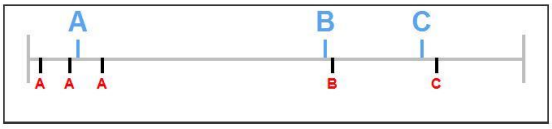


In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Cardinal voting system**?

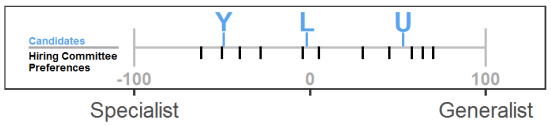
[*Remember: In Cardinal voting a winning candidate is the candidate who is closest on average to all of the voters.*]

- Candidate **Y**
- Candidate **L** (*correct answer*)
- Candidate **U**

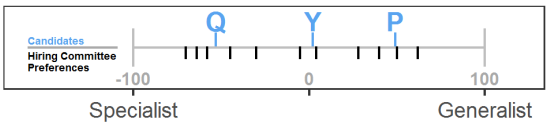
Learning item 2b:
This question is about **Plurality Voting**



Version A



Version B

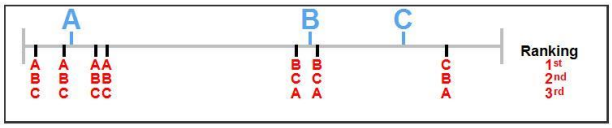


In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Plurality voting system**?

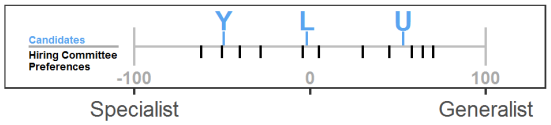
[*Remember: In plurality voting a winning candidate needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.*]

- Candidate **Y** (*correct answer*)
- Candidate **L**
- Candidate **U**

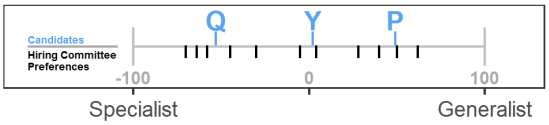
Learning item 2c:
This question is about **Ranked-Choice Voting**



Version A



Version B



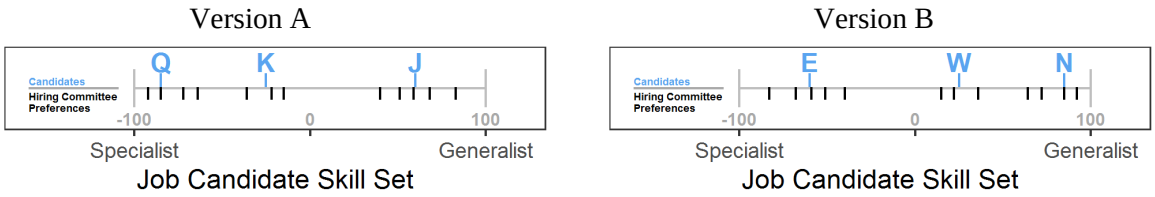
In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Ranked-Choice voting system**?

[*Remember: In Ranked-Choice voting a winning candidate must have over 50% of the total votes. This works by first counting all voters first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support.*]

- Candidate **Y**
- Candidate **L**

- Candidate **U** (*correct answer*)

Learning item 3a:
This question is about **Cardinal Voting**

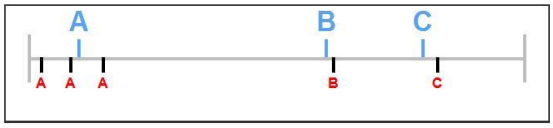


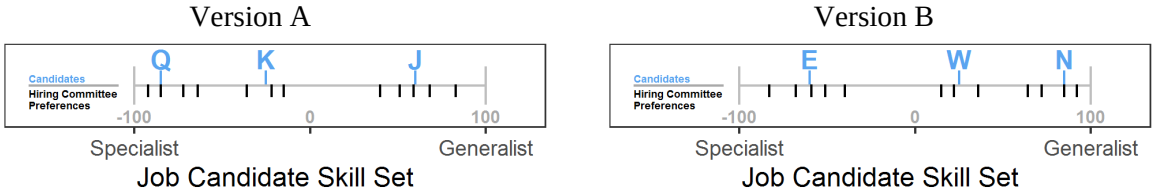
In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Cardinal voting system**?

[*Remember: In Cardinal voting a winning candidate is the candidate who is closest on average to all of the voters.*]

- Candidate **Q**
- Candidate **K** (*correct answer*)
- Candidate **J**

Learning item 3b:
This question is about **Plurality Voting**





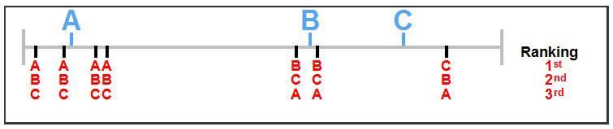
In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Plurality voting system**?

[*Remember: In plurality voting a winning candidate needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.*]

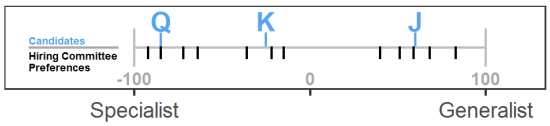
- Candidate **Q**
- Candidate **K**

- Candidate **J** (*correct answer*)

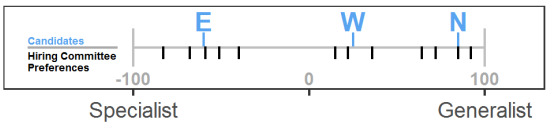
Learning item 3c:
This question is about **Ranked-Choice Voting**



Version A



Version B



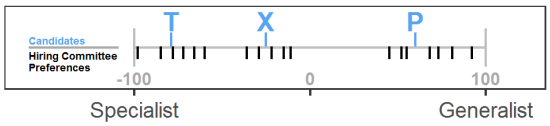
In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Ranked-Choice voting system**?

[*Remember: In Ranked-Choice voting a winning candidate must have over 50% of the total votes. This works by first counting all voters first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support.*]

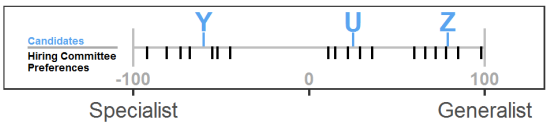
- Candidate **Q** (*correct answer*)
- Candidate **K**
- Candidate **J**

Learning item 4a:
This question is about **Cardinal Voting**

Version A



Version B

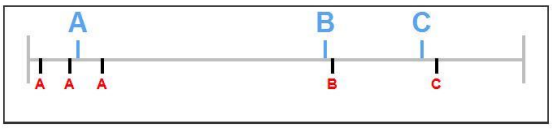


In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Cardinal voting system**?

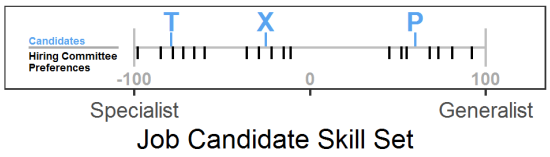
[*Remember: In Cardinal voting a winning candidate is the candidate who is closest on average to all of the voters.*]

- Candidate **T**
- Candidate **X** (*correct answer*)
- Candidate **P**

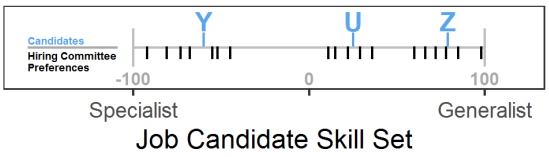
Learning item 4b:
This question is about **Plurality Voting**



Version A



Version B

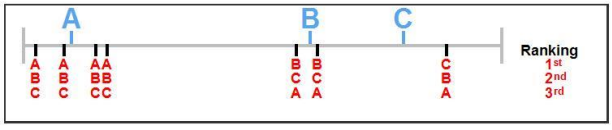


In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Plurality voting system**?

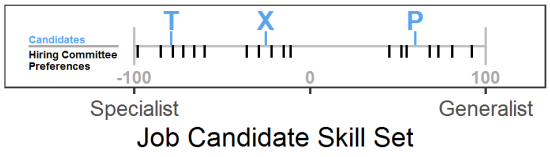
[*Remember: In plurality voting a winning candidate needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.*]

- Candidate **T**
- Candidate **X**
- Candidate **P** (*correct answer*)

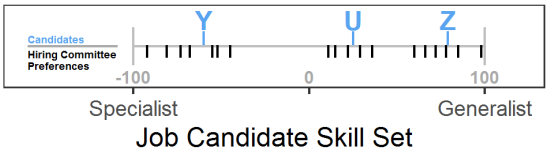
Learning item 4c:
This question is about **Ranked-Choice Voting**



Version A



Version B



In regards to the information presented in the graph above... which candidate would be chosen as the winner in a **Ranked-Choice voting system**?

[*Remember: In Ranked-Choice voting a winning candidate must have over 50% of the total votes. This works by first counting all voters first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support.*]

- Candidate **T** (*correct answer*)
- Candidate **X**
- Candidate **P**

B6. Declared voting system preferences

Q1. **Which of the three voting systems below do you think is most fair?**
[see figures and descriptions below if you need a reminder]

- Ranked-Choice
- Plurality
- Cardinal

Q2. Order the voting systems below (drag and drop list items), from most fair (top) to least fair (bottom).

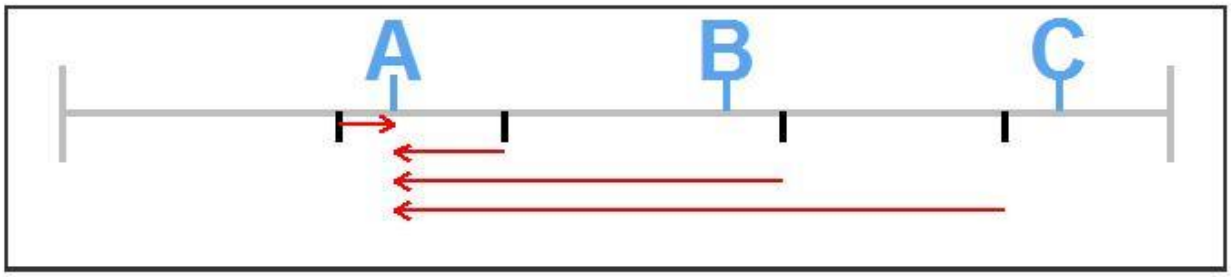
- Ranked-Choice
- Plurality
- Cardinal

Q3. Rate the voting systems below, from very fair (100) to not very fair (0).

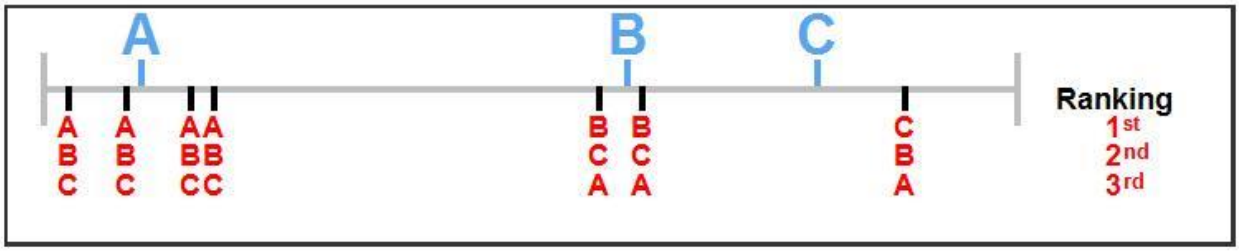
- Ranked-Choice
- Plurality
- Cardinal

Remember:

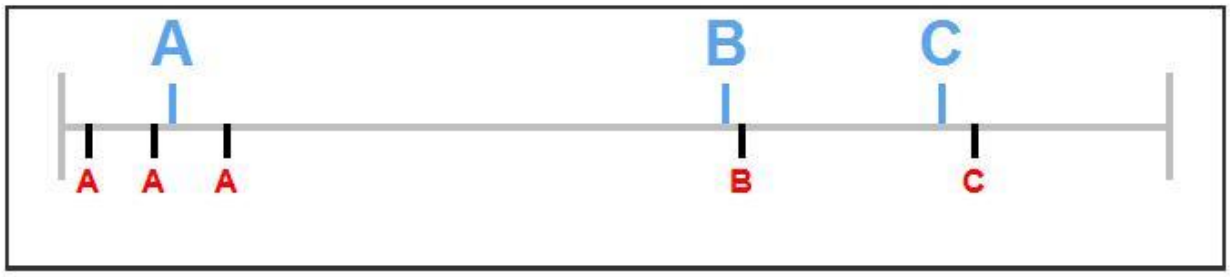
- In **cardinal** voting a winning candidate is the candidate who is closest on average to the all of the voters. The figure below provides an example of how cardinal voting works for calculating the score of a Candidate by measuring the distance from each member of the hiring committee.



- *In **ranked-choice** voting a winning candidate must have over 50% of the total votes. This works by first counting all voters first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support. The figure below shows how ranked-choice voting works. The red letters represent preference rankings for each member of the hiring committee. The top row is the first pick, the middle row is the second favorite, and the bottom row is the least* favorite.



- In **plurality** voting a winning candidate needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win. The figure below shows how plurality voting works.



B7. Teaching Interventions

Note: Participants were randomly assigned to one of 4 teaching interventions: retrieval practice, re-study, discovery learning, and a control (no intervention) condition. Every condition (except the control condition) was exposed to three different "voting system quirks" which were the center squeeze phenomenon, violation of the majority criterion, and the spoiler effect—each of which show notable weaknesses in ranked-choice, cardinal, and plurality voting, respectively.

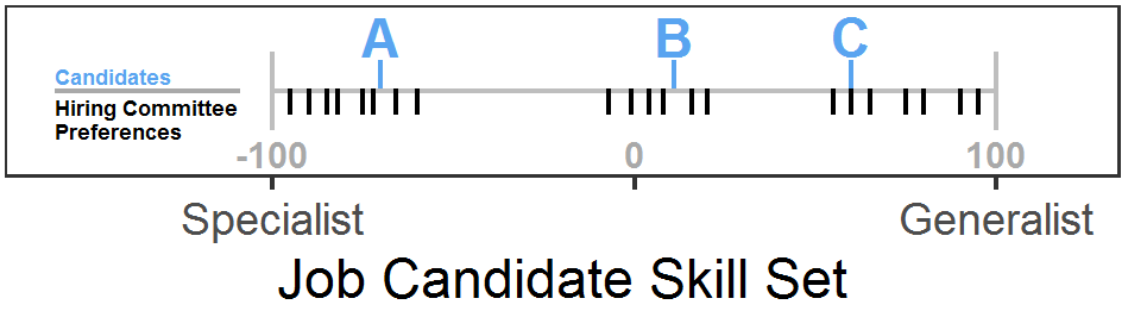
Condition: Retrieval Practice

[Retrieval Practice: Center Squeeze Example]

On the next few pages you will be walked through a specific voting example called the "Center Squeeze Phenomenon," which shows how ranked-choice voting can hurt moderate or "middle-of-the-road" candidates.

...but first we will ask you to determine the results of an election using different voting systems.

Item 1A (Retrieval Practice, Center Squeeze, Cardinal item).



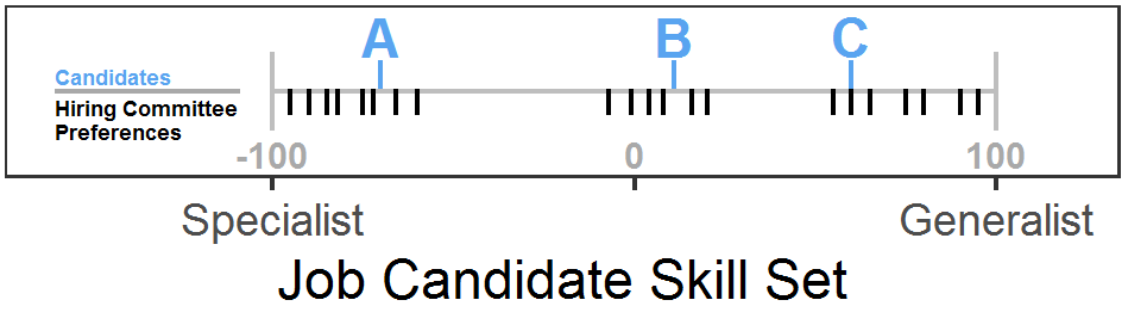
Based on the above graph... which candidate would be chosen as the winner in a **cardinal voting system**?

- Candidate **A**
- Candidate **B** (*correct answer*)
- Candidate **C**

If correct answer is chosen, following message was displayed:
Correct! **Candidate B** would be chosen as the winner in a **cardinal voting system.**

If incorrect answer is chosen, following message was displayed:
Incorrect! **Candidate B** would be chosen as the winner in a **cardinal voting system.**

Item 1B (Retrieval Practice, Center Squeeze, Plurality item).



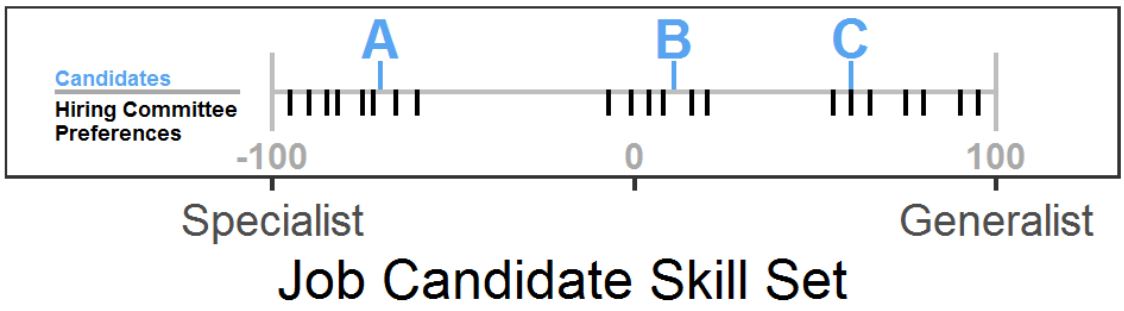
Based on the above graph... which candidate would be chosen as the winner in a **plurality voting system**?

- Candidate **A** (*correct answer*)
- Candidate **B**
- Candidate **C**

If correct answer is chosen, following message was displayed:
Correct! **Candidate A** would be chosen as the winner in a **plurality voting system.**

If incorrect answer is chosen, following message was displayed:
Incorrect! **Candidate A** would be chosen as the winner in a **plurality voting system.**

Item 1C (Retrieval Practice, Center Squeeze, RCV item).



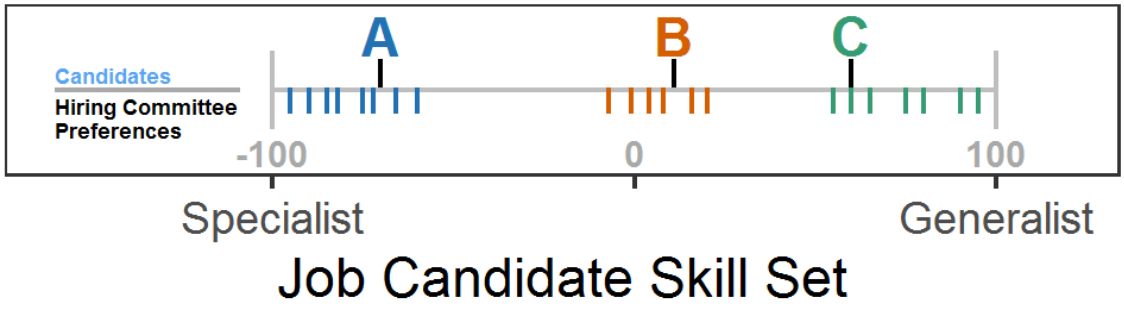
Based on the above graph... which candidate would be chosen as the winner in a **ranked-choice voting system**?

- Candidate **A**
- Candidate **B**
- Candidate **C** (*correct answer*)

If correct answer is chosen, following message was displayed:
Correct! **Candidate C** would be chosen as the winner in a **ranked-choice voting system.**

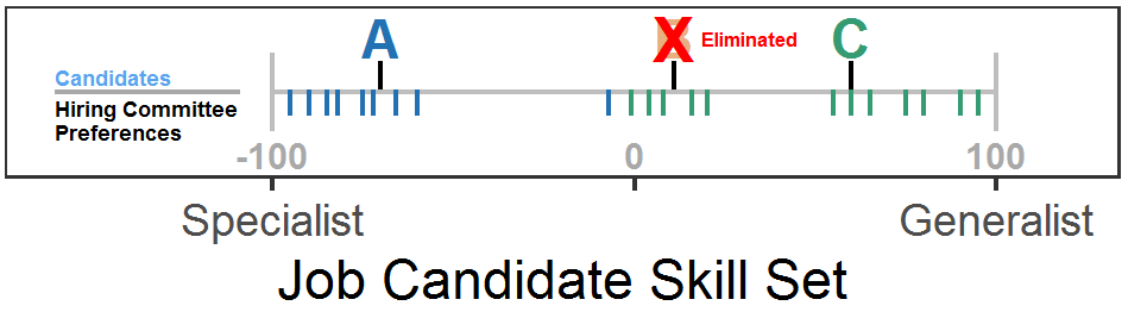
If incorrect answer is chosen, following message was displayed:
Incorrect! **Candidate C** would be chosen as the winner in a **ranked-choice voting system.**

You will now be walked through the “Center Squeeze Phenomenon."



Who wins according to plurality and ranked-choice voting?

- 8 voters have candidate A as their favorite (8/21 voters or 38%). Therefore, **A wins according to plurality**.
- 6 voters have candidate B as their favorite (6/21 voters or 29%), and
- 7 voters have candidate C as their favorite (7/21 voters or 33%). **C wins according to ranked-choice voting**, because we would eliminate B, since no candidate has over 50%, giving C 12 of the 21 votes (57%).



Who wins according to cardinal voting?

- Candidate A is the furthest from the voters with an average distance of 74.4,
- Candidate B is the closest to the voters with an average distance of 57.6. Therefore, **B wins according to cardinal voting**.
- Candidate C is the second closest to the voters with an average distance of 73.1.

Why is this "center squeeze" example problematic for ranked-choice and plurality voting?
In this "center squeeze" example, there is a fairly good middle-of-the-road candidate, but ranked-choice and plurality voting instead choose extreme candidates.

Think about it this way - imagine that the voters were the same but instead of having three candidates there were only two. If this was just an election with A vs B, B would win. If this was

just an election of B vs C, B would win. If there was an election of A vs C, C would win. So in the head-to-head comparisons B wins two of the three elections.

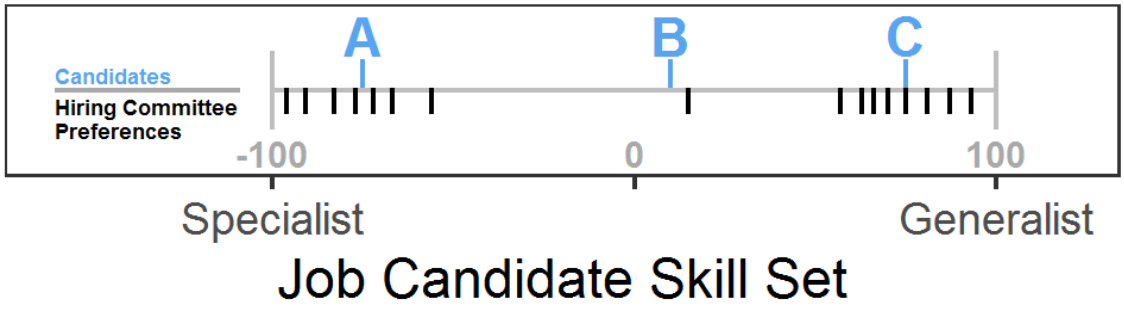
In sum, this example shows that ranked-choice voting can favor a more extreme candidate (C), and plurality can also favor a more extreme candidate (A), as opposed to more middle-of-the-road candidate picked by cardinal (B).

[Retrieval Practice: Majority Criterion Violation Example]

On the next few pages you will be walked through a specific voting example called the "Majority Criterion Violation," which demonstates how in cardinal voting it is possible for a candidate to lose despite having a majority of the first places votes.

...but first we will ask you to determine the results of the an election using different voting systems.

Item 1A (Retrieval Practice, Majority Criterion, Cardinal item).
Based on the above graph... which candidate would be chosen as the winner in a **cardinal voting system**?

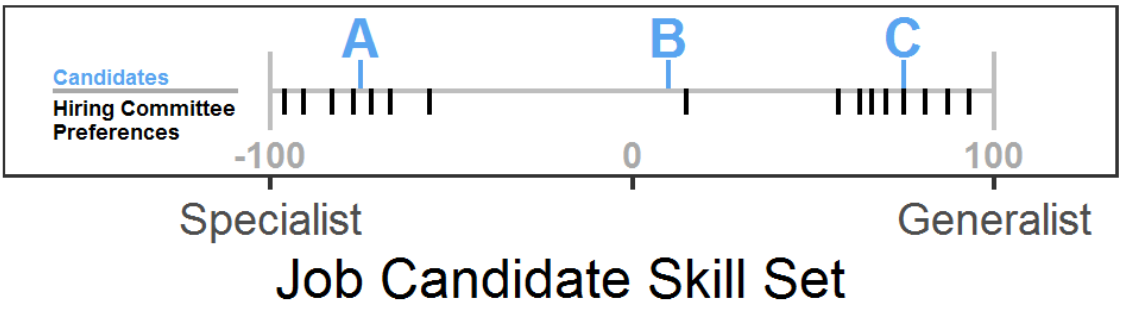


- Candidate **A**
- Candidate **B** (*correct answer*)
- Candidate **C**

If correct answer is chosen, following message was displayed:
Correct! **Candidate B** would be chosen as the winner in a **cardinal voting system.**

If incorrect answer is chosen, following message was displayed:
Incorrect! **Candidate B** would be chosen as the winner in a **cardinal voting system.**

Item 1B (Retrieval Practice, Majority Criterion, Plurality item).
Based on the above graph... which candidate would be chosen as the winner in a **plurality voting system**?

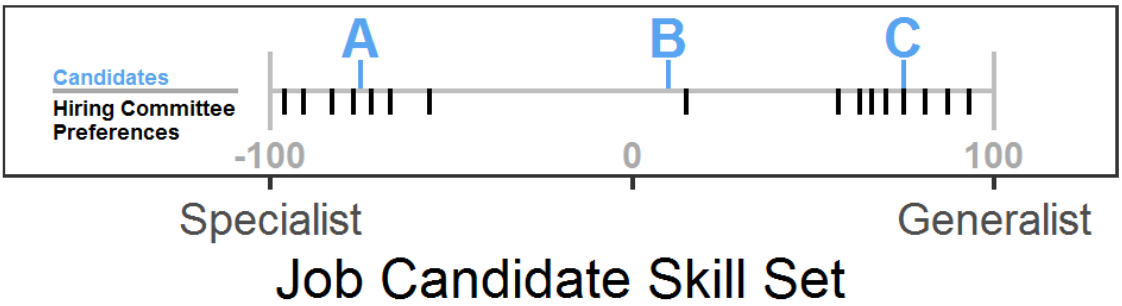


- Candidate **A**
- Candidate **B**
- Candidate **C** (*correct answer*)

If correct answer is chosen, following message was displayed:
Correct! **Candidate C** would be chosen as the winner in a **plurality voting system.**

If incorrect answer is chosen, following message was displayed:
Incorrect! **Candidate C** would be chosen as the winner in a **plurality voting system.**

Item 1C (Retrieval Practice, Majority Criterion, RCV item).
Based on the above graph... which candidate would be chosen as the winner in a **ranked-choice voting system**?

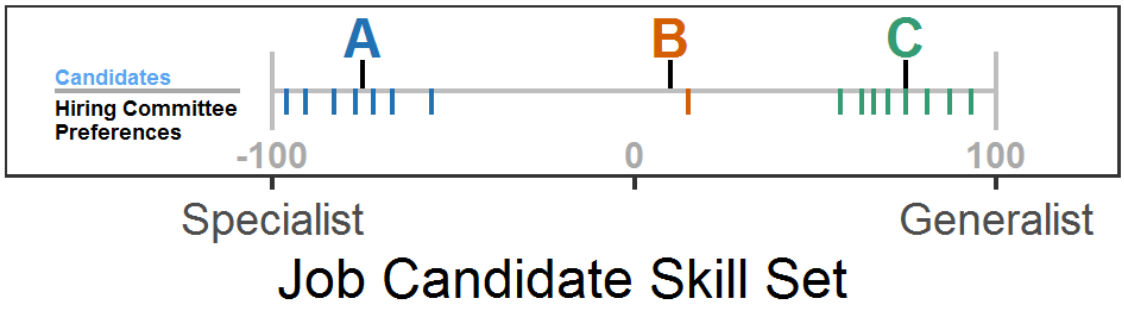


- Candidate **A**
- Candidate **B**
- Candidate **C** (*correct answer*)

If correct answer is chosen, following message was displayed:
Correct! **Candidate C** would be chosen as the winner in a **ranked-choice voting system.**

If incorrect answer is chosen, following message was displayed:
Incorrect! **Candidate C** would be chosen as the winner in a **ranked-choice voting system.**

You will now be walked through the “Majority Criterion Violation."



Who wins according to plurality and ranked-choice voting?

- 7 voters have candidate A as their favorite (7/16 voters or 44%), and
- 1 voter has candidate B as their favorite (1/16 voters or 6%), and
- 8 voters have candidate C as their favorite (8/16 voters or 50%). Therefore, **C wins according to plurality**. **C also wins according to ranked-choice voting**, because we would eliminate B, since no candidate has over 50% of the total votes, giving C 9 of the 16 votes (56%).

Who wins according to cardinal voting?

- Candidate A is the furthest from all the voters with an average distance of 84.9,
- Candidate B is the closest to the voters with an average distance of 70.6. Therefore, **B wins according to cardinal voting**.
- Candidate C is the second closest to the voters with an average distance of 75.4.

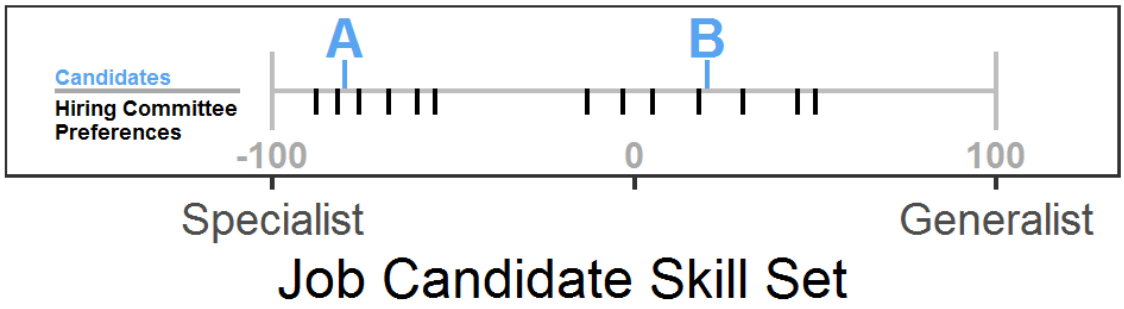
Why is this example about the "majority criterion" problematic for carinal voting?
In certain situations cardinal voting can lead to winners that very few voters would choose as their first pick. In the example above, B won according to cardinal, but only one voter was close to B. Additionally, candidate C lost according to cardinal voting but many more voters are close to C than B.

[Retrieval Practice: Spoiler Effect Example]

On the next few pages you will be walked through a specific voting example called the “Spoiler Effect," which shows how outcomes in plurality voting can be altered by the addition of a third candidate.

...but first we will ask you to determine the results of the an election using different voting systems.

Item 1A (Retrieval Practice, Spoiler Effect, Cardinal item).
Based on the above graph... which candidate would be chosen as the winner in a **cardinal voting system**?

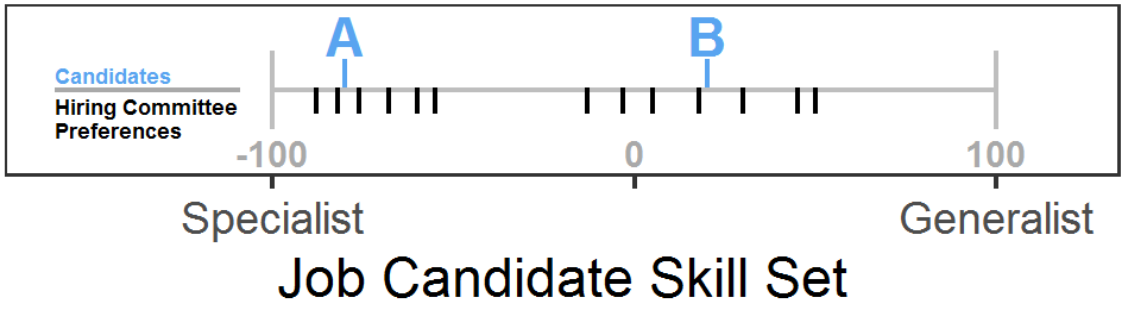


- Candidate **A**
- Candidate **B** (*correct answer*)

If correct answer is chosen, following message was displayed:
Correct! **Candidate B** would be chosen as the winner in a **cardinal voting system.**

If incorrect answer is chosen, following message was displayed:
Incorrect! **Candidate B** would be chosen as the winner in a **cardinal voting system.**

Q1B (Retrieval Practice, Spoiler Effect, Plurality item).
Based on the above graph... which candidate would be chosen as the winner in a **plurality voting system**?



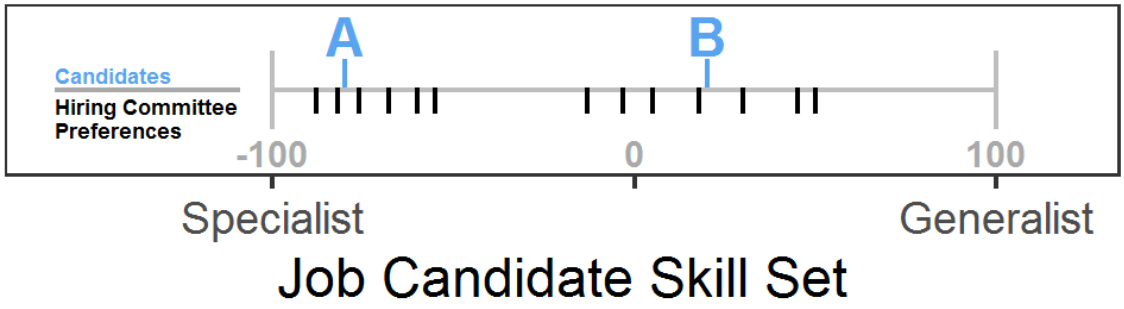
- Candidate **A**
- Candidate **B** (*correct answer*)

If correct answer is chosen, following message was displayed:
Correct! **Candidate B** would be chosen as the winner in a **plurality voting system.**

If incorrect answer is chosen, following message was displayed:
Incorrect! **Candidate B** would be chosen as the winner in a **plurality voting system.**

Q1C (Retrieval Practice, Spoiler Effect, RCV item).

Based on the above graph... which candidate would be chosen as the winner in a **ranked-choice voting system**?

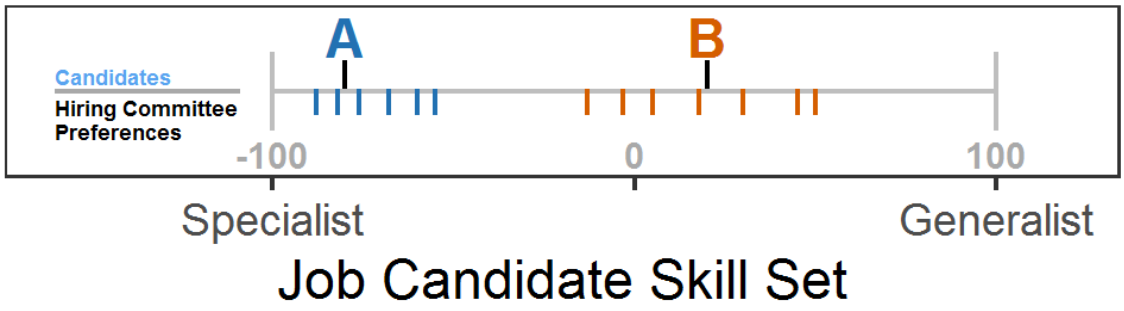


- Candidate **A**
- Candidate **B** (*correct answer*)

If correct answer is chosen, following message was displayed:
Correct! **Candidate B** would be chosen as the winner in a **ranked-choice voting system.**

If incorrect answer is chosen, following message was displayed:
Incorrect! **Candidate B** would be chosen as the winner in a **ranked-choice voting system.**

You will now be walked through a specific voting example called the "Spoiler Effect."



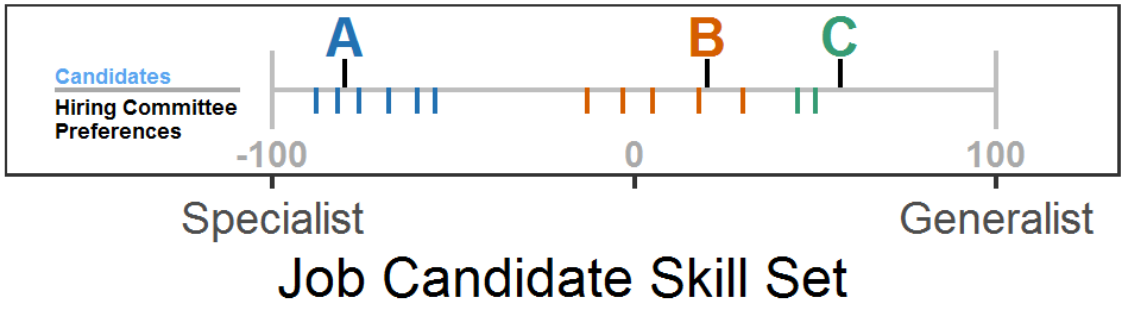
Who wins according to plurality and ranked-choice voting?

- 6 voters have candidate A as their favorite (6/13 voters or 46%), and
- 7 voters have candidate B as their favorite (7/13 voters or 54%). Therefore, **B wins according to both plurality and ranked-choice voting**, since the candidate has over 50% of the total votes.

Who wins according to cardinal voting?

- Candidate A is the furthest from the voters with an average distance of 58.7,
- Candidate B is closest to the voters with an average distance of 52.8. Therefore, **B wins according to cardinal voting**.

Imagine the same scenario with a slight difference…
Imagine having three candidates instead of two. Specifically, imagine if there was a third candidate who was not very popular, like candidate C in the graph below.



Why is this example about having a third candidate problematic for plurality voting?
In plurality voting, the presence of an unpopular third candidate (like Candidate C) can change who wins. This could be viewed as strange given that very few people like Candidate C - so why should C's decision to join the election affect who wins if C has no chance of winning? This does not happen with ranked-choice or cardinal voting.

Condition: Re-Study
[Re-Study: Center Squeeze Example]

On the next few pages you will be walked through a specific voting example called the "Center Squeeze Phenomenon," which shows how ranked-choice voting can hurt moderate or "middle-of-the-road" candidates

---

Item 1A (Re-Study, Center Squeeze, Cardinal item).

As a reminder:
Cardinal voting focuses on the distances between each candidate and the voters. The candidate that is closest on average to the voters wins.

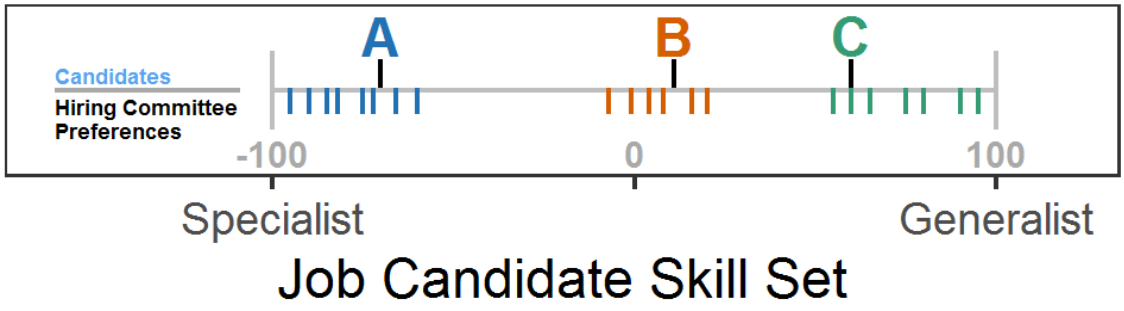
Item 1B (Re-Study, Center Squeeze, Plurality item).

As a reminder:
Plurality voting has each voter pick one candidate and the winning candidate is the one with the most votes, even if they do not have over 50% of the total votes.

Item 1C (Re-Study, Center Squeeze, RCV item).

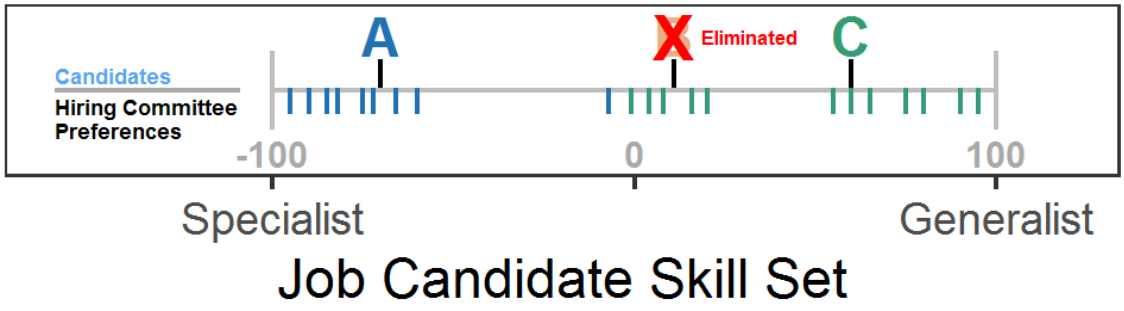
As a reminder:
Ranked-Choice voting first counts all voters' first preference and sees if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is “eliminated” and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support.

You will now be walked through the “Center Squeeze Phenomenon."



Who wins according to plurality and ranked-choice voting?

- 8 voters have candidate A as their favorite (8/21 voters or 38%). Therefore, **A wins according to plurality**.
- 6 voters have candidate B as their favorite (6/21 voters or 29%), and
- 7 voters have candidate C as their favorite (7/21 voters or 33%). **C wins according to ranked-choice voting**, because we would eliminate B, since no candidate has over 50%, giving C 12 of the 21 votes (57%).



Who wins according to cardinal voting?

- Candidate A is the furthest from the voters with an average distance of 74.4,
- Candidate B is the closest to the voters with an average distance of 57.6. Therefore, **B wins according to cardinal voting**.
- Candidate C is the second closest to the voters with an average distance of 73.1.

Why is this "center squeeze" example problematic for ranked-choice and plurality voting?
In this "center squeeze" example, there is a fairly good middle-of-the-road candidate, but ranked-choice and plurality voting instead choose extreme candidates.

Think about it this way - imagine that the voters were the same but instead of having three candidates there were only two. If this was just an election with A vs B, B would win. If this was just an election of B vs C, B would win. If there was an election of A vs C, C would win. So in the head-to-head comparisons B wins two of the three elections.

In sum, this example shows that ranked-choice voting can favor a more extreme candidate (C), and plurality can also favor a more extreme candidate (A), as opposed to more middle-of-the-road candidate picked by cardinal (B).

[Re-Study: Majority Criterion Violation]

On the next few pages you will be walked through a specific voting example called the “Majority Criterion Violation," which demonstrates how in cardinal voting it is possible for a candidate to lose despite having a majority of the first places votes.

Item 1A (Re-Study, Majority Criterion, Cardinal item).

As a reminder:
Cardinal voting focuses on the distances between each candidate and the voters. The candidate that is closest on average to the voters wins.

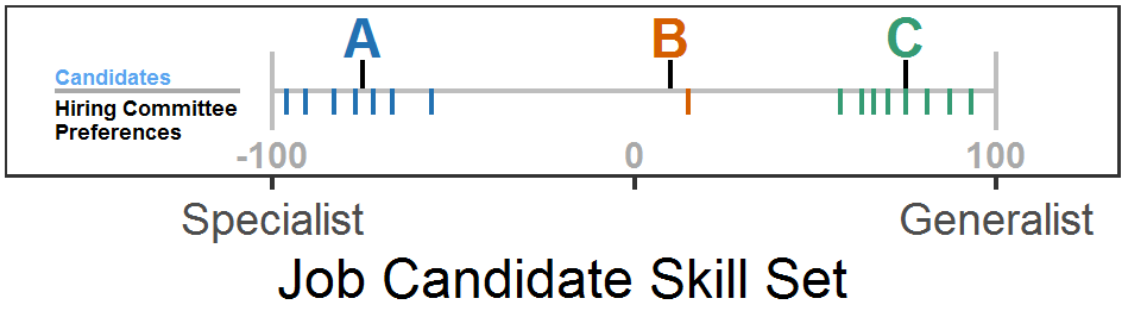
Item 1B (Re-Study, Majority Criterion, Plurality item).

As a reminder:
Plurality voting has each voter pick one candidate and the winning candidate is the one with the most votes, even if they do not have over 50% of the total votes.

Item 1C (Re-Study, Majority Criterion, RCV item).

As a reminder:
Ranked-Choice voting first counts all voters' first preference and sees if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is “eliminated” and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support.

You will now be walked through the “Majority Criterion Violation."



Who wins according to plurality and ranked-choice voting?

- 7 voters have candidate A as their favorite (7/16 voters or 44%), and
- 1 voter has candidate B as their favorite (1/16 voters or 6%), and
- 8 voters have candidate C as their favorite (8/16 voters or 50%). Therefore, **C wins according to plurality**. **C also wins according to ranked-choice voting**, because we would eliminate B, since no candidate has over 50% of the total votes, giving C 9 of the 16 votes (56%).

Who wins according to cardinal voting?

- Candidate A is the furthest from all the voters with an average distance of 84.9,
- Candidate B is the closest to the voters with an average distance of 70.6. Therefore, **B wins according to cardinal voting**.
- Candidate C is the second closest to the voters with an average distance of 75.4.

Why is this example about the "majority criterion" problematic for carinal voting?
In certain situations cardinal voting can lead to winners that very few voters would choose as their first pick. In the example above, B won according to cardinal, but only one voter was close to B. Additionally, candidate C lost according to cardinal voting but many more voters are close to C than B.

[Re-Study: Spoiler Effect]

On the next few pages you will be walked through a specific voting example called the “Spoiler Effect," which shows how outcomes in plurality voting can be altered by the addition of a third candidate.

Item 1A (Re-Study, Center Squeeze, Cardinal item).

As a reminder:
Cardinal voting focuses on the distances between each candidate and the voters. The candidate that is closest on average to the voters wins.

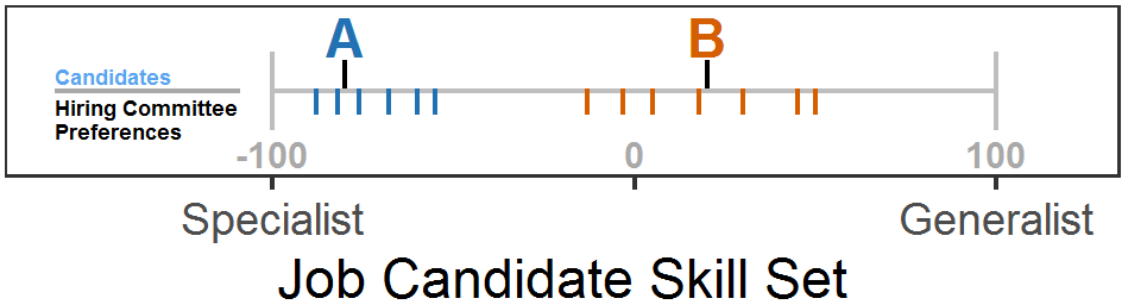
Item 1B (Re-Study, Center Squeeze, Plurality item).

As a reminder:
Plurality voting has each voter pick one candidate and the winning candidate is the one with the most votes, even if they do not have over 50% of the total votes.

Item 1C (Re-Study, Center Squeeze, RCV item).

As a reminder:
Ranked-Choice voting first counts all voters' first preference and sees if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support.

You will now be walked through a specific voting example called the "Spoiler Effect."



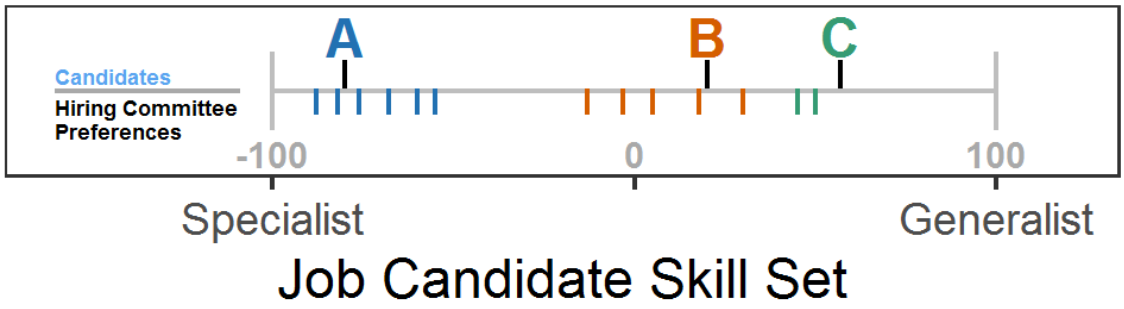
Who wins according to plurality and ranked-choice voting?

- 6 voters have candidate A as their favorite (6/13 voters or 46%), and
- 7 voters have candidate B as their favorite (7/13 voters or 54%). Therefore, **B wins according to both plurality and ranked-choice voting**, since the candidate has over 50% of the total votes.

Who wins according to cardinal voting?

- Candidate A is the furthest from the voters with an average distance of 58.7,
- Candidate B is closest to the voters with an average distance of 52.8. Therefore, **B wins according to cardinal voting**.

Imagine the same scenario with a slight difference…
Imagine having three candidates instead of two. Specifically, imagine if there was a third candidate who was not very popular, like candidate C in the graph below.

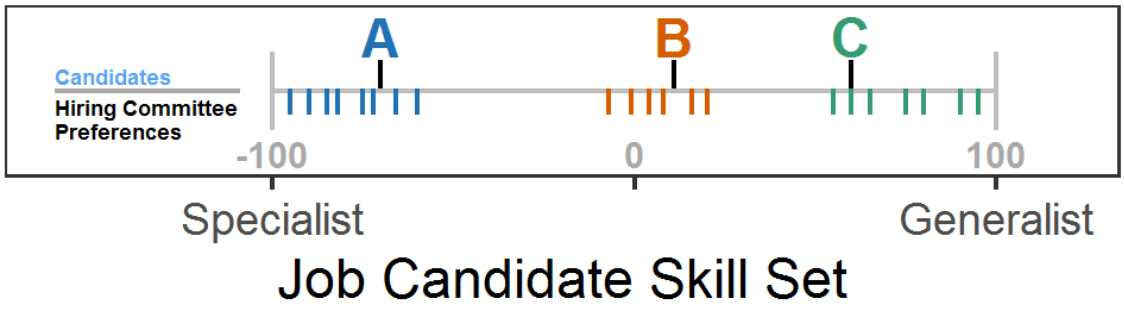


Why is this example about having a third candidate problematic for plurality voting?
In plurality voting, the presence of an unpopular third candidate (like Candidate C) can change who wins. This could be viewed as strange given that very few people like Candidate C - so why should C's decision to join the election affect who wins if C has no chance of winning? This does not happen with ranked-choice or cardinal voting.

Condition: Discovery Learning
[Discovery Learning: Center Squeeze Example]

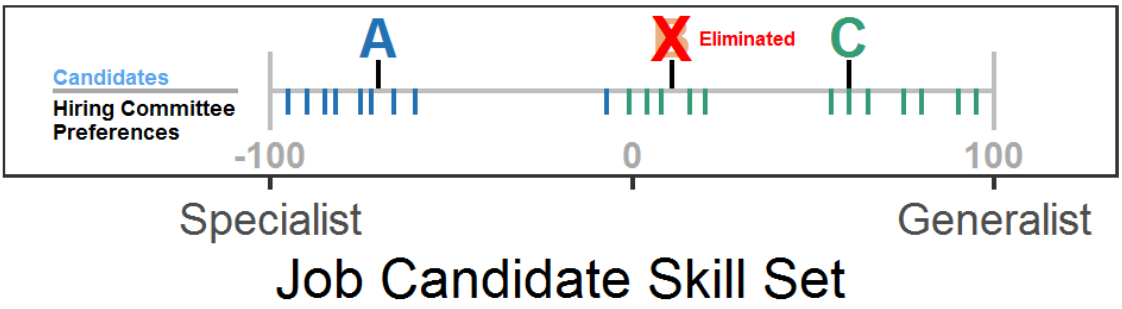
On the next few pages you will be walked through a specific voting example called the "Center Squeeze Phenomenon," which shows how ranked-choice voting can hurt moderate or "middle-of-the-road" candidates

---



Who wins according to plurality and ranked-choice voting?

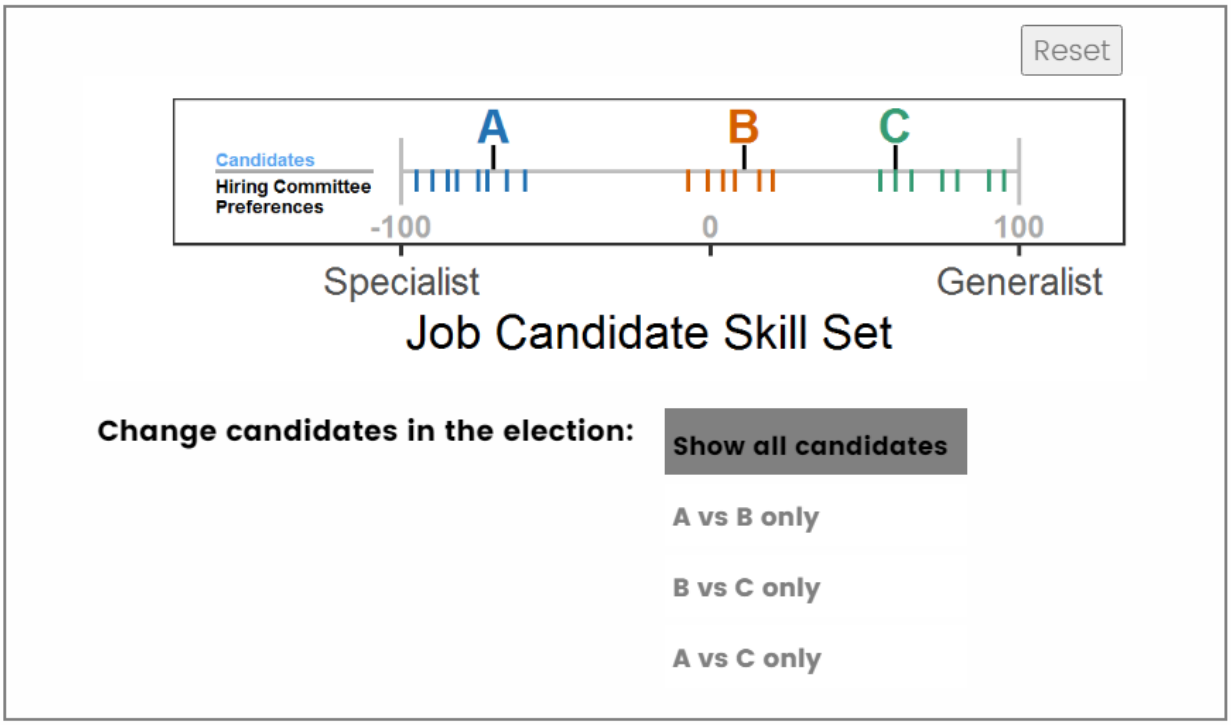
- 8 voters have candidate A as their favorite (8/21 voters or 38%). Therefore, **A wins according to plurality**.
- 6 voters have candidate B as their favorite (6/21 voters or 29%), and
- 7 voters have candidate C as their favorite (7/21 voters or 33%). **C wins according to ranked-choice voting**, because we would eliminate B, since no candidate has over 50%, giving C 12 of the 21 votes (57%).



Who wins according to cardinal voting?

- Candidate A is the furthest from the voters with an average distance of 74.4,
- Candidate B is the closest to the voters with an average distance of 57.6. Therefore, **B wins according to cardinal voting**.
- Candidate C is the second closest to the voters with an average distance of 73.1.

We created the interactive graph below to help you answer the following questions.



[Note: The above figure is a screenshot of the applet. The participant can show/hide candidates by clicking one of the four buttons next to "change candidates in the election." When a candidate is removed the hiring committee member preferences are re-colored by which (present) candidate they are closest to.]

One way to assess the strength of competing candidate is to examine the outcomes from head-to-head matchups (i.e., limiting the options to two candidates, in a 1 vs. 1 competition). Using the

interactive graph above, examine the different 1 vs. 1 comparisons. Are these comparisons useful? And if so, what do they reveal about problem with ranked-choice and plurality voting?

The name center squeeze phenomenon refers to conditions where middle-of-the-road candidates are disadvantaged in ranked-choice voting (by middle, we mean candidates somewhere between two other candidates). What conditions do you think lead to middle candidates being disadvantaged in ranked-choice voting?

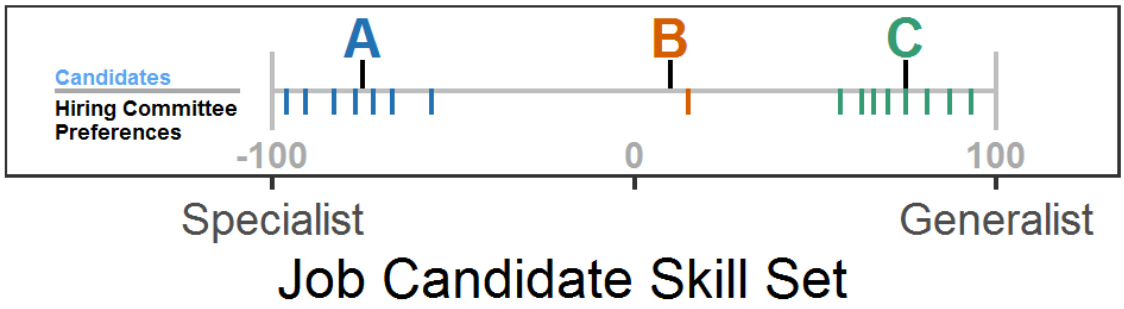
Why is this "center squeeze" example problematic for ranked-choice and plurality voting?
In this "center squeeze" example, there is a fairly good middle-of-the-road candidate, but ranked-choice and plurality voting instead choose extreme candidates.

Think about it this way - imagine that the voters were the same but instead of having three candidates there were only two. If this was just an election with A vs B, B would win. If this was just an election of B vs C, B would win. If there was an election of A vs C, C would win. So in the head-to-head comparisons B wins two of the three elections.

In sum, this example shows that ranked-choice voting can favor a more extreme candidate (C), and plurality can also favor a more extreme candidate (A), as opposed to more middle-of-the-road candidate picked by cardinal (B).

[**Discovery Learning**: Majority Criterion Violation]

On the next few pages you will be walked through a specific voting example called the “Majority Criterion Violation," which demonstrates how in cardinal voting it is possible for a candidate to lose despite having a majority of the first places votes.



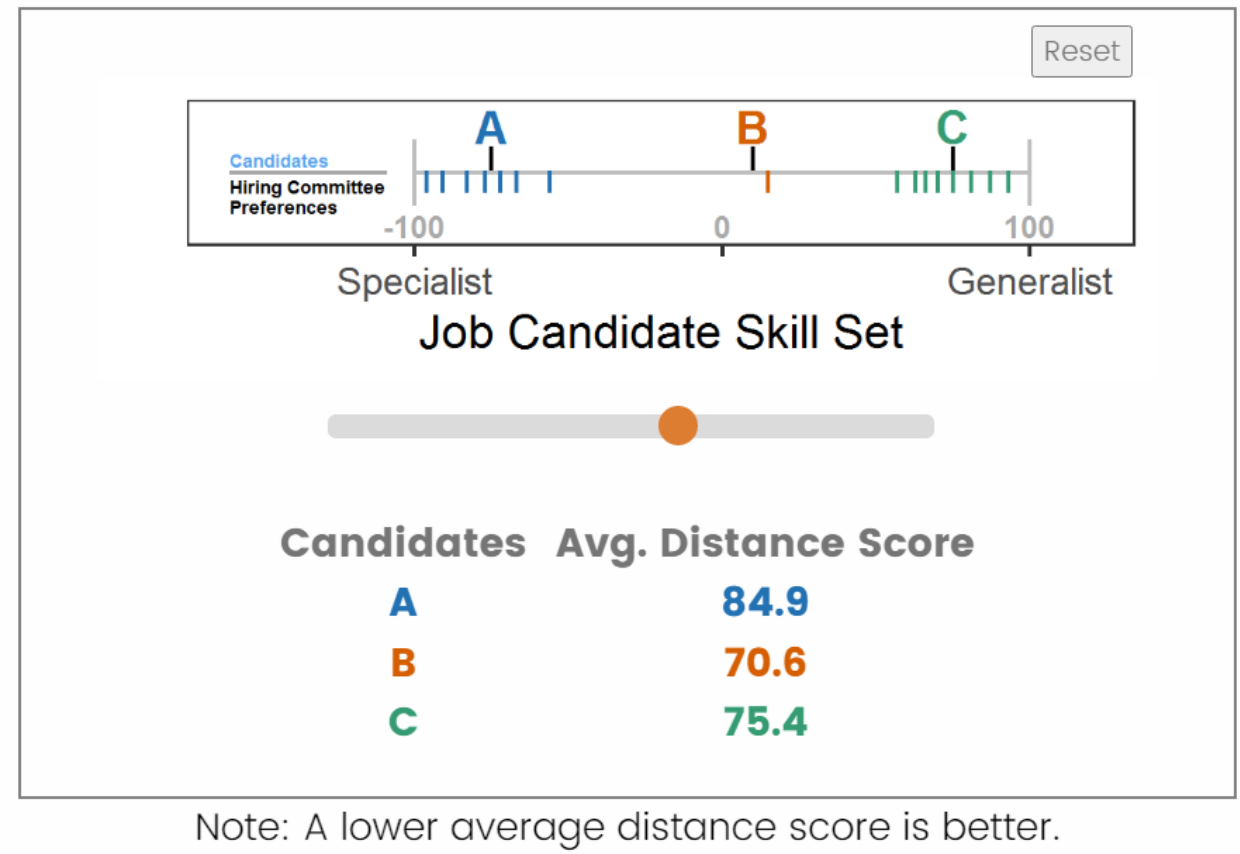
Who wins according to plurality and ranked-choice voting?

- 7 voters have candidate A as their favorite (7/16 voters or 44%), and
- 1 voter has candidate B as their favorite (1/16 voters or 6%), and
- 8 voters have candidate C as their favorite (8/16 voters or 50%). Therefore, C wins according to plurality. C also wins according to ranked-choice voting, because we would eliminate B, since no candidate has over 50% of the total votes, giving C 9 of the 16 votes (56%).

Who wins according to cardinal voting?

- Candidate A is the furthest from all the voters with an average distance of 84.9,
- Candidate B is the closest to the voters with an average distance of 70.6. Therefore, B wins according to cardinal voting.
- Candidate C is the second closest to the voters with an average distance of 75.4.

We created the interactive graph below to help you answer the following questions.



[Note: The above figure is a screenshot of the applet. The participant can manipulate the slider with the orange dot to change the position of the job candidate "B" in the graph. At the far left position the B's average score is 79.8 and at the far right it is 70.3.]

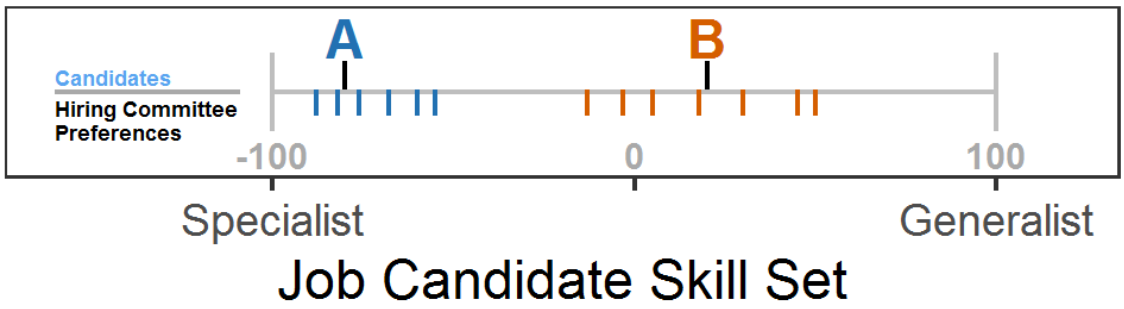
B wins according to cardinal. Do you see any problems with that outcome?

Move around the slider and see how it affects the average distance score for candidate B. What position on the slider is best for candidate B (i.e., has the lowest score) and what does that tell us about cardinal voting?

Why is this example about the 'majority criterion' problematic for carinal voting?
In certain situations cardinal voting can lead to winners that very few voters would choose as their first pick. In the example above, B won according to cardinal, but only one voter was close to B. Additionally, candidate C lost according to cardinal voting but many more voters are close to C than B.

[**Discovery Learning:** Spoiler Effect]

On the next few pages you will be walked through a specific voting example called the "Spoiler Effect," which shows how outcomes in plurality voting can be altered by the addition of a third candidate.



Who wins according to plurality and ranked-choice voting?

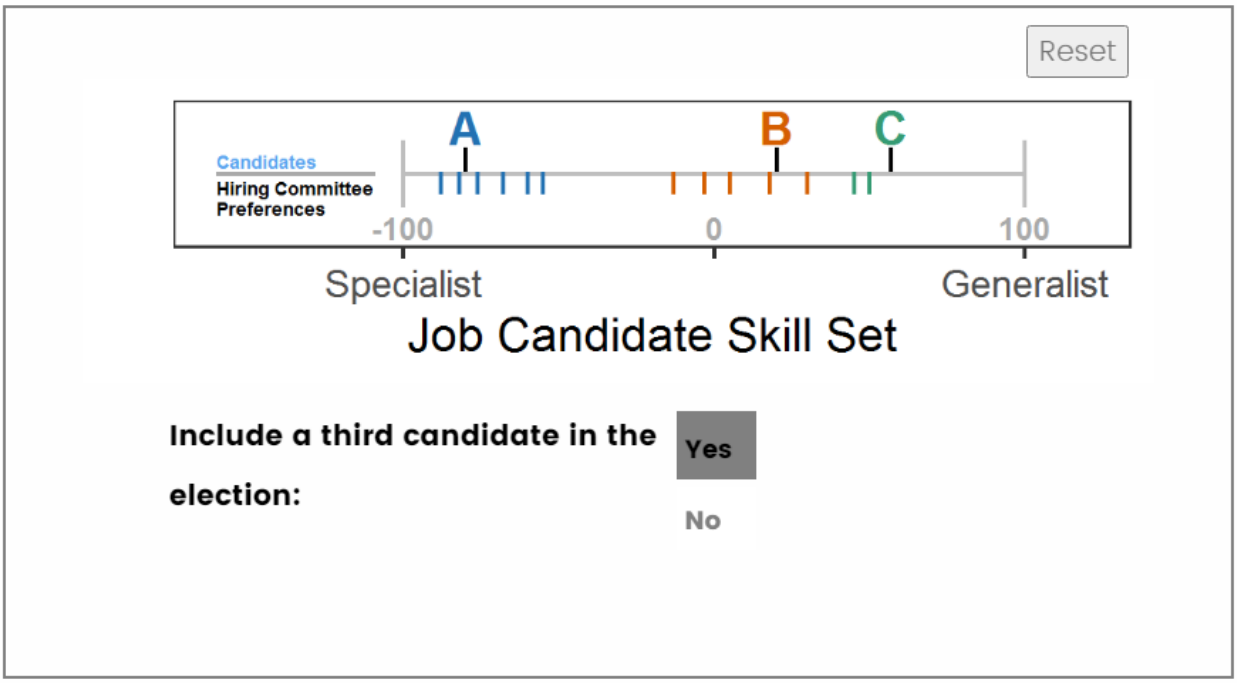
- 6 voters have candidate A as their favorite (6/13 voters or 46%), and
- 7 voters have candidate B as their favorite (7/13 voters or 54%). Therefore, **B wins according to both plurality and ranked-choice voting**, since the candidate has over 50% of the total votes.

Who wins according to cardinal voting?

- Candidate A is the furthest from the voters with an average distance of 58.7,
- Candidate B is closest to the voters with an average distance of 52.8. Therefore, **B wins according to cardinal voting**.

Imagine the same scenario with a slight difference…
Imagine having three candidates instead of two. Specifically, imagine if there was a third candidate who was not very popular, like candidate C in the graph below.

We created the interactive graph below to help you answer the following questions.



[Note: The above figure is a screenshot of the applet. Participants can show/hide the "third candidate" (C), by clicking yes or no. The two right-most hiring committee members switch from supporting candidate C to B, depending on C's presence.]

What is one possible consequence of adding a candidate to an election using a plurality voting system?

Why is this example about having a third candidate problematic for plurality voting?
In plurality voting, the presence of an unpopular third candidate (like Candidate C) can change who wins. This could be viewed as strange given that very few people like Candidate C - so why should C's decision to join the election affect who wins if C has no chance of winning? This does not happen with ranked-choice or cardinal voting.

B8. Voting System Metacognition

Rate how well you think you understand the different voting systems. (0=Not at all; 100=completely)

- Cardinal
- Plurality
- Ranked-choice voting

Appendix C. Study 3 Survey Materials

C1. How to read tables introduction

This study is about how groups of people make collective decisions. We will be using an example of a company hiring a job candidate.

Scenario:
There is a hiring committee of people who work for the company who have different preferences about which candidates would be best to hire. The company has realized that there are different ways that the hiring committee can vote for the candidates. One way is that each member of the hiring committee simply chooses their favorite candidate. We call this the 'Pick One Candidate' voting system. Another way is that each member of the hiring committee ranks each candidate as their first, second, or third favorite. We call this the 'Rank All Candidates' voting system. Another way is that each member of the hiring committee rates each candidate from 100 (best) to 0 (worst). We call this the 'Rate All Candidates' voting system.

The company decides that instead of just using one voting system, the hiring committee use all three voting systems. **Your job is to decide which candidate won based on the results from the three different voting systems.**

Pick One Candidate

Voter	Choice
1	B

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	B	A	C

Rate All Candidates

	Ratings		
Voter	A	B	C
1	64	97	21

The three tables above represent three different types of ballots (or voting methods).

- In the "Pick One Candidate" voter system the voters are asked to pick their favorite candidate.
- In the "Rank All Candidates" voter system the voters rank the candidates by preference (for example, 1st, 2nd, 3rd)
- In the "Rate All Candidates" voter system the voters rate the candidates on a 0-100 scale, where 0 is the lowest or worst rating and 100 is the highest or best rating.

The example above has three job candidates ("A", "B", & "C") and one member on the hiring committee (Voter 1). We will hide the names of job candidates and voters just to make the tables less cluttered.

Note that in this example Voter 1 has chosen Candidate B as the best candidate on all three ballots.

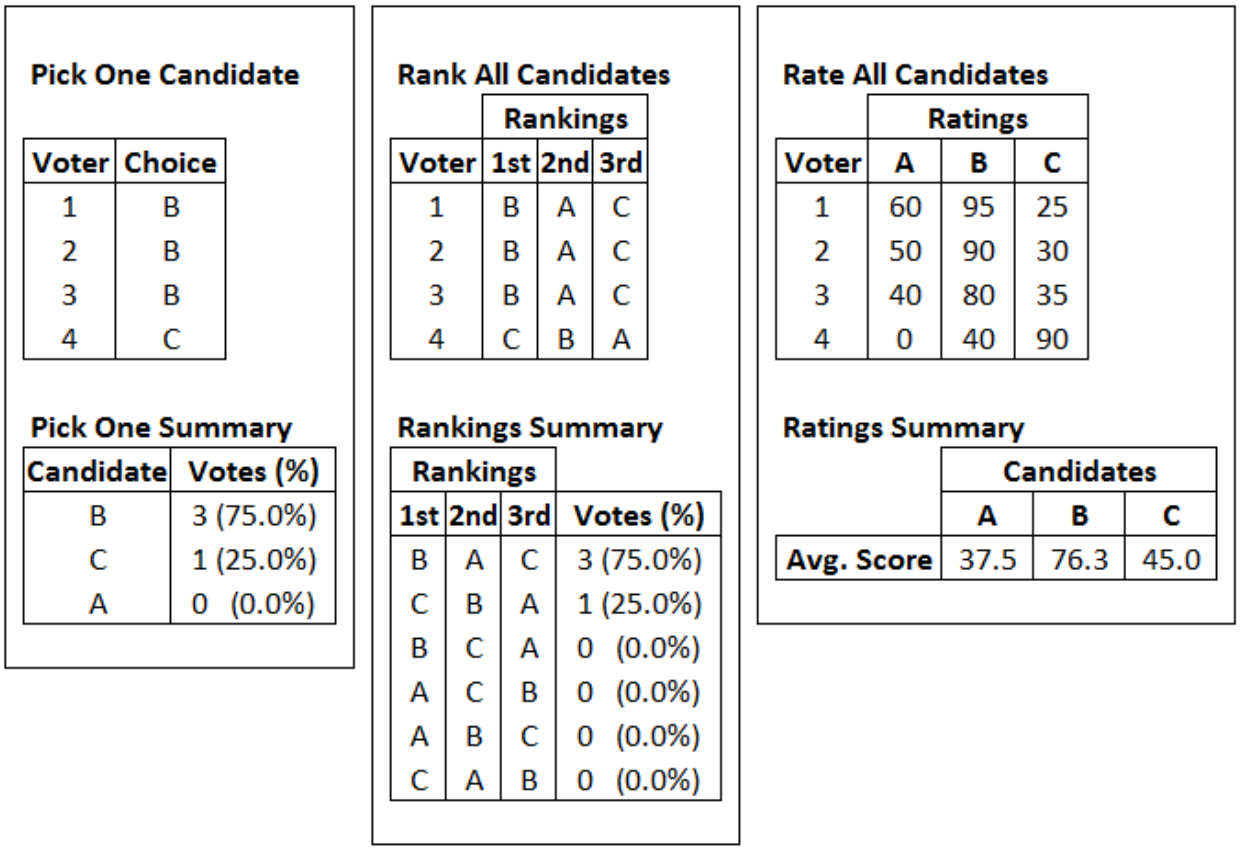
On the next 13 pages you will be presented with 13 different elections and **you need to decide which candidate should be hired**.

C2. Table Comprehension

[Summary and Raw Data Condition]
Here is an example with three job candidates ("A", "B", & "C") and four members on the hiring committee.

Additionally, we added three summary tables which summarize the voting data for each ballot type.

- The "Pick One Summary" table shows that 3 of the hiring committee members voted to hire Candidate B and 1 member voted for Candidate C. No voter chose Candidate A.
- The "Rankings Summary" table shows that 3 out of the 4 voters preferred Candidate B, followed by Candidate A, with Candidate C as their 3rd pick.
- The "Ratings Summary" table shows the average rating score for the three job candidates. Candidate B had the highest average rating with a score of 76.3



Pick One Candidate

Voter	Choice
1	B
2	B
3	B
4	C

Pick One Summary

Candidate	Votes (%)
B	3 (75.0%)
C	1 (25.0%)
A	0 (0.0%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	B	A	C
2	B	A	C
3	B	A	C
4	C	B	A

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
B	A	C	3 (75.0%)
C	B	A	1 (25.0%)
B	C	A	0 (0.0%)
A	C	B	0 (0.0%)
A	B	C	0 (0.0%)
C	A	B	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	A	B	C
1	60	95	25
2	50	90	30
3	40	80	35
4	0	40	90

Ratings Summary

	Candidates		
	A	B	C
Avg. Score	37.5	76.3	45.0

[**Raw Data Only Condition**]
Here is an example with three job candidates ("A", "B", & "C") and four members on the hiring committee. From now on we will hide the names of job candidates just to make the tables less cluttered.

In this example:

- The "Pick One Candidate" table shows that 3 of the hiring committee members voted to

hire Candidate B and 1 member voted for Candidate C. No voter chose Candidate A.

- The "Rank All Candidates" table shows that 3 out of the 4 voters preferred Candidate B, followed by Candidate A, with Candidate C as their 3rd pick.
- The "Rate All Candidates" table shows higher ratings on average for Candidate B, which has an average rating score of 76.3, if you were to calcuate the average.

Pick One Candidate

Voter	Choice
1	B
2	B
3	B
4	C

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	B	A	C
2	B	A	C
3	B	A	C
4	C	B	A

Rate All Candidates

	Ratings		
Voter	A	B	C
1	60	95	25
2	50	90	30
3	40	80	35
4	0	40	90

[**Questions for both conditions below their respective stimuli**]

Q1a. Based on the information above, **rate the candidates based on who should be hired**.
(100=most fair to hire; 0=least fair to hire)
[website had a range slider]
[website allowed single choice)

- Candidate **A**
- Candidate **B** (*Correct Answer*)
- Candidate **C**

Q1b. **Rank order the candidates based on who you think should be hired**.
(Top = first to hire; bottom = last to hire)
[website allowed click-and-drag functionality for rankings]

- Candidate **A**
- Candidate **B** (*Correct Answer*)
- Candidate **C**

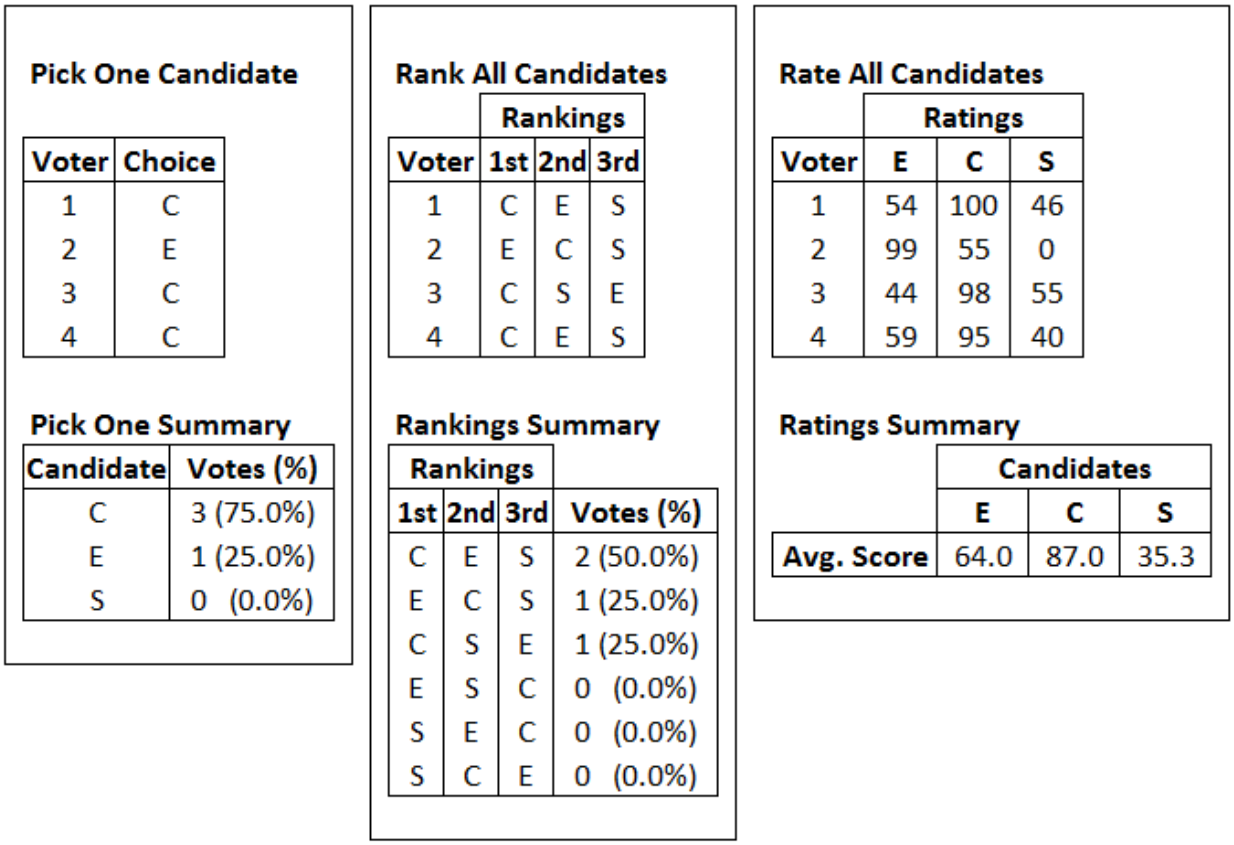
Q1c. **Which job candidate do you think should be hired**?
(choose one)

- Candidate **A**
- Candidate **B** (*Correct Answer*)
- Candidate **C**

Q1d. Which table did you primarily use to make your decisions?

- “Pick One Candidate” (Left Table)
- “Rank All Candidates” (Middle Table)
- “Rate All Candidates” (Right Table)

[Summary and Raw Data Condition]



Pick One Candidate

Voter	Choice
1	C
2	E
3	C
4	C

Pick One Summary

Candidate	Votes (%)
C	3 (75.0%)
E	1 (25.0%)
S	0 (0.0%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	C	E	S
2	E	C	S
3	C	S	E
4	C	E	S

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
C	E	S	2 (50.0%)
E	C	S	1 (25.0%)
C	S	E	1 (25.0%)
E	S	C	0 (0.0%)
S	E	C	0 (0.0%)
S	C	E	0 (0.0%)

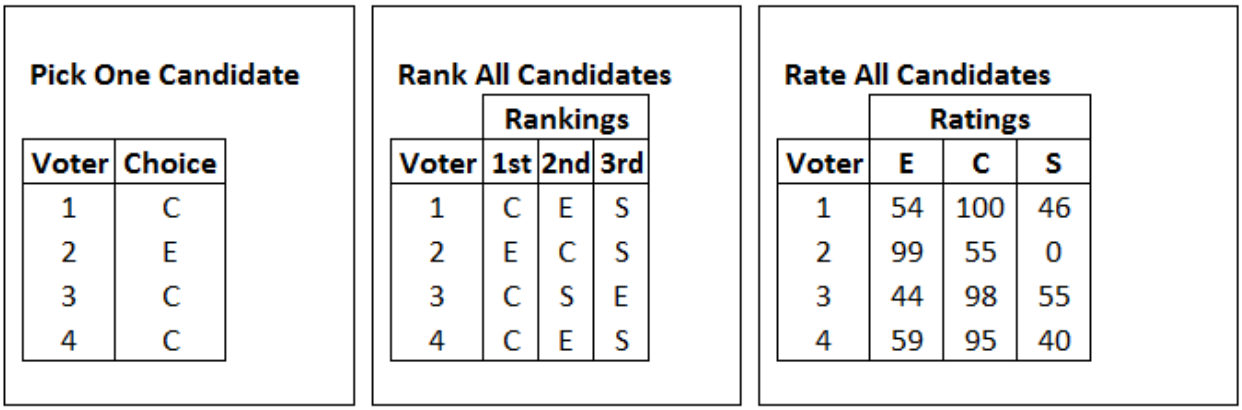
Rate All Candidates

	Ratings		
Voter	E	C	S
1	54	100	46
2	99	55	0
3	44	98	55
4	59	95	40

Ratings Summary

	Candidates		
	E	C	S
Avg. Score	64.0	87.0	35.3

[Raw Data Only Condition]



Pick One Candidate

Voter	Choice
1	C
2	E
3	C
4	C

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	C	E	S
2	E	C	S
3	C	S	E
4	C	E	S

Rate All Candidates

	Ratings		
Voter	E	C	S
1	54	100	46
2	99	55	0
3	44	98	55
4	59	95	40

[**Questions for both conditions below their respective stimuli**]

Q2a. Based on the information above, **rate the candidates based on who should be hired**.
(100=most fair to hire; 0=least fair to hire)
[website had a range slider]
[website allowed single choice)

- Candidate **E**
- Candidate **C** (*Correct Answer*)
- Candidate **S**

Q2b. **Rank order the candidates based on who you think should be hired**.
(Top = first to hire; bottom = last to hire)
[website allowed click-and-drag functionality for rankings]

- Candidate **E**
- Candidate **C** (*Correct Answer*)
- Candidate **S**

Q2c. **Which job candidate do you think should be hired**?
(choose one)

- Candidate **E**
- Candidate **C** (*Correct Answer*)
- Candidate **S**

Q2d. Which table did you primarily use to make your decisions?

- "Pick One Candidate" (Left Table)
- "Rank All Candidates" (Middle Table)
- "Rate All Candidates" (Right Table)

[Summary and Raw Data Condition]

Pick One Candidate

Voter	Choice
1	R
2	J
3	P
4	R
5	R
6	R

Pick One Summary

Candidate	Votes (%)
R	4 (66.7%)
J	1 (16.7%)
P	1 (16.7%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	R	P	J
2	J	R	P
3	P	R	J
4	R	P	J
5	R	P	J
6	R	P	J

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
R	P	J	4 (66.7%)
P	R	J	1 (16.7%)
J	R	P	1 (16.7%)
J	P	R	0 (0.0%)
P	J	R	0 (0.0%)
R	J	P	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	J	R	P
1	32	96	65
2	69	59	27
3	0	80	96
4	35	92	61
5	20	100	76
6	27	100	69

Ratings Summary

	Candidates		
	J	R	P
Avg. Score	30.5	87.8	65.7

[Raw Data Only Condition]

Pick One Candidate

Voter	Choice
1	R
2	J
3	P
4	R
5	R
6	R

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	R	P	J
2	J	R	P
3	P	R	J
4	R	P	J
5	R	P	J
6	R	P	J

Rate All Candidates

	Ratings		
Voter	J	R	P
1	32	96	65
2	69	59	27
3	0	80	96
4	35	92	61
5	20	100	76
6	27	100	69

[**Questions for both conditions below their respective stimuli**]

Q3a. Based on the information above, **rate the candidates based on who should be hired**. (100=most fair to hire; 0=least fair to hire)

[website had a range slider]
[website allowed single choice)

- Candidate **J**
- Candidate **R** (*Correct Answer*)
- Candidate **P**

Q3b. **Rank order the candidates based on who you think should be hired**.
(Top = first to hire; bottom = last to hire)
[website allowed click-and-drag functionality for rankings]

- Candidate **J**
- Candidate **R** (*Correct Answer*)
- Candidate **P**

Q3c. **Which job candidate do you think should be hired**?
(choose one)

- Candidate **J**
- Candidate **R** (*Correct Answer*)
- Candidate **P**

Q3d. Which table did you primarily use to make your decisions?

- "Pick One Candidate" (Left Table)
- "Rank All Candidates" (Middle Table)
- "Rate All Candidates" (Right Table)

C3. Who Won the Election Task

[Note: Participants were randomized to see stimuli with either "Raw+Summary Data" or "Raw Data Only." The major difference between the two conditions is whether or not participants saw summary tables for each ballot type. For simplicity, only the "Raw+Summary Data" stimuli are shown below. See Appendix C2 for side-by-side comparisons.]

For each item participants answered three questions which appeared below each figure.

1. Based on the above graph, **rate the candidates based on who should be hired**.
(100=most fair to hire; 0=least fair to hire)
- Candidate **A**
- Candidate **B**
- Candidate **C**

2. Rank order the candidates based on who you think should be hired.
(Top = first to hire; bottom = last to hire)
- Candidate **A**
- Candidate **B**
- Candidate **C**

3. Which job candidate do you think should be hired?
(choose one)
- Candidate **A**
- Candidate **B**
- Candidate **C**

4. Which table did you primarily use to make your decisions?
- "Pick One Candidate" (Left Table)
- "Rank All Candidates" (Middle Table)
- "Rate All Candidates" (Right Table)

The stimuli items are presented below.

Item 1:

Pick One Candidate

Voter	Choice
1	K
2	S
3	K
4	K
5	A
6	A
7	A
8	K
9	S
10	K
11	A

Pick One Summary

Candidate	Votes (%)
K	5 (45.5%)
A	4 (36.4%)
S	2 (18.2%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	K	A	S
2	S	A	K
3	K	A	S
4	K	A	S
5	A	K	S
6	A	K	S
7	A	S	K
8	K	A	S
9	S	A	K
10	K	A	S
11	A	K	S

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
K	A	S	5 (45.5%)
A	K	S	3 (27.3%)
S	A	K	2 (18.2%)
A	S	K	1 (9.1%)
S	K	A	0 (0.0%)
K	S	A	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	K	A	S
1	95	54	8
2	9	50	96
3	100	60	14
4	87	46	0
5	62	97	51
6	73	87	41
7	46	87	68
8	92	68	22
9	16	57	97
10	91	50	4
11	67	93	47

Ratings Summary

	Candidates		
	K	A	S
Avg. Score	67.1	68.1	40.7

Item 2:

Pick One Candidate

Voter	Choice
1	G
2	C
3	Z
4	C
5	C
6	G
7	G

Pick One Summary

Candidate	Votes (%)
C	3 (42.9%)
G	3 (42.9%)
Z	1 (14.3%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	G	Z	C
2	C	Z	G
3	Z	G	C
4	C	Z	G
5	C	Z	G
6	G	Z	C
7	G	Z	C

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
G	Z	C	3 (42.9%)
C	Z	G	3 (42.9%)
Z	G	C	1 (14.3%)
G	C	Z	0 (0.0%)
Z	C	G	0 (0.0%)
C	G	Z	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	C	Z	G
1	16	63	92
2	90	65	19
3	46	92	63
4	92	46	0
5	95	60	14
6	9	55	100
7	25	71	84

Ratings Summary

	Candidates		
	C	Z	G
Avg. Score	53.3	64.6	53.1

Item 3:

Pick One Candidate

Voter	Choice
1	A
2	F
3	F
4	A
5	A
6	F
7	F

Pick One Summary

Candidate	Votes (%)
F	4 (57.1%)
A	3 (42.9%)
O	0 (0.0%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	A	O	F
2	F	O	A
3	F	O	A
4	A	O	F
5	A	O	F
6	F	O	A
7	F	O	A

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
F	O	A	5 (57.1%)
A	O	F	3 (42.9%)
O	A	F	0 (0.0%)
A	F	O	0 (0.0%)
F	A	O	0 (0.0%)
O	F	A	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	A	O	F
1	88	44	23
2	12	82	98
3	9	79	100
4	91	21	0
5	94	37	16
6	16	86	95
7	19	89	92

Ratings Summary

	Candidates		
	A	O	F
Avg. Score	47.0	62.6	60.6

Item 4:

Pick One Candidate

Voter	Choice
1	L
2	L
3	W
4	N
5	W
6	N
7	W
8	N
9	N

Pick One Summary

Candidate	Votes (%)
N	4 (44.4%)
W	3 (33.3%)
L	2 (22.2%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	L	W	N
2	L	W	N
3	W	L	N
4	N	L	W
5	W	L	N
6	N	L	W
7	W	L	N
8	N	L	W
9	N	L	W

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
N	L	W	4 (44.4%)
W	L	N	3 (33.3%)
L	W	N	2 (22.2%)
L	N	W	0 (0.0%)
N	W	L	0 (0.0%)
W	N	L	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	N	L	W
1	44	94	69
2	53	84	60
3	3	66	90
4	91	47	23
5	0	63	87
6	91	28	4
7	21	83	93
8	94	44	19
9	100	38	13

Ratings Summary

	Candidates		
	N	L	W
Avg. Score	55.2	60.8	50.9

Item 5:

Pick One Candidate

Voter	Choice
1	I
2	Q
3	I
4	I
5	Z
6	Q
7	Z

Pick One Summary

Candidate	Votes (%)
I	3 (42.9%)
Q	2 (28.6%)
Z	2 (28.6%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	I	Q	Z
2	Q	I	Z
3	I	Q	Z
4	I	Q	Z
5	Z	Q	I
6	Q	I	Z
7	Z	Q	I

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
I	Q	Z	3 (42.9%)
Q	I	Z	2 (28.6%
Z	Q	I	2 (28.6%
I	Z	Q	0 (0.0%)
Z	I	Q	0 (0.0%)
Q	Z	I	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	I	Q	Z
1	81	75	0
2	82	88	48
3	84	79	3
4	86	80	5
5	41	47	89
6	80	86	50
7	30	36	100

Ratings Summary

	Candidates		
	I	Q	Z
Avg. Score	69.1	70.1	42.1

Item 6:

Pick One Candidate

Voter	Choice
1	J
2	J
3	S
4	J
5	J
6	J
7	U
8	U
9	S
10	U
11	J
12	U
13	U
14	S

Pick One Summary

Candidate	Votes (%)
J	6 (42.9%)
U	5 (35.7%)
S	3 (21.4%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	J	S	U
2	J	S	U
3	S	U	J
4	J	S	U
5	J	S	U
6	J	S	U
7	U	S	J
8	U	S	J
9	S	U	J
10	U	S	J
11	J	S	U
12	U	S	J
13	U	S	J
14	S	U	J

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
J	S	U	6 (42.9%)
U	S	J	5 (35.7%)
S	U	J	3 (21.4%)
U	J	S	0 (0.0%)
J	U	S	0 (0.0%)
S	J	U	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	J	S	U
1	95	45	4
2	84	67	26
3	43	93	67
4	94	58	17
5	100	51	10
6	97	55	14
7	8	58	99
8	18	68	92
9	54	97	56
10	22	72	88
11	97	47	5
12	11	61	99
13	0	50	91
14	45	95	65

Ratings Summary

	Candidates		
	J	S	U
Avg. Score	54.9	65.5	52.4

Item 7:

Pick One Candidate

Voter	Choice
1	K
2	T
3	K
4	T
5	T
6	T
7	K
8	K
9	K

Pick One Summary

Candidate	Votes (%)
K	5 (55.6%)
T	4 (44.4%)

Rank All Candidates

	Rankings	
Voter	1st	2nd
1	K	T
2	T	K
3	K	T
4	T	K
5	T	K
6	T	K
7	K	T
8	K	T
9	K	T

Rankings Summary

Rankings		
1st	2nd	Votes (%)
K	T	5 (55.6%)
T	K	4 (44.4%)

Rate All Candidates

	Ratings	
Voter	T	K
1	51	53
2	95	8
3	29	75
4	99	4
5	98	0
6	90	14
7	41	63
8	3	100
9	18	85

Ratings Summary

	Candidates	
	T	K
Avg. Score	58.2	44.7

Item 8:

Pick One Candidate

Voter	Choice
1	M
2	L
3	M
4	U
5	M
6	U
7	L
8	L
9	L
10	L
11	U
12	M
13	U
14	L
15	U
16	U
17	U
18	M

Pick One Summary

Candidate	Votes (%)
U	7 (38.9%)
L	6 (33.3%)
M	5 (27.8%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	M	U	L
2	L	M	U
3	M	L	U
4	U	M	L
5	M	L	U
6	U	M	L
7	L	M	U
8	L	M	U
9	L	M	U
10	L	M	U
11	U	M	L
12	M	L	U
13	U	M	L
14	L	M	U
15	U	M	L
16	U	M	L
17	U	M	L
18	M	L	U

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
U	M	L	7 (38.9%)
L	M	U	6 (33.3%)
M	L	U	4 (22.2%)
M	U	L	1 (5.6%)
U	L	M	0 (0.0%)
L	U	M	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	L	M	U
1	60	92	61
2	94	74	26
3	66	99	55
4	6	39	87
5	71	97	49
6	16	48	96
7	90	79	31
8	97	71	24
9	96	63	15
10	99	67	19
11	32	64	89
12	74	94	47
13	10	42	90
14	87	82	34
15	25	58	96
16	0	33	80
17	21	53	100
18	62	95	58

Ratings Summary

	Candidates		
	L	M	U
Avg. Score	55.9	69.4	58.7

Item 9:

Pick One Candidate

Voter	Choice
1	Z
2	X
3	Z
4	Z
5	Z
6	C
7	X
8	X
9	X
10	C
11	X
12	Z

Pick One Summary

Candidate	Votes (%)
X	5 (41.7%)
Z	5 (41.7%)
C	2 (16.7%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	Z	C	X
2	X	C	Z
3	Z	C	X
4	Z	C	X
5	Z	C	X
6	C	X	Z
7	X	C	Z
8	X	C	Z
9	X	C	Z
10	C	X	Z
11	X	C	Z
12	Z	C	X

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
Z	C	X	5 (41.7%)
X	C	Z	5 (41.7%)
C	X	Z	2 (16.7%)
X	Z	C	0 (0.0%)
Z	X	C	0 (0.0%)
C	Z	X	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	Z	C	X
1	92	60	16
2	6	56	100
3	90	63	19
4	94	44	0
5	99	49	5
6	47	97	61
7	11	60	98
8	2	52	96
9	20	70	89
10	43	92	66
11	18	67	91
12	96	57	13

Ratings Summary

	Candidates		
	Z	C	X
Avg. Score	51.5	63.9	54.5

Item 10:

Pick One Candidate

Voter	Choice
1	J
2	N
3	J
4	J
5	N
6	Q
7	N
8	J
9	J
10	Q
11	N

Pick One Summary

Candidate	Votes (%)
J	5 (45.5%)
N	4 (36.4%)
Q	2 (18.2%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	J	Q	N
2	N	Q	J
3	J	Q	N
4	J	Q	N
5	N	Q	J
6	Q	N	J
7	N	Q	J
8	J	Q	N
9	J	Q	N
10	Q	N	J
11	N	Q	J

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
J	Q	N	5 (45.5%)
N	Q	J	4 (36.4%)
Q	N	J	2 (18.2%)
Q	J	N	0 (0.0%)
J	N	Q	0 (0.0%)
N	J	Q	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	N	Q	I
1	7	16	50
2	55	47	12
3	15	23	58
4	24	33	67
5	47	39	4
6	85	93	76
7	60	52	17
8	0	9	43
9	35	43	78
10	91	100	69
11	68	60	25

Ratings Summary

	Candidates		
	N	Q	I
Avg. Score	44.3	46.8	45.4

C4. Voting System Introduction

[Note: Participants were formally introduced to cardinal, plurality, and ranked-choice voting in this block. The voting system order was randomized for each participant.]

After each voting system was introduced the participant was asked:

> Q. How fair is [cardinal/plurality/RCV] voting? Rate from very fair (100) to not very fair (0). [slider question format]

The stimuli follows:

The next three pages will teach you about three different ways to determine who should win when multiple people are voting.

[Raw+Summary Data Condition]

[**Cardinal**]
One method to determine election outcomes is to have voters rate each candidate. This is called **cardinal voting**. In cardinal voting the candidate that has the highest average rating wins. In the information presented below we show the ratings of four voters for three candidates.

As you can see, C has the lowest voter rating on average [math: (100+67+0+25)/4=48.0]. A and B both have higher voter ratings on average. However, B has a better voter rating on average [math: (67+100+50+75)/4=73.0] than A [math: (17+50+100+92)/4=64.8]. Thus, according to cardinal voting, **candidate B would win**.

Rate All Candidates

	Ratings		
Voter	**A**	**B**	**C**
1	17	67	100
2	50	100	67
3	100	50	0
4	92	75	25

Ratings Summary

	Candidates		
	A	**B**	**C**
Avg. Score	64.8	73.0	48.0

[**Plurality**]
One method to determine election outcomes is to allow all voters to choose only one candidate and the candidate with the most votes wins. This is called **plurality voting**. Plurality voting

prioritizes selecting a winner who has the single most support. A winning candidate in a plurality system needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.

For example, in the information presented below "Candidate A" has 4 out of 9 hiring committee members support. "Candidate B" has 2 out of the 9 hiring committee members support. "Candidate C" has 3 out of the 9 hiring committee members support. Using a Plurality voting system, Candidate A would win, even though they did not receive more than 50% of the votes.

Pick One Candidate

Voter	Choice
1	B
2	A
3	A
4	A
5	C
6	C
7	B
8	A
9	C

Pick One Summary

Candidate	Votes (%)
A	4 (44.4%)
C	3 (33.3%)
B	2 (22.2%)

[**Ranked-Choice Voting**]
One method to determine election outcomes is to focus on how each voter ranks the candidates (e.g., first, second, third preference). This is called **ranked-choice voting**. Ranked-choice voting prioritizes majority approval and a winning candidate must have over 50% of the total votes. This works by first counting all voters' first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support. If so, this candidate wins. If not, the process is repeated.

For example, in the information presented below, there are four voters who rank Candidate A as their first preference. There are three voters who rank C as their first preference and there are two voters who rank B as their first preference.

Although A has the most first place votes (4 out of 9), A does not have over 50% of the votes. Because B has the fewest first place votes, B would be removed as a candidate, and B's votes would be reassigned. In this case, the two voters who most preferred B would be reassigned to C, because this is their 2nd choice. Once B is removed, there are 5 votes for C, and 4 votes for A, so Candidate C would win.

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	C	B	A
2	B	C	A
3	C	B	A
4	A	B	C
5	B	C	A
6	A	B	C
7	C	B	A
8	A	B	C
9	A	B	C

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
A	B	C	4 (44.4%)
C	B	A	3 (33.3%)
B	C	A	2 (22.2%)
A	C	B	0 (0.0%)
B	A	C	0 (0.0%)
C	A	B	0 (0.0%)

[Raw Data Only Condition]

[**Cardinal**]
One method to determine election outcomes is to have voters rate each candidate. This is called **cardinal voting**. In cardinal voting the candidate that has the highest average rating wins. In the information presented below we show the ratings of four voters for three candidates.

As you can see, C has the lowest voter rating on average [math: (100+67+0+25)/4=48.0]. A and B both have higher voter ratings on average. However, B has a better voter rating on average [math: (67+100+50+75)/4=73.0] than A [math: (17+50+100+92)/4=64.8]. Thus, according to cardinal voting, **candidate B would win**.

Rate All Candidates

	Ratings		
Voter	A	B	C
1	17	67	100
2	50	100	67
3	100	50	0
4	92	75	25

[**Plurality**]
One method to determine election outcomes is to allow all voters to choose only one candidate and the candidate with the most votes wins. This is called **plurality voting**. Plurality voting

prioritizes selecting a winner who has the single most support. A winning candidate in a plurality system needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.

For example, in the information presented below “Candidate A” has 4 out of 9 hiring committee members support. “Candidate B” has 2 out of the 9 hiring committee members support. “Candidate C” has 3 out of the 9 hiring committee members support. Using a Plurality voting system, Candidate A would win, even though they did not receive more than 50% of the votes.

Pick One Candidate

Voter	Choice
1	B
2	A
3	A
4	A
5	C
6	C
7	B
8	A
9	C

[**Ranked-Choice Voting**]
One method to determine election outcomes is to focus on how each voter ranks the candidates (e.g., first, second, third preference). This is called **ranked-choice voting**. Ranked-choice voting prioritizes majority approval and a winning candidate must have over 50% of the total votes. This works by first counting all voters' first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is “eliminated” and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support. If so, this candidate wins. If not, the process is repeated.

For example, in the information presented below, there are four voters who rank Candidate A as their first preference. There are three voters who rank C as their first preference and there are two voters who rank B as their first preference.

Although A has the most first place votes (4 out of 9), A does not have over 50% of the votes. Because B has the fewest first place votes, B would be removed as a candidate, and B’s votes would be reassigned. In this case, the two voters who most preferred B would be reassigned to C, because this is their 2nd choice. Once B is removed, there are 5 votes for C, and 4 votes for A, so Candidate C would win.

Rank All Candidates

	Rankings		
Voter	**1st**	**2nd**	**3rd**
1	C	B	A
2	B	C	A
3	C	B	A
4	A	B	C
5	B	C	A
6	A	B	C
7	C	B	A
8	A	B	C
9	A	B	C

C5. Declared voting system preferences

Q1. **Which of the three voting systems below do you think is most fair?**
[see descriptions below if you need a reminder]

- Cardinal / “Rate Candidates”
- Plurality / “Pick One”
- Ranked-choice voting / “Rank Candidates”

Q2. Order the voting systems below (drag and drop list items), from most fair (top) to least fair (bottom).

- Cardinal / “Rate Candidates”
- Plurality / “Pick One”
- Ranked-choice voting / “Rank Candidates”

Q3. Rate the voting systems below, from very fair (100) to not very fair (0).

- Cardinal / “Rate Candidates”
- Plurality / “Pick One”
- Ranked-choice voting / “Rank Candidates”

Remember:

- *In **cardinal** voting a winning candidate is the candidate with the best average score. This involves voters rating all candidates and calculating the average (mean) rating.*

- *In **ranked-choice** voting a winning candidate must have over 50% of the total votes. This works by first counting all voters first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support.*

- *In **plurality** voting a winning candidate needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.*

Q4. Choose the voting system you would prefer to use.
[see descriptions below if you need a reminder]

- Cardinal / “Rate Candidates”
- Plurality / “Pick One”
- Ranked-choice voting / “Rank Candidates”

Q5. Rank order the voting systems by how much you would prefer to use them.
Drag and drop them from most like to use for voting (top) to least like to use for voting (bottom).

- Cardinal / “Rate Candidates”
- Plurality / “Pick One”
- Ranked-choice voting / “Rank Candidates”

Q6. Rate each voting system by how much you would like to use them. (0=Do not like at all; 100=Like a lot)

- Cardinal / “Rate Candidates”
- Plurality / “Pick One”
- Ranked-choice voting / “Rank Candidates”

C6. Voting system comprehension

[Note: Participants were randomized to see stimuli with either "Raw+Summary Data" or "Raw Data Only." The major difference between the two conditions is whether or not participants saw summary tables for each ballot type. For simplicity, only the "Raw+Summary Data" stimuli are shown below. See Appendix C2 for side-by-side comparisons.]

For each item participants answered three questions which appeared below each figure.

Based on the information above... which candidate would be chosen as the winner in a **cardinal voting system**?

- Candidate **A**
- Candidate **B**
- Candidate **C**

The next few pages will ask you to identify which job candidate should be hired based on the different ways to determine elections that you just learned about.

Learning item 1a:
This question is about **cardinal voting**.

Rate All Candidates

	Ratings		
Voter	**R**	**T**	**M**
1	98	42	16
2	45	95	69
3	34	97	80
4	98	36	10
5	3	65	91
6	88	26	0
7	15	77	100
8	21	84	93
9	91	49	23

Ratings Summary

	Candidates		
	R	**T**	**M**
Avg. Score	54.8	63.4	53.6

Learning item 1b:
This question is about **plurality voting**.

Pick One Candidate

Voter	Choice
1	R
2	T
3	T
4	R
5	M
6	R
7	M
8	M
9	R

Pick One Summary

Candidate	Votes (%)
R	4 (44.4%)
M	3 (33.3%)
T	2 (22.2%)

Learning item 1c:
This question is about **ranked-choice voting**.

Pick One Candidate

Voter	Choice
1	R
2	T
3	T
4	R
5	M
6	R
7	M
8	M
9	R

Pick One Summary

Candidate	Votes (%)
R	4 (44.4%)
M	3 (33.3%)
T	2 (22.2%)

Learning item 2a:
This question is about **cardinal voting**.

Rate All Candidates

	Ratings		
Voter	**Y**	**L**	**U**
1	93	69	22
2	63	99	53
3	5	45	92
4	34	74	81
5	90	50	3
6	10	50	97
7	21	61	94
8	55	95	60
9	83	79	32
10	100	60	14
11	0	40	86

Ratings Summary

	Candidates		
	Y	**L**	**U**
Avg. Score	50.4	65.6	57.6

Learning item 2b:
This question is about **plurality voting**.

Pick One Candidate

Voter	Choice
1	Y
2	L
3	U
4	U
5	Y
6	U
7	U
8	L
9	Y
10	Y
11	U

Pick One Summary

Candidate	Votes (%)
U	5 (45.5%)
Y	4 (36.4%)
L	2 (18.2%)

Learning item 2c:
This question is about **ranked-choice voting**.

Rank All Candidates

	Rankings		
Voter	**1st**	**2nd**	**3rd**
1	Y	L	U
2	L	Y	U
3	U	L	Y
4	U	L	Y
5	Y	L	U
6	U	L	Y
7	U	L	Y
8	L	U	Y
9	Y	L	U
10	Y	L	U
11	U	L	Y

Rankings Summary

Rankings			
1st	**2nd**	**3rd**	**Votes (%)**
U	L	Y	5 (45.5%)
Y	L	U	4 (36.4%)
L	U	Y	1 (9.1%)
L	Y	U	1 (9.1%)
Y	U	L	0 (0.0%)
U	Y	L	0 (0.0%)

Learning item 3a:
This question is about **cardinal voting**.

Rate All Candidates

	Ratings		
Voter	**Q**	**K**	**J**
1	26	61	88
2	71	94	43
3	9	45	95
4	88	77	26
5	100	64	14
6	14	50	99
7	63	98	51
8	0	36	86
9	58	94	55
10	92	72	21
11	96	60	10
12	19	55	95

Ratings Summary

	Candidates		
	Q	**K**	**J**
Avg. Score	53.0	67.2	56.9

Learning item 3b:
This question is about **plurality voting**.

Pick One Candidate

Voter	Choice
1	J
2	K
3	J
4	Q
5	Q
6	J
7	K
8	J
9	K
10	Q
11	Q
12	J

Pick One Summary

Candidate	Votes (%)
J	5 (41.7%)
Q	4 (33.3%)
K	3 (25.0%)

Learning item 3c:
This question is about **ranked-choice voting**.

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	J	K	Q
2	K	Q	J
3	J	K	Q
4	Q	K	J
5	Q	K	J
6	J	K	Q
7	K	Q	J
8	J	K	Q
9	K	Q	J
10	Q	K	J
11	Q	K	J
12	J	K	Q

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
J	K	Q	5 (41.7%)
Q	K	J	4 (33.3%)
K	Q	J	3 (25.0%)
K	J	Q	0 (0.0%)
J	Q	K	0 (0.0%)
Q	J	K	0 (0.0%)

Learning item 4a:
This question is about **cardinal voting**.

Rate All Candidates

	Ratings		
Voter	**T**	**X**	**P**
1	89	58	8
2	28	59	92
3	11	43	93
4	97	73	23
5	93	77	27
6	75	94	44
7	61	92	59
8	97	65	15
9	67	99	52
10	22	54	98
11	24	55	96
12	0	32	82
13	14	46	96
14	89	80	30
15	100	69	19
16	71	98	48
17	63	95	57
18	7	38	88

Ratings Summary

	Candidates		
	T	**X**	**P**
Avg. Score	56.0	68.2	57.1

Learning item 4b:
This question is about **plurality voting**.

Pick One Candidate

Voter	Choice
1	T
2	P
3	P
4	T
5	T
6	X
7	X
8	T
9	X
10	P
11	P
12	P
13	P
14	T
15	T
16	X
17	X
18	P

Pick One Summary

Candidate	Votes (%)
P	7 (38.9%)
T	6 (33.3%)
X	5 (27.8%)

Learning item 4c:
This question is about **ranked-choice voting**.

Rank All Candidates

	Rankings		
Voter	**1st**	**2nd**	**3rd**
1	T	X	P
2	P	X	T
3	P	X	T
4	T	X	P
5	T	X	P
6	X	T	P
7	X	T	P
8	T	X	P
9	X	T	P
10	P	X	T
11	P	X	T
12	P	X	T
13	P	X	T
14	T	X	P
15	T	X	P
16	X	T	P
17	X	T	P
18	P	X	T

Rankings Summary

Rankings			
1st	**2nd**	**3rd**	**Votes (%)**
P	X	T	7 (38.9%)
T	X	P	6 (33.3%)
X	T	P	5 (27.8%)
P	T	X	0 (0.0%)
T	P	X	0 (0.0%)
X	P	T	0 (0.0%)

C7. Teaching Intervention

[Note: Participants were randomized to see stimuli with either "Raw+Summary Data" or "Raw Data Only." The major difference between the two conditions is whether or not participants saw summary tables for each ballot type. For simplicity, only the "Raw+Summary Data" stimuli are shown below. See Appendix C2 for side-by-side comparisons. Voting system quirks used a counterbalanced presentation order across participants.]

[Center Squeeze Phenomenon]
On the next page you will be walked through a specific voting example called the "Center Squeeze Phenomenon," which shows how ranked-choice voting can hurt moderate or "middle-of-the-road" candidates.

Pick One Candidate

Voter	Choice
1	A
2	C
3	C
4	C
5	C
6	B
7	B
8	A
9	A
10	A
11	B
12	B
13	A
14	C
15	A
16	B
17	A
18	A
19	B
20	C
21	C

Pick One Summary

Candidate	Votes (%)
A	8 (38.1%)
C	7 (33.3%)
B	6 (28.6%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	A	B	C
2	C	B	A
3	C	B	A
4	C	B	A
5	C	B	A
6	B	C	A
7	B	C	A
8	A	B	C
9	A	B	C
10	A	B	C
11	B	A	C
12	B	C	A
13	A	B	C
14	C	B	A
15	A	B	C
16	B	C	A
17	A	B	C
18	A	B	C
19	B	C	A
20	C	B	A
21	C	B	A

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
A	B	C	8 (38.1%)
C	B	A	7 (33.3%)
B	C	A	5 (23.8%)
B	A	C	1 (4.8%)
C	A	B	0 (0.0%)
A	C	B	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	A	B	C
1	85	36	6
2	18	67	97
3	0	49	79
4	9	58	88
5	24	73	97
6	55	96	66
7	46	94	76
8	94	57	27
9	98	53	24
10	97	48	18
11	62	89	59
12	53	98	68
13	99	50	20
14	12	61	91
15	88	39	9
16	58	93	63
17	93	44	14
18	91	42	12
19	48	97	73
20	21	70	100
21	3	52	82

Ratings Summary

	Candidates		
	A	B	C
Avg. Score	55.0	65.0	55.7

Who wins according to plurality and ranked-choice voting?

- 8 voters have candidate A as their favorite (8/21 voters or 38%). Therefore, **A wins according to plurality**.
- 6 voters have candidate B as their favorite (6/21 voters or 29%), and
- 7 voters have candidate C as their favorite (7/21 voters or 33%). **C wins according to ranked-choice voting**, because we would eliminate B, since no candidate has over 50%, giving C 12 of the 21 votes (57%).

Who wins according to cardinal voting?

- Candidate A has the lowest average voter rating with a score of 55.0,
- Candidate B has the highest average voter rating with a score of 65.0. Therefore, **B wins according to cardinal voting**.
- Candidate C has the second highest average voter rating with a score of 55.7.

Why is this "center squeeze" example problematic for ranked-choice and plurality voting?
In this "center squeeze" example, there is a fairly good middle-of-the-road candidate, but ranked-choice and plurality voting instead choose extreme candidates.

Think about it this way - imagine that the voters were the same but instead of having three candidates there were only two. If this was just an election with A vs B, B would win. If this was just an election of B vs C, B would win. If there was an election of A vs C, C would win. So in the head-to-head comparisons B wins two of the three elections.

In sum, this example shows that ranked-choice voting can favor a more extreme candidate (C), and plurality can also favor a more extreme candidate (A), as opposed to more middle-of-the-road candidate picked by cardinal (B).

[Majority Criterion Violation]
On the next page you will be walked through a specific voting example called the “Majority Criterion Violation," which demonstrates how in cardinal voting it is possible for a candidate to lose despite having a majority of the first places votes.

Pick One Candidate

Voter	Choice
1	B
2	A
3	A
4	C
5	A
6	A
7	C
8	C
9	C
10	A
11	A
12	C
13	C
14	A
15	C
16	C

Pick One Summary

Candidate	Votes (%)
C	8 (50.0%)
A	7 (43.8%)
B	1 (6.2%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	B	C	A
2	A	B	C
3	A	B	C
4	C	B	A
5	A	B	C
6	A	B	C
7	C	B	A
8	C	B	A
9	C	B	A
10	A	B	C
11	A	B	C
12	C	B	A
13	C	B	A
14	A	B	C
15	C	B	A
16	C	B	A

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
C	B	A	8 (50.0%)
A	B	C	7 (43.8%)
B	C	A	1 (6.2%)
B	A	C	0 (0.0%)
A	C	B	0 (0.0%)
C	A	B	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	A	B	C
1	47	97	65
2	88	32	0
3	91	35	3
4	12	68	100
5	89	56	23
6	99	43	11
7	2	57	90
8	5	61	93
9	15	71	97
10	98	46	14
11	95	40	8
12	19	75	93
13	23	78	90
14	95	49	17
15	18	73	95
16	9	64	96

Ratings Summary

	Candidates		
	A	B	C
Avg. Score	50.3	59.1	55.9

Who wins according to plurality and ranked-choice voting?

- 7 voters have candidate A as their favorite (7/16 voters or 44%), and
- 1 voter has candidate B as their favorite (1/16 voters or 6%), and
- 8 voters have candidate C as their favorite (8/16 voters or 50%). Therefore, **C wins according to plurality**. **C also wins according to ranked-choice voting**, because we would eliminate B, since no candidate has over 50% of the total votes, giving C 9 of the 16 votes (56%).

Who wins according to cardinal voting?

- Candidate A has the lowest average voter rating with a score of 50.3,
- Candidate B has the highest average voter rating with a score of 59.1. Therefore, **B wins according to cardinal voting**.
- Candidate C has the second highest average voter rating with a score of 55.9.

Why is this example about the "majority criterion" problematic for cardinal voting?

In certain situations cardinal voting can lead to winners that very few voters would choose as their first pick. In the example above, B won according to cardinal, but only one voter preferred B over the two other candidates. Additionally, candidate C lost according to cardinal voting but many more voters had C as their first pick compared to B.

[Spoiler Effect]
On the next page you will be walked through a specific voting example called the “Spoiler Effect," which shows how outcomes in plurality voting can be altered by the addition of a third candidate.

Pick One Candidate

Voter	Choice
1	B
2	A
3	A
4	A
5	B
6	A
7	A
8	C
9	B
10	B
11	B
12	C
13	A

Pick One Summary

Candidate	Votes (%)
B	7 (53.8%)
A	6 (46.2%)

Rank All Candidates

	Rankings	
Voter	1st	2nd
1	B	A
2	A	B
3	A	B
4	A	B
5	B	A
6	A	B
7	A	B
8	B	A
9	B	A
10	B	A
11	B	A
12	B	A
13	A	B

Rankings Summary

Rankings		
1st	2nd	Votes (%)
B	A	7 (53.8%)
A	B	6 (46.2%)

Rate All Candidates

	Ratings	
Voter	A	B
1	24	94
2	99	34
3	87	46
4	100	30
5	54	78
6	93	40
7	84	49
8	10	80
9	48	85
10	33	100
11	42	91
12	14	84
13	96	26

Ratings Summary

	Candidates	
	A	B
Avg. Score	60.3	64.4

Who wins according to plurality and ranked-choice voting?

- 6 voters have candidate A as their favorite (6/13 voters or 46%), and
- 7 voters have candidate B as their favorite (7/13 voters or 54%). Therefore, **B wins according to both plurality and ranked-choice voting**, since the candidate has over 50% of the total votes.

Who wins according to cardinal voting?

- Candidate A has the lowest average voting rating with a score of 60.3.
- Candidate B has the highest average voter rating with a score of 64.4. Therefore, **B wins according to cardinal voting**.

Imagine the same scenario with a slight difference…
Imagine having three candidates instead of two. Specifically, imagine if there was a third candidate who was not very popular, like candidate C in the information presented below.

Pick One Candidate

Voter	Choice
1	B
2	A
3	A
4	A
5	B
6	A
7	A
8	C
9	B
10	B
11	B
12	C
13	A

Pick One Summary

Candidate	Votes (%)
A	6 (46.2%)
B	5 (38.5%)
C	2 (15.4%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	B	C	A
2	A	B	C
3	A	B	C
4	A	B	C
5	B	A	C
6	A	B	C
7	A	B	C
8	C	B	A
9	B	C	A
10	B	C	A
11	B	C	A
12	C	B	A
13	A	B	C

Rankings Summary

Rankings			
1st	2nd	3rd	Votes (%)
A	B	C	6 (46.2%)
B	C	A	4 (30.8%)
C	B	A	2 (15.4%)
B	A	C	1 (7.7%)
C	A	B	0 (0.0%)
A	C	B	0 (0.0%)

Rate All Candidates

	Ratings		
Voter	A	B	C
1	24	94	82
2	99	34	8
3	87	46	20
4	100	30	4
5	54	78	52
6	93	40	14
7	84	49	23
8	10	80	96
9	48	85	59
10	33	100	74
11	42	91	65
12	14	84	93
13	96	26	0

Ratings Summary

	Candidates		
	A	B	C
Avg. Score	60.3	64.4	45.4

Why is this example about having a third candidate problematic for plurality voting?
In plurality voting, the presence of an unpopular third candidate (like Candidate C) can change who wins. This could be viewed as strange given that very few people like Candidate C - so why should C's decision to join the election affect who wins if C has no chance of winning? This does not happen with ranked-choice or cardinal voting.

C8. Probing questions about voting system preferences

Q1. Prior to participating in this study, how familiar were you with the different voting systems?
(0=Not at all familiar; 100=Very familiar)
[rating scale items]

- Cardinal / "Rate Candidates"
- Plurality / "Pick One"
- Ranked-choice voting / "Rank Candidates"

Q2. How much have you heard about the different voting systems in the news?
(0=Not at all; 100=A lot)
[rating scale items]

- Cardinal / "Rate Candidates"
- Plurality / "Pick One"
- Ranked-choice voting / "Rank Candidates"

Q3. Prior to participating in this study, how positive/negative were you views about the different voting systems?
(-100 = Very negative; 100 = Very positive)
[rating scale items]

- Cardinal / "Rate Candidates"
- Plurality / "Pick One"
- Ranked-choice voting / "Rank Candidates"

Q4. Which voting system, if any, do you believe would be **most likely to benefit your preferred political candidate(s) or political party**?

- Cardinal / "Rate Candidates"
- Plurality / "Pick One"
- Ranked-choice voting / "Rank Candidates"
- None / They would all equally benefit my preferred political candidates or political party

Q5. Which voting system, if any, do you believe **your preferred political candidate(s) or political party would be most likely to endorse**?

- Cardinal / "Rate Candidates"
- Plurality / "Pick One"
- Ranked-choice voting / "Rank Candidates"
- They would be equally likely to endorse any of the above voting systems
- I don't know / Unsure

Q6. Is there anything else you would like to add or comment on about voting systems?
(If not you can skip this question)

[textbox]

Appendix D. Study 4 Survey Materials

D1. How to read tables introduction

[Raw+Summary Data Condition]
This study is about how groups of people make collective decisions. We will be using an example of a company hiring a job candidate.

Scenario:
There is a hiring committee of people who work for the company who have different preferences about which candidates would be best to hire. The company has realized that there are different ways that the hiring committee can vote for the candidates. One way is that each member of the hiring committee simply chooses their favorite candidate. We call this the 'Pick One Candidate' voting system. Another way is that each member of the hiring committee ranks each candidate as their first, second, or third favorite. We call this the 'Rank All Candidates' voting system. Another way is that each member of the hiring committee rates each candidate from 100 (best) to 0 (worst). We call this the 'Rate All Candidates' voting system. The example below has three job candidates ("A", "B", & "C") and six members on the hiring committee (Voters 1-6). We will hide the names of job candidates and voters just to make the tables less cluttered.

The company decides that instead of just using one voting system, the hiring committee will use all three voting systems. **Your job is to decide which candidate won based on the results from the three different voting systems.**

Pick One Candidate

Voter	Choice
1	A
2	B
3	A
4	A
5	C
6	B

Pick One Summary

Candidate	Votes (%)
A	3 (60.0%)
B	2 (33.3%)
C	1 (16.7%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	A	B	C
2	B	A	C
3	A	B	C
4	A	B	C
5	C	A	B
6	B	A	C

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
A	3 (60.0%)	4 (66.7%)
B	2 (33.3%)	2 (33.3%)
C	1 (16.7%)	X

Rate All Candidates

	Ratings		
Voter	A	B	C
1	100	50	15
2	82	90	30
3	95	70	25
4	87	60	20
5	73	25	100
6	70	80	50

Ratings Summary

	Candidates		
	A	B	C
Avg. Score	84.5	62.5	40.0

In the top half of the above boxes, we are showing you each voters' choices for the three different types of ballots (or voting methods).

- In the "Pick One Candidate" voter system the voters are asked to pick their favorite candidate.
- In the "Rank All Candidates" voter system the voters rank the candidates by preference (for example, 1st, 2nd, 3rd)
- In the "Rate All Candidates" voter system the voters rate the candidates on a 0-100 scale, where 0 is the lowest or worst rating and 100 is the highest or best rating.

In the bottom half of the above boxes, we are showing you the results from different types of ballots (or voting methods).

- The "Pick One Summary" table shows counts of voters' choices for their favorite candidate.
- The "Rankings Summary" table shows the results of voters ranking the candidates by preference. **This method works by first counting all "1st choice" votes. Next the candidate with the least support (here, Candidate C) is eliminated ("X" notes elimination). Votes that were cast for an eliminated candidate are re-distributed based on the next favorite candidate. This final voting count is seen in the "Final Choice" column.**
- The "Ratings Summary" table shows the average ratings after having voters rate the candidates on a 0-100 scale, where 0 is the lowest or worst rating and 100 is the highest or best rating.

Note that in this example Candidate A is the best candidate on all three ballots:

- Candidate A has 3 (60.0%) of the votes in the "Pick One Summary" table.
- Candidate A has 4 (66.7%) of the votes in the "Rankings Summary" table.
- Candidate A has an average rating score of 84.5, higher than any other candidate, in the

"Ratings Summary" table.

[Raw Data Only Condition]
This study is about how groups of people make collective decisions. We will be using an example of a company hiring a job candidate.

Scenario:
There is a hiring committee of people who work for the company who have different preferences about which candidates would be best to hire. The company has realized that there are different ways that the hiring committee can vote for the candidates. One way is that each member of the hiring committee simply chooses their favorite candidate. We call this the 'Pick One Candidate' voting system. Another way is that each member of the hiring committee ranks each candidate as their first, second, or third favorite. We call this the 'Rank All Candidates' voting system. Another way is that each member of the hiring committee rates each candidate from 100 (best) to 0 (worst). We call this the 'Rate All Candidates' voting system. The example below has three job candidates ("A", "B", & "C") and six members on the hiring committee (Voters 1-6). We will hide the names of job candidates and voters just to make the tables less cluttered.

The company decides that instead of just using one voting system, the hiring committee will use all three voting systems. **Your job is to decide which candidate won based on the results from the three different voting systems.**

Pick One Candidate

Voter	Choice
1	A
2	B
3	A
4	A
5	C
6	B

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	A	B	C
2	B	A	C
3	A	B	C
4	A	B	C
5	C	A	B
6	B	A	C

Rate All Candidates

	Ratings		
Voter	A	B	C
1	100	50	15
2	82	90	30
3	95	70	25
4	87	60	20
5	73	25	100
6	70	80	50

In the tables above, we are showing you each voters' choices for the three different types of ballots (or voting methods).

- In the "Pick One Candidate" voter system the voters are asked to pick their favorite candidate.
- In the "Rank All Candidates" voter system the voters rank the candidates by preference (for example, 1st, 2nd, 3rd)
- In the "Rate All Candidates" voter system the voters rate the candidates on a 0-100 scale, where 0 is the lowest or worst rating and 100 is the highest or best rating.

Note that in this example Candidate A is the best candidate on all three ballots.

[Summary Data Only Condition]
This study is about how groups of people make collective decisions. We will be using an example of a company hiring a job candidate.

Scenario:
There is a hiring committee of people who work for the company who have different preferences about which candidates would be best to hire. The company has realized that there are different ways that the hiring committee can vote for the candidates. One way is that each member of the hiring committee simply chooses their favorite candidate. We call this the 'Pick One Candidate' voting system. Another way is that each member of the hiring committee ranks each candidate as their first, second, or third favorite. We call this the 'Rank All Candidates' voting system. Another way is that each member of the hiring committee rates each candidate from 100 (best) to 0 (worst). We call this the 'Rate All Candidates' voting system. The example below has three job candidates ("A", "B", & "C") and six members on the hiring committee (Voters 1-6). We will hide the names of job candidates and voters just to make the tables less cluttered.

The company decides that instead of just using one voting system, the hiring committee will use all three voting systems. **Your job is to decide which candidate won based on the results from the three different voting systems.**

Pick One Summary

Candidate	Votes (%)
A	3 (60.0%)
B	2 (33.3%)
C	1 (16.7%)

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
A	3 (60.0%)	4 (66.7%)
B	2 (33.3%)	2 (33.3%)
C	1 (16.7%)	X

Ratings Summary

	Candidates		
	A	B	C
Avg. Score	84.5	62.5	40.0

In the tables above, we are showing you the results from different types of ballots (or voting methods).

- The "Pick One Summary" table shows counts of voters' choices for their favorite candidate.
- The "Rankings Summary" table shows the results of voters ranking the candidates by preference. **This method works by first counting all "1st choice" votes. Next the candidate with the least support (here, Candidate C) is eliminated ("X" notes elimination). Votes that were cast for an eliminated candidate are re-distributed based on the next favorite candidate. This final voting count is seen in the "Final Choice" column.**
- The "Ratings Summary" table shows the average ratings after having voters rate the candidates on a 0-100 scale, where 0 is the lowest or worst rating and 100 is the highest or best rating.

Note that in this example Candidate A is the best candidate on all three ballots:

- Candidate A has 3 (60.0%) of the votes in the "Pick One Summary" table.
- Candidate A has 4 (66.7%) of the votes in the "Rankings Summary" table.
- Candidate A has an average rating score of 84.5, higher than any other candidate, in the "Ratings Summary" table.

D2. Table Comprehension

[Note: Participants were assigned to one of three conditions: Summary and Raw Data, Raw Data Only, and Summary Data Only. Each condition had slightly different information presented to participants.]

[Summary and Raw Data Condition]
This study is about how groups of people make collective decisions. We will be using an example of a company hiring a job candidate.

Scenario:
There is a hiring committee of people who work for the company who have different preferences about which candidates would be best to hire. The company has realized that there are different ways that the hiring committee can vote for the candidates. One way is that each member of the hiring committee simply chooses their favorite candidate. We call this the 'Pick One Candidate' voting system. Another way is that each member of the hiring committee ranks each candidate as their first, second, or third favorite. We call this the 'Rank All Candidates' voting system. Another way is that each member of the hiring committee rates each candidate from 100 (best) to 0 (worst). We call this the 'Rate All Candidates' voting system. The example below has three job candidates ("A", "B", & "C") and six members on the hiring committee (Voters 1-6). We will hide the names of job candidates and voters just to make the tables less cluttered.

The company decides that instead of just using one voting system, the hiring committee will use all three voting systems. **Your job is to decide which candidate won based on the results from the three different voting systems.**

Pick One Candidate

Voter	Choice
1	A
2	B
3	A
4	A
5	C
6	B

Pick One Summary

Candidate	Votes (%)
A	3 (60.0%)
B	2 (33.3%)
C	1 (16.7%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	A	B	C
2	B	A	C
3	A	B	C
4	A	B	C
5	C	A	B
6	B	A	C

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
A	3 (60.0%)	4 (66.7%)
B	2 (33.3%)	2 (33.3%)
C	1 (16.7%)	X

Rate All Candidates

	Ratings		
Voter	A	B	C
1	100	50	15
2	82	90	30
3	95	70	25
4	87	60	20
5	73	25	100
6	70	80	50

Ratings Summary

	Candidates		
	A	B	C
Avg. Score	84.5	62.5	40.0

In the top half of the above boxes, we are showing you each voters' choices for the three different types of ballots (or voting methods).

- In the "Pick One Candidate" voter system the voters are asked to pick their favorite candidate.
- In the "Rank All Candidates" voter system the voters rank the candidates by preference (for example, 1st, 2nd, 3rd)
- In the "Rate All Candidates" voter system the voters rate the candidates on a 0-100 scale,

where 0 is the lowest or worst rating and 100 is the highest or best rating.

In the bottom half of the above boxes, we are showing you the results from different types of ballots (or voting methods).

- The "Pick One Summary" table shows counts of voters' choices for their favorite candidate.
- The "Rankings Summary" table shows the results of voters ranking the candidates by preference. **This method works by first counting all "1st choice" votes. Next the candidate with the least support (here, Candidate C) is eliminated ("X" notes elimination). Votes that were cast for an eliminated candidate are re-distributed based on the next favorite candidate. This final voting count is seen in the "Final Choice" column.**
- The "Ratings Summary" table shows the average ratings after having voters rate the candidates on a 0-100 scale, where 0 is the lowest or worst rating and 100 is the highest or best rating.

Note that in this example Candidate A is the best candidate on all three ballots:

- Candidate A has 3 (60.0%) of the votes in the "Pick One Summary" table.
- Candidate A has 4 (66.7%) of the votes in the "Rankings Summary" table.
- Candidate A has an average rating score of 84.5, higher than any other candidate, in the "Ratings Summary" table.

[**Raw Data Only Condition**]
This study is about how groups of people make collective decisions. We will be using an example of a company hiring a job candidate.

Scenario:
There is a hiring committee of people who work for the company who have different preferences about which candidates would be best to hire. The company has realized that there are different ways that the hiring committee can vote for the candidates. One way is that each member of the hiring committee simply chooses their favorite candidate. We call this the 'Pick One Candidate' voting system. Another way is that each member of the hiring committee ranks each candidate as their first, second, or third favorite. We call this the 'Rank All Candidates' voting system. Another way is that each member of the hiring committee rates each candidate from 100 (best) to 0 (worst). We call this the 'Rate All Candidates' voting system. The example below has three job candidates ("A", "B", & "C") and six members on the hiring committee (Voters 1-6). We will hide the names of job candidates and voters just to make the tables less cluttered.

The company decides that instead of just using one voting system, the hiring committee will use all three voting systems. **Your job is to decide which candidate won based on the results from the three different voting systems.**

Pick One Candidate

Voter	Choice
1	A
2	B
3	A
4	A
5	C
6	B

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	A	B	C
2	B	A	C
3	A	B	C
4	A	B	C
5	C	A	B
6	B	A	C

Rate All Candidates

	Ratings		
Voter	A	B	C
1	100	50	15
2	82	90	30
3	95	70	25
4	87	60	20
5	73	25	100
6	70	80	50

In the tables above, we are showing you each voters' choices for the three different types of ballots (or voting methods).

- In the "Pick One Candidate" voter system the voters are asked to pick their favorite candidate.
- In the "Rank All Candidates" voter system the voters rank the candidates by preference (for example, 1st, 2nd, 3rd)
- In the "Rate All Candidates" voter system the voters rate the candidates on a 0-100 scale, where 0 is the lowest or worst rating and 100 is the highest or best rating.

Note that in this example Candidate A is the best candidate on all three ballots.

[**Summary Data Only Condition**]
This study is about how groups of people make collective decisions. We will be using an example of a company hiring a job candidate.

Scenario:
There is a hiring committee of people who work for the company who have different preferences about which candidates would be best to hire. The company has realized that there are different ways that the hiring committee can vote for the candidates. One way is that each member of the hiring committee simply chooses their favorite candidate. We call this the 'Pick One Candidate' voting system. Another way is that each member of the hiring committee ranks each candidate as their first, second, or third favorite. We call this the 'Rank All Candidates' voting system. Another way is that each member of the hiring committee rates each candidate from 100 (best) to 0 (worst). We call this the 'Rate All Candidates' voting system. The example below has three job candidates ("A", "B", & "C") and six members on the hiring committee (Voters 1-6). We will hide the names of job candidates and voters just to make the tables less cluttered.

The company decides that instead of just using one voting system, the hiring committee will use all three voting systems. **Your job is to decide which candidate won based on the results from the three different voting systems.**

Pick One Summary

Candidate	Votes (%)
A	3 (60.0%)
B	2 (33.3%)
C	1 (16.7%)

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
A	3 (60.0%)	4 (66.7%)
B	2 (33.3%)	2 (33.3%)
C	1 (16.7%)	X

Ratings Summary

	Candidates		
	A	B	C
Avg. Score	84.5	62.5	40.0

In the tables above, we are showing you the results from different types of ballots (or voting methods).

- The "Pick One Summary" table shows counts of voters' choices for their favorite candidate.
- The "Rankings Summary" table shows the results of voters ranking the candidates by preference. **This method works by first counting all "1st choice" votes. Next the candidate with the least support (here, Candidate C) is eliminated ("X" notes elimination). Votes that were cast for an eliminated candidate are re-distributed based on the next favorite candidate. This final voting count is seen in the "Final Choice" column.**
- The "Ratings Summary" table shows the average ratings after having voters rate the candidates on a 0-100 scale, where 0 is the lowest or worst rating and 100 is the highest or best rating.

Note that in this example Candidate A is the best candidate on all three ballots:

- Candidate A has 3 (60.0%) of the votes in the "Pick One Summary" table.
- Candidate A has 4 (66.7%) of the votes in the "Rankings Summary" table.
- Candidate A has an average rating score of 84.5, higher than any other candidate, in the "Ratings Summary" table.

[This question was shown to participants in either: “Raw+Summary Data” or “Summary Data Only” conditions]

What does the "X" represent in the "Final Choice" column within the "Rankings Summary" table?

This method works by first counting all "1st choice" votes. Next the candidate with the least support (here, Candidate C) is eliminated ("X" notes elimination). Votes that were cast for an eliminated candidate are re-distributed based on the next favorite candidate. This final voting count is seen in the "Final Choice" column.

- "X" notes a candidate's elimination, where their votes are redistributed to the next favorite candidate in the "Final Choice" column. [*correct answer*]
- "X" notes that a candidate did not receive any votes.

[Feedback from the prior question was shown to participants in either: “Raw+Summary Data” or “Summary Data Only” conditions. Participants were shown one of the two messages below, depending if they selected the correct answer, or not.]

Correct! "X" notes a candidate's elimination, where their votes are redistributed to the next favorite candidate in the "Final Choice" column.

Incorrect! "X" notes a candidate's elimination, where their votes are redistributed to the next favorite candidate in the "Final Choice" column.

On the next 13 pages you will be presented with 13 different elections and **you need to decide which candidate should be hired**.

[Note: on the subsequent items, participants answered the next four questions, which were presented below the table stimuli for each condition]

Q1. Based on the information above, **rate the candidates based on who should be hired.**
(100=Should hire; 0=Should not hire)

- Candidate **A**
- Candidate **B**
- Candidate **C**

Q2. **Rank order the candidates based on who you think should be hired.**
(Top = first to hire; bottom = last to hire)

- Candidate **A**
- Candidate **B**
- Candidate **C**

Q3. **Which job candidate do you think should be hired?**
(choose one)

- Candidate **A**
- Candidate **B**
- Candidate **C**

Q4. Which table did you primarily use to make your decisions?

- “Pick One Candidate” (Left Table)
- “Rank All Candidates” (Middle Table)

- “Rate All Candidates” (Right Table)

[**Raw+Summary Data Condition**]

Item 1

Here is an example with three job candidates ("A", "B", & "C") and four members on the hiring committee.

Additionally, we added three summary tables which summarize the voting data for each ballot type.

- The "Pick One Summary" table shows that 3 of the hiring committee members voted to hire Candidate B and 1 member voted for Candidate C. No voter chose Candidate A.
- The "Rankings Summary" table shows that after eliminating the candidate with the least support (candidate A), 3 out of the 4 voters preferred Candidate B.
- The "Ratings Summary" table shows the average rating score for the three job candidates. Candidate B had the highest average rating with a score of 76.3

Pick One Candidate

Voter	Choice
1	B
2	B
3	B
4	C

Pick One Summary

Candidate	Votes (%)
B	3 (75.0%)
C	1 (25.0%)
A	0 (0.0%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	B	A	C
2	B	A	C
3	B	A	C
4	C	B	A

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
B	3 (75.0%)	3 (75.0%)
C	1 (25.0%)	1 (25.0%)
A	0 (0.0%)	X

Rate All Candidates

	Ratings		
Voter	A	B	C
1	60	95	25
2	50	90	30
3	40	80	35
4	0	40	90

Ratings Summary

	Candidates		
	A	B	C
Avg. Score	37.5	76.3	45.0

[**Raw Data Only Condition**]

Item 1

Here is an example with three job candidates ("A", "B", & "C") and four members on the hiring committee.

In this example:

- The "Pick One Candidate" table shows that 3 of the hiring committee members voted to hire Candidate B and 1 member voted for Candidate C. No voter chose Candidate A.
- The "Rank All Candidates" table shows that 3 out of the 4 voters preferred Candidate B, followed by Candidate A, with Candidate C as their 3rd pick.
- The "Rate All Candidates" table shows higher ratings on average for Candidate B, which

has an average rating score of 76.3, if you were to calculate the average.

Pick One Candidate

Voter	Choice
1	B
2	B
3	B
4	C

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	B	A	C
2	B	A	C
3	B	A	C
4	C	B	A

Rate All Candidates

	Ratings		
Voter	A	B	C
1	60	95	25
2	50	90	30
3	40	80	35
4	0	40	90

[**Summary Data Only Condition**]

Item 1

Here is an example with three job candidates ("A", "B", & "C") and four members on the hiring committee.

The three tables summarize the voting data for each ballot type.

- The "Pick One Summary" table shows that 3 of the hiring committee members voted to hire Candidate B and 1 member voted for Candidate C. No voter chose Candidate A.
- The "Rankings Summary" table shows that after eliminating the candidate with the least support (candidate A), 3 out of the 4 voters preferred Candidate B.
- The "Ratings Summary" table shows the average rating score for the three job candidates. Candidate B had the highest average rating with a score of 76.3

Pick One Summary

Candidate	Votes (%)
B	3 (75.0%)
C	1 (25.0%)
A	0 (0.0%)

Rankings Summary

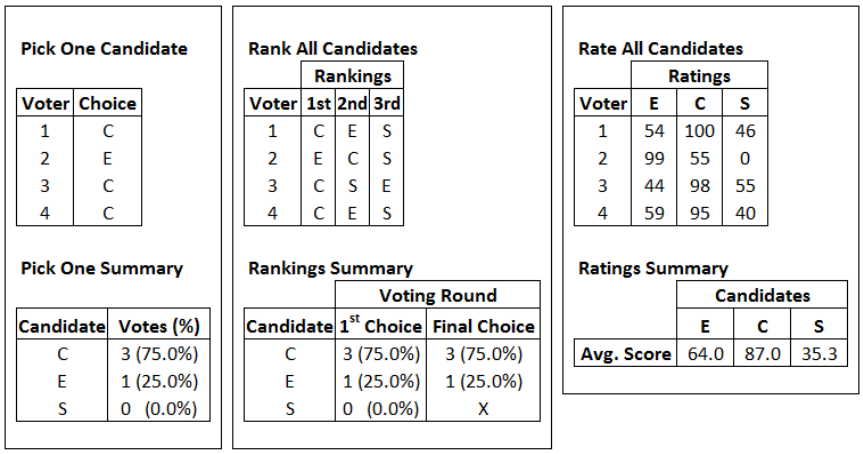
	Voting Round	
Candidate	1st Choice	Final Choice
B	3 (75.0%)	3 (75.0%)
C	1 (25.0%)	1 (25.0%)
A	0 (0.0%)	X

Ratings Summary

	Candidates		
	A	B	C
Avg. Score	37.5	76.3	45.0

[**Raw+Summary Data Condition**]

Item 2



Pick One Candidate

Voter	Choice
1	C
2	E
3	C
4	C

Pick One Summary

Candidate	Votes (%)
C	3 (75.0%)
E	1 (25.0%)
S	0 (0.0%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	C	E	S
2	E	C	S
3	C	S	E
4	C	E	S

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
C	3 (75.0%)	3 (75.0%)
E	1 (25.0%)	1 (25.0%)
S	0 (0.0%)	X

Rate All Candidates

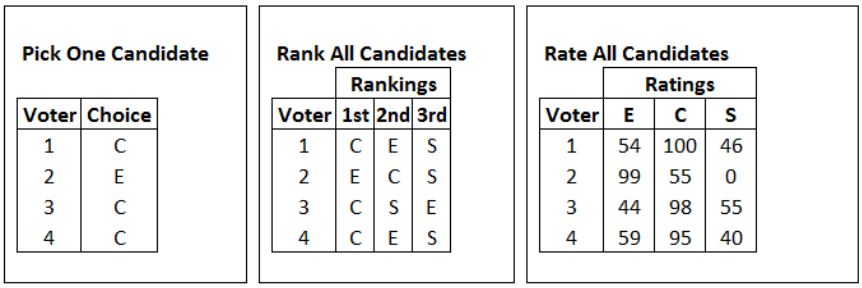
	Ratings		
Voter	E	C	S
1	54	100	46
2	99	55	0
3	44	98	55
4	59	95	40

Ratings Summary

	Candidates		
	E	C	S
Avg. Score	64.0	87.0	35.3

[**Raw Data Only Condition**]

Item 2



Pick One Candidate

Voter	Choice
1	C
2	E
3	C
4	C

Rank All Candidates

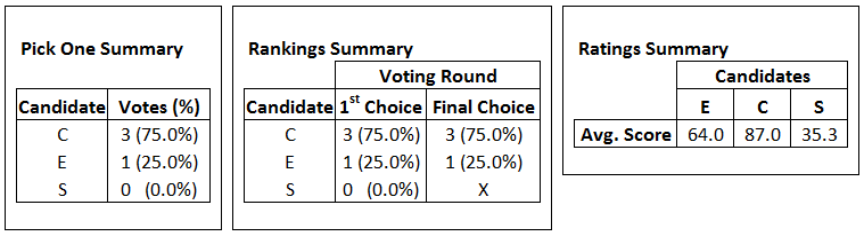
	Rankings		
Voter	1st	2nd	3rd
1	C	E	S
2	E	C	S
3	C	S	E
4	C	E	S

Rate All Candidates

	Ratings		
Voter	E	C	S
1	54	100	46
2	99	55	0
3	44	98	55
4	59	95	40

[**Summary Data Only Condition**]

Item 2



Pick One Summary

Candidate	Votes (%)
C	3 (75.0%)
E	1 (25.0%)
S	0 (0.0%)

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
C	3 (75.0%)	3 (75.0%)
E	1 (25.0%)	1 (25.0%)
S	0 (0.0%)	X

Ratings Summary

	Candidates		
	E	C	S
Avg. Score	64.0	87.0	35.3

[**Raw+Summary Data Condition**]

Item 3

Pick One Candidate

Voter	Choice
1	R
2	J
3	P
4	R
5	R
6	R

Pick One Summary

Candidate	Votes (%)
R	4 (66.7%)
J	1 (16.7%)
P	1 (16.7%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	R	P	J
2	J	R	P
3	P	R	J
4	R	P	J
5	R	P	J
6	R	P	J

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
R	4 (66.7%	5 (83.3%)
J	1 (16.7%)	1 (16.7%)
P	1 (16.7%)	X

Rate All Candidates

	Ratings		
Voter	J	R	P
1	32	96	65
2	69	59	27
3	0	80	96
4	35	92	61
5	20	100	76
6	27	100	69

Ratings Summary

	Candidates		
	J	R	P
Avg. Score	30.5	87.8	65.7

[Raw Data Only Condition]

Item 3

Pick One Candidate

Voter	Choice
1	R
2	J
3	P
4	R
5	R
6	R

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	R	P	J
2	J	R	P
3	P	R	J
4	R	P	J
5	R	P	J
6	R	P	J

Rate All Candidates

	Ratings		
Voter	J	R	P
1	32	96	65
2	69	59	27
3	0	80	96
4	35	92	61
5	20	100	76
6	27	100	69

[Summary Data Only Condition]

Item 3

Pick One Summary

Candidate	Votes (%)
R	4 (66.7%)
J	1 (16.7%)
P	1 (16.7%)

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
R	4 (66.7%	5 (83.3%)
J	1 (16.7%)	1 (16.7%)
P	1 (16.7%)	X

Ratings Summary

	Candidates		
	J	R	P
Avg. Score	30.5	87.8	65.7

D3. Who Won the Election Task

[Note: Participants were randomized to see stimuli with either 1) "Raw+Summary Data", 2) "Raw Data Only", or 3) "Summary Data Only." The major difference between the conditions is whether or not participants saw summary tables for each ballot type. For simplicity, only the "Raw+Summary Data" stimuli are shown below. See Appendix D2 for condition comparisons.]

For each item participants answered three questions which appeared below each figure.

1. Based on the above graph, **rate the candidates based on who should be hired**.
(100=most fair to hire; 0=least fair to hire)
- Candidate **A**
- Candidate **B**
- Candidate **C**

2. Rank order the candidates based on who you think should be hired.
(Top = first to hire; bottom = last to hire)
- Candidate **A**
- Candidate **B**
- Candidate **C**

3. Which job candidate do you think should be hired?
(choose one)
- Candidate **A**
- Candidate **B**
- Candidate **C**

4. Which table did you primarily use to make your decisions?
- "Pick One Candidate" (Left Table)
- "Rank All Candidates" (Middle Table)
- "Rate All Candidates" (Right Table)

The stimuli items are presented below.

Item 1

Pick One Candidate

Voter	Choice
1	K
2	S
3	K
4	K
5	A
6	A
7	A
8	K
9	S
10	K
11	A

Pick One Summary

Candidate	Votes (%)
K	5 (45.5%)
A	4 (36.4%)
S	2 (18.2%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	K	A	S
2	S	A	K
3	K	A	S
4	K	A	S
5	A	K	S
6	A	K	S
7	A	S	K
8	K	A	S
9	S	A	K
10	K	A	S
11	A	K	S

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
A	4 (36.4%)	6 (54.5%)
K	5 (45.5%)	5 (45.5%)
S	2 (18.2%)	X

Rate All Candidates

	Ratings		
Voter	K	A	S
1	95	54	8
2	9	50	96
3	100	60	14
4	87	46	0
5	62	97	51
6	73	87	41
7	46	87	68
8	92	68	22
9	16	57	97
10	91	50	4
11	67	93	47

Ratings Summary

	Candidates		
	K	A	S
Avg. Score	67.1	68.1	40.7

Item 2

Pick One Candidate

Voter	Choice
1	G
2	C
3	Z
4	C
5	C
6	G
7	G

Pick One Summary

Candidate	Votes (%)
C	3 (42.9%)
G	3 (42.9%)
Z	1 (14.3%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	G	Z	C
2	C	Z	G
3	Z	G	C
4	C	Z	G
5	C	Z	G
6	G	Z	C
7	G	Z	C

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
G	3 (42.9%)	4 (57.1%)
C	3 (42.9%)	3 (42.9%)
Z	1 (14.3%)	X

Rate All Candidates

	Ratings		
Voter	C	Z	G
1	16	63	92
2	90	65	19
3	46	92	63
4	92	46	0
5	95	60	14
6	9	55	100
7	25	71	84

Ratings Summary

	Candidates		
	C	Z	G
Avg. Score	53.3	64.6	53.1

Item 3

Pick One Candidate

Voter	Choice
1	A
2	F
3	F
4	A
5	A
6	F
7	F

Pick One Summary

Candidate	Votes (%)
F	4 (57.1%)
A	3 (42.9%)
O	0 (0.0%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	A	O	F
2	F	O	A
3	F	O	A
4	A	O	F
5	A	O	F
6	F	O	A
7	F	O	A

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
F	4 (57.1%)	4 (57.1%)
A	3 (42.9%)	3 (42.9%)
O	0 (0.0%)	X

Rate All Candidates

	Ratings		
Voter	A	O	F
1	88	44	23
2	12	82	98
3	9	79	100
4	91	21	0
5	94	37	16
6	16	86	95
7	19	89	92

Ratings Summary

	Candidates		
	A	O	F
Avg. Score	47.0	62.6	60.6

Item 4

Pick One Candidate

Voter	Choice
1	L
2	L
3	W
4	N
5	W
6	N
7	W
8	N
9	N

Pick One Summary

Candidate	Votes (%)
N	4 (44.4%)
W	3 (33.3%)
L	2 (22.2%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	L	W	N
2	L	W	N
3	W	L	N
4	N	L	W
5	W	L	N
6	N	L	W
7	W	L	N
8	N	L	W
9	N	L	W

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
W	3 (33.3%)	5 (55.6%)
N	4 (44.4%)	4 (44.4%)
L	2 (22.2%)	X

Rate All Candidates

	Ratings		
Voter	N	L	W
1	44	94	69
2	53	84	60
3	3	66	90
4	91	47	23
5	0	63	87
6	91	28	4
7	21	83	93
8	94	44	19
9	100	38	13

Ratings Summary

	Candidates		
	N	L	W
Avg. Score	55.2	60.8	50.9

Item 5

Pick One Candidate

Voter	Choice
1	I
2	Q
3	I
4	I
5	Z
6	Q
7	Z

Pick One Summary

Candidate	Votes (%)
I	3 (42.9%)
Q	2 (28.6%)
Z	2 (28.6%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	I	Q	Z
2	Q	I	Z
3	I	Q	Z
4	I	Q	Z
5	Z	Q	I
6	Q	I	Z
7	Z	Q	I

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
Q	2 (28.6%)	4 (57.1%)
I	3 (42.9%)	3 (42.9%)
Z	2 (28.6%)	X

Rate All Candidates

	Ratings		
Voter	I	Q	Z
1	81	75	0
2	82	88	48
3	84	79	3
4	86	80	5
5	41	47	89
6	80	86	50
7	30	36	100

Ratings Summary

	Candidates		
	I	Q	Z
Avg. Score	69.1	70.1	42.1

Item 6

Pick One Candidate

Voter	Choice
1	J
2	J
3	S
4	J
5	J
6	J
7	U
8	U
9	S
10	U
11	J
12	U
13	U
14	S

Pick One Summary

Candidate	Votes (%)
J	6 (42.9%)
U	5 (35.7%)
S	3 (21.4%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	J	S	U
2	J	S	U
3	S	U	J
4	J	S	U
5	J	S	U
6	J	S	U
7	U	S	J
8	U	S	J
9	S	U	J
10	U	S	J
11	J	S	U
12	U	S	J
13	U	S	J
14	S	U	J

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
U	5 (35.7%)	8 (57.1%)
J	6 (42.9%)	6 (42.9%)
S	3 (21.4%)	X

Rate All Candidates

	Ratings		
Voter	J	S	U
1	95	45	4
2	84	67	26
3	43	93	67
4	94	58	17
5	100	51	10
6	97	55	14
7	8	58	99
8	18	68	92
9	54	97	56
10	22	72	88
11	97	47	5
12	11	61	99
13	0	50	91
14	45	95	65

Ratings Summary

	Candidates		
	J	S	U
Avg. Score	54.9	65.5	52.4

Item 7

Pick One Candidate

Voter	Choice
1	K
2	T
3	K
4	T
5	T
6	T
7	K
8	K
9	K

Pick One Summary

Candidate	Votes (%)
K	5 (55.6%)
T	4 (44.4%)

Rank All Candidates

	Rankings	
Voter	1st	2nd
1	K	T
2	T	K
3	K	T
4	T	K
5	T	K
6	T	K
7	K	T
8	K	T
9	K	T

Rankings Summary

	Voting Round
Candidate	1st Choice
K	5 (55.6%)
T	4 (44.4%)

Rate All Candidates

	Ratings	
Voter	T	K
1	51	53
2	95	8
3	29	75
4	99	4
5	98	0
6	90	14
7	41	63
8	3	100
9	18	85

Ratings Summary

	Candidates	
	T	K
Avg. Score	58.2	44.7

Item 8

Pick One Candidate

Voter	Choice
1	M
2	L
3	M
4	U
5	M
6	U
7	L
8	L
9	L
10	L
11	U
12	M
13	U
14	L
15	U
16	U
17	U
18	M

Pick One Summary

Candidate	Votes (%)
U	7 (38.9%)
L	6 (33.3%)
M	5 (27.8%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	M	U	L
2	L	M	U
3	M	L	U
4	U	M	L
5	M	L	U
6	U	M	L
7	L	M	U
8	L	M	U
9	L	M	U
10	L	M	U
11	U	M	L
12	M	L	U
13	U	M	L
14	L	M	U
15	U	M	L
16	U	M	L
17	U	M	L
18	M	L	U

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
L	6 (33.3%)	10
U	7 (38.9%)	8
M	5 (27.8%)	X

Rate All Candidates

	Ratings		
Voter	L	M	U
1	60	92	61
2	94	74	26
3	66	99	55
4	6	39	87
5	71	97	49
6	16	48	96
7	90	79	31
8	97	71	24
9	96	63	15
10	99	67	19
11	32	64	89
12	74	94	47
13	10	42	90
14	87	82	34
15	25	58	96
16	0	33	80
17	21	53	100
18	62	95	58

Ratings Summary

	Candidates		
	L	M	U
Avg. Score	55.9	69.4	58.7

Item 9

Pick One Candidate

Voter	Choice
1	Z
2	X
3	Z
4	Z
5	Z
6	C
7	X
8	X
9	X
10	C
11	X
12	Z

Pick One Summary

Candidate	Votes (%)
X	5 (41.7%)
Z	5 (41.7%)
C	2 (16.7%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	Z	C	X
2	X	C	Z
3	Z	C	X
4	Z	C	X
5	Z	C	X
6	C	X	Z
7	X	C	Z
8	X	C	Z
9	X	C	Z
10	C	X	Z
11	X	C	Z
12	Z	C	X

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
X	5 (41.7%)	7 (58.3%)
Z	5 (41.7%)	5 (41.7%)
C	2 (16.7%)	X

Rate All Candidates

	Ratings		
Voter	Z	C	X
1	92	60	16
2	6	56	100
3	90	63	19
4	94	44	0
5	99	49	5
6	47	97	61
7	11	60	98
8	2	52	96
9	20	70	89
10	43	92	66
11	18	67	91
12	96	57	13

Ratings Summary

	Candidates		
	Z	C	X
Avg. Score	51.5	63.9	54.5

Item 10

Pick One Candidate

Voter	Choice
1	J
2	N
3	J
4	J
5	N
6	Q
7	N
8	J
9	J
10	Q
11	N

Pick One Summary

Candidate	Votes (%)
J	5 (45.5%)
N	4 (36.4%)
Q	2 (18.2%)

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	J	Q	N
2	N	Q	J
3	J	Q	N
4	J	Q	N
5	N	Q	J
6	Q	N	J
7	N	Q	J
8	J	Q	N
9	J	Q	N
10	Q	N	J
11	N	Q	J

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
N	4 (36.4%)	6 (54.5%)
J	5 (45.5%)	5 (45.5%)
Q	2 (18.2%)	X

Rate All Candidates

	Ratings		
Voter	N	Q	J
1	7	16	50
2	55	47	12
3	15	23	58
4	24	33	67
5	47	39	4
6	85	93	76
7	60	52	17
8	0	9	43
9	35	43	78
10	91	100	69
11	68	60	25

Ratings Summary

	Candidates		
	N	Q	J
Avg. Score	44.3	46.8	45.4

D4. Voting system introduction

[Note: Participants were formally introduced to cardinal, plurality, and ranked-choice voting in this block. The voting system order was randomized for each participant.]

After each voting system was introduced the participant was asked:

> Q. How fair is [cardinal/plurality/RCV] voting? Rate from very fair (100) to not very fair (0). [slider question format]

The stimuli follows:

The next three pages will teach you about three different ways to determine who should win when multiple people are voting.

[Raw+Summary Data]

[Cardinal]

One method to determine election outcomes is to have voters rate each candidate. This is called **cardinal voting**. In cardinal voting the candidate that has the highest average rating wins. In the information presented below we show the ratings of four voters for three candidates.

As you can see, C has the lowest voter rating on average [math: (100+67+0+25)/4=48.0]. A and B both have higher voter ratings on average. However, B has a better voter rating on average [math: (67+100+50+75)/4=73.0] than A [math: (17+50+100+92)/4=64.8]. Thus, according to cardinal voting, **candidate B would win**.

Rate All Candidates

	Ratings		
Voter	**A**	**B**	**C**
1	17	67	100
2	50	100	67
3	100	50	0
4	92	75	25

Ratings Summary

	Candidates		
	A	**B**	**C**
Avg. Score	64.8	73.0	48.0

[Plurality]

One method to determine election outcomes is to allow all voters to choose only one candidate and the candidate with the most votes wins. This is called **plurality voting**. Plurality voting

prioritizes selecting a winner who has the single most support. A winning candidate in a plurality system needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.

For example, in the information presented below “Candidate A” has 4 out of 9 hiring committee members support. “Candidate B” has 2 out of the 9 hiring committee members support. “Candidate C” has 3 out of the 9 hiring committee members support. Using a Plurality voting system, Candidate A would win, even though they did not receive more than 50% of the votes.

Pick One Candidate

Voter	Choice
1	B
2	A
3	A
4	A
5	C
6	C
7	B
8	A
9	C

Pick One Summary

Candidate	Votes (%)
A	4 (44.4%)
C	3 (33.3%)
B	2 (22.2%)

[Ranked-Choice Voting]

One method to determine election outcomes is to focus on how each voter ranks the candidates (e.g., first, second, third preference). This is called **ranked-choice voting**. Ranked-choice voting prioritizes majority approval and a winning candidate must have over 50% of the total votes. This works by first counting all voters' first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is “eliminated” and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support. If so, this candidate wins. If not, the process is repeated.

For example, in the information presented below, there are four voters who rank Candidate A as their first choice. There are three voters who rank C as their first choice and there are two voters who rank B as their first choice.

Although A has the most first choice votes (4 out of 9), A does not have over 50% of the votes.

Because B has the fewest first choice votes, B would be removed as a candidate, and B's votes would be reassigned. In this case, the two voters who most preferred B would be reassigned to C, because this is their 2nd choice. Once B is removed, there are 5 votes for C, and 4 votes for A, so Candidate C would win.

Rank All Candidates

	Rankings		
Voter	1st	2nd	3rd
1	C	B	A
2	B	C	A
3	C	B	A
4	A	B	C
5	B	C	A
6	A	B	C
7	C	B	A
8	A	B	C
9	A	B	C

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
C	3 (33.3%)	5 (55.6%)
A	4 (44.4%)	4 (44.4%)
B	2 (22.2%)	X

[Raw Data Only Condition]

[Cardinal]

One method to determine election outcomes is to have voters rate each candidate. This is called **cardinal voting**. In cardinal voting the candidate that has the highest average rating wins. In the information presented below we show the ratings of four voters for three candidates.

As you can see, C has the lowest voter rating on average [math: (100+67+0+25)/4=48.0]. A and B both have higher voter ratings on average. However, B has a better voter rating on average [math: (67+100+50+75)/4=73.0] than A [math: (17+50+100+92)/4=64.8]. Thus, according to cardinal voting, **candidate B would win**.

Rate All Candidates

	Ratings		
Voter	A	B	C
1	17	67	100
2	50	100	67
3	100	50	0
4	92	75	25

[Plurality]
One method to determine election outcomes is to allow all voters to choose only one candidate and the candidate with the most votes wins. This is called **plurality voting**. Plurality voting prioritizes selecting a winner who has the single most support. A winning candidate in a plurality system needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.

For example, in the information presented below "Candidate A" has 4 out of 9 hiring committee members support. "Candidate B" has 2 out of the 9 hiring committee members support. "Candidate C" has 3 out of the 9 hiring committee members support. Using a Plurality voting system, Candidate A would win, even though they did not receive more than 50% of the votes.

Pick One Candidate

Voter	Choice
1	B
2	A
3	A
4	A
5	C
6	C
7	B
8	A
9	C

[Ranked-Choice Voting]
One method to determine election outcomes is to focus on how each voter ranks the candidates (e.g., first, second, third preference). This is called **ranked-choice voting**. Ranked-choice voting prioritizes majority approval and a winning candidate must have over 50% of the total votes. This works by first counting all voters' first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is "eliminated" and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support. If so, this candidate wins. If not, the process is repeated.

For example, in the information presented below, there are four voters who rank Candidate A as their first choice. There are three voters who rank C as their first choice and there are two voters who rank B as their first choice.

Although A has the most first choice votes (4 out of 9), A does not have over 50% of the votes. Because B has the fewest first choice votes, B would be removed as a candidate, and B's votes

would be reassigned. In this case, the two voters who most preferred B would be reassigned to C, because this is their 2nd choice. Once B is removed, there are 5 votes for C, and 4 votes for A, so Candidate C would win.

Rank All Candidates

	Rankings		
Voter	**1st**	**2nd**	**3rd**
1	C	B	A
2	B	C	A
3	C	B	A
4	A	B	C
5	B	C	A
6	A	B	C
7	C	B	A
8	A	B	C
9	A	B	C

[Summary Data Only Condition]

[Cardinal]
One method to determine election outcomes is to have voters rate each candidate. This is called **cardinal voting**. In cardinal voting the candidate that has the highest average rating wins. In the information presented below we show the ratings of four voters for three candidates.

As you can see, C has the lowest voter rating on average [math: (100+67+0+25)/4=48.0]. A and B both have higher voter ratings on average. However, B has a better voter rating on average [math: (67+100+50+75)/4=73.0] than A [math: (17+50+100+92)/4=64.8]. Thus, according to cardinal voting, **candidate B would win**.

Ratings Summary

	Candidates		
	A	**B**	**C**
Avg. Score	64.8	73.0	48.0

[Plurality]
One method to determine election outcomes is to allow all voters to choose only one candidate and the candidate with the most votes wins. This is called **plurality voting**. Plurality voting prioritizes selecting a winner who has the single most support. A winning candidate in a plurality system needs to have more votes than any other candidate but does not necessarily need to have over 50% of the votes to win.

For example, in the information presented below “Candidate A” has 4 out of 9 hiring committee members support. “Candidate B” has 2 out of the 9 hiring committee members support. “Candidate C” has 3 out of the 9 hiring committee members support. Using a Plurality voting system, Candidate A would win, even though they did not receive more than 50% of the votes.

Pick One Summary

Candidate	Votes (%)
A	4 (44.4%)
C	3 (33.3%)
B	2 (22.2%)

[Ranked-Choice Voting]

One method to determine election outcomes is to focus on how each voter ranks the candidates (e.g., first, second, third preference). This is called **ranked-choice voting**. Ranked-choice voting prioritizes majority approval and a winning candidate must have over 50% of the total votes. This works by first counting all voters' first preference and seeing if any candidate has over 50% support. If so, this candidate wins. If not, then the candidate with the least amount of support is “eliminated” and their votes are re-assigned to their next favorite candidate. The votes are then counted again to see if a candidate has over 50% support. If so, this candidate wins. If not, the process is repeated.

For example, in the information presented below, there are four voters who rank Candidate A as their first choice. There are three voters who rank C as their first choice and there are two voters who rank B as their first choice.

Although A has the most first choice votes (4 out of 9), A does not have over 50% of the votes. Because B has the fewest first choice votes, B would be removed as a candidate, and B’s votes would be reassigned. In this case, the two voters who most preferred B would be reassigned to C, because this is their 2nd choice. Once B is removed, there are 5 votes for C, and 4 votes for A, so Candidate C would win.

Rankings Summary

	Voting Round	
Candidate	1st Choice	Final Choice
C	3 (33.3%)	5 (55.6%)
A	4 (44.4%)	4 (44.4%)
B	2 (22.2%)	X

D5. Open-Ended Voting System Pros/Cons Questions

In this study you have learned about three different voting systems (cardinal/"rate candidates", plurality/"pick one", and ranked-choice voting/"rank candidates").

On the next few pages you will be asked to explain what you like and don't like about these voting systems.

[Note: The presentation for 3 voting systems' open-ended questions was counterbalanced.]

[Cardinal]
What are the reasons you like cardinal/"rate candidates" voting?
(Please explain your reasoning by writing at least a few sentences.)

What are the reasons you don't like cardinal/"rate candidates" voting?
(Please explain your reasoning by writing at least a few sentences.)

[Plurality]
What are the reasons you like plurality/"pick one" voting?
(Please explain your reasoning by writing at least a few sentences.)

What are the reasons you don't like plurality/"pick one" voting?
(Please explain your reasoning by writing at least a few sentences.)

[Ranked-Choice]
What are the reasons you like ranked-choice/"rank candidates" voting?
(Please explain your reasoning by writing at least a few sentences.)

What are the reasons you don't like ranked-choice/"rank candidates" voting?
(Please explain your reasoning by writing at least a few sentences.)

Appendix E. Choice Consistency Linear Discriminant Analysis Results

Equal prior probabilities (.33) were set for participant's voting system group membership in both analyses. The reason for this decision (in contrast to using proportional group membership) is because voting system preferences were measured after the Who Won the Election Task and voting system preferences may have been influenced by the experience. Additionally, a "leave-one-out" internal classification system was used, where each case is classified by the discriminant functions from all other cases other than itself (Lachenbruch & Mickey, 1968). This method reduces bias and provides more conservative estimates than an uncorrected classification method.

Study 1

Predicting Voting System Preferences by Choices in the Who Won the Election Task

We found that the 3 predictors of representing competing voting mechanisms for determining electoral outcomes (plurality, ranked-choice, and cardinal voting choices in the Who Won the Election Task) were able to significantly distinguish participants' voting system preferences, $\lambda = .87$, $\chi^2(6) = 39.18$, $p < .001$. The second function was not found to be significant, $\lambda = .99$, $\chi^2(2) = 2.59$, $p = .275$. Of the variance explained, 93.8% is explained by Function 1, with only 6.2% explained by Function 2.

Function 1 had high standardized discriminant function coefficients (i.e., a measure of unique variance for predictors) for WWET cardinal scores (.75) and WWET plurality scores (-.52) and a lower coefficient for WWET RCV scores (.24). Additionally, Function 1 had high structure matrix coefficients (i.e., a measure of individual, non-unique, contributions) for Who Won the Election Task choice scores for cardinal (.98), plurality (-.90), and RCV (-.82).

Together, these results indicate that Function 1 distinguished participants who chose the cardinal option and those who chose the plurality option, with RCV choices being somewhere in between these two distributions. However, the structure matrix coefficients indicate that Ranked-Choice voting choices tended to be more similar to Plurality choices, than cardinal (see Table E1).

Table E1. Discriminant Functions for Who Won the Election Task Choices Predicting Voting System Preferences

	Standardized Discriminant Function Coefficients		Structure Matrix Coefficients	
Predictor Variables	*Function 1*	*Function 2*	*Function 1*	*Function 2*
WWET Cardinal	.75	1.44	.98	.12
WWET Plurality	-.52	.44	-.90	.42
WWET RCV	.24	1.23	-.82	.53

The overall hit rate for successfully predicting voting system preferences from choices in the Who Won the Election Task was 41.50%. A Huberty's *I* of .13 was found, indicating a small effect size in the Who Won the Election Task's ability to predict voting system preferences (Huberty, & Lowman, 2000). The hit rate for three groups was 52.10% for plurality ($z = 3.42$, $p < .001$), 30.2% for RCV ($z = -.73$, $p = .465$) and 56.30% ($z = 3.96$, $p < .001$) for cardinal voting. From the Who Won the Election Task choices we could correctly classify participants who preferred plurality and cardinal voting systems, but the predictions for RCV participants was at

chance levels.

Predicting Voting System Preferences by Choices in the Who Won the Election Task and Individual Difference Measures

Beyond examining if choices in the Who Won the Election Task alone can predict participants' voting system preferences, we were interested if the addition of important individual difference measures could increase predictability of voting system preferences. Additionally, we want to assess the relative contribution of competing measures. Two questions are being addressed here. The first, what are the patterns of relations across the individual difference measures, including Who Won the Election Task choices, for predicting voting system preferences. The second, how does the introduction of individual difference measures, beyond just the inclusion participants' choices in the Who Won the Election Task, affect our ability to predict voting preferences.

We found that the choices in the Who Won the Election Task and the individual difference measures together were able to significantly distinguish participants' voting system preferences, $\lambda = .78$, $\chi^2(26) = 69.37$, $p < .001$. Additionally, the second function was also found to be significant, $\lambda = .92$, $\chi^2(12) = 22.46$, $p = .033$. Of the variance explained, 68.6% is explained by Function 1 and 31.4% is explained by Function 2.

Function 1 had high standardized discriminant function coefficients (i.e., a measure of unique variance for predictors) for Who Won the Election Task Plurality scores (-.38), WWET cardinal scores (.53), and lower standardized discriminant coefficients for utilitarianism (.28) and CRT (.20). Additionally, Function 1 had high structure matrix coefficients (i.e., a measure of individual, non-unique, contributions) for WWET cardinal choices (.85), WWET plurality choices (-.78), WWET RCV (-.71) and lower structure matrix coefficients for CRT (.37), learning comprehension score (.34), utilitarianism (.26), and numeracy (.23).

Function 2 had high standardized discriminant function coefficients for Who Won the Election Task plurality choices (.53), fairness/reciprocity (.70), and authority/respect (-.87), and lower standardized discriminant function coefficients for Who Won the Election Task cardinal choices (.34), harm/care (-.46), and conservatism (.41). Additionally, Function 2 had high structure matrix coefficients for authority/respect (-.65), in-group/loyalty (-.43), and purity/sanctity (-.40), and lower structure matrix coefficients for WWET plurality choices (.19), WWET RCV choices (.21), fairness/reciprocity (.29), conservatism (-.21), and harm/care (-.10). See Table E2 for complete results.

Table E2. Discriminant Functions for Individual Difference Measures Predicting Voting System Preferences

	Standardized Discriminant Function Coefficients		Structure Matrix Coefficients	
Predictor Variables	*Function 1*	*Function 2*	*Function 1*	*Function 2*
WWET Cardinal	.54	.34	.85	-.01
WWET Plurality	-.38	.53	-.78	.19
WWET RCV	.06	.20	-.71	.21
Authority/Respect	-.29	-.87	-.24	-.65
Fairness/Reciprocity	-.30	.70	-.11	.29
Harm/Care	.19	-.46	-.01	-.10
In-group/Loyalty	.03	.07	-.12	-.43
Purity/Sanctity	-.01	-.06	-.15	-.40
Conservatism	.27	.41	-.01	-.21
Utilitarianism	.28	-.17	.26	-.11
CRT	.20	.27	.37	.24
Numeracy	.05	.02	.23	.02
Learning Comp. Score	.10	.22	.34	.28

Function 1, which accounted for most of the explained variance, distinguished participants who tended to choose less of the Election Task Plurality choices and more of the Election Task Cardinal choices. These participants tended to be higher in utilitarianism, cognitive reflection (CRT), numeracy, and learning comprehension scores. Function 2 found meaningful patterns among the moral/ideological variables. Function 2 best distinguished participants who chose Election Task Plurality choices and were high in fairness but low in authority/respect. These participants also tended to be lower in harm/care, in-group/loyalty, and purity/sanctity.

The overall hit rate for successfully predicting voting system preferences from choices in the Election Task was 41.90%. A Huberty's *I* of .13 was found, indicating a small effect size in the Election Task choices and individual difference measures' ability to predict voting system preferences (Huberty, & Lowman, 2000). The hit rate for three groups was 52.10% for plurality ($z = 3.42$, $p < .001$), 32.20% for RCV ($z = -.21$, $p = .833$) and 53.10% ($z = 3.42$, $p < .001$) for cardinal voting. From the included predictor variables, we could correctly classify participants who preferred plurality and cardinal voting systems, but the predictions for RCV participants was again at chance levels.

Study 2

A similar analysis to Study 1 was completed for Study 2. However, because participants completed the Who Won the Election Task twice, followed by two separate sessions of rating the fairness of the competing voting systems, two linear discriminant analysis were conducted. The first predicting voting system preferences (Time 1) from WWET choices (Time 1) and the second predicting voting system preferences (Time 2) from WWET choices (Time 2).

Time Point 1 (Before Formally Learning About Voting Systems)

We found that the 3 predictors of representing competing voting mechanisms for determining electoral outcomes (plurality, ranked-choice, and cardinal voting choices in the Who Won the Election Task) were able to significantly distinguish participants' voting system preferences, $\lambda =$

.88, $\chi^2(6) = 40.60$, $p < .001$. The second function was not found to be significant, $\lambda = 1.00$, $\chi^2(2) = .69$, $p = .709$. Of the variance explained, 98.4% is explained by Function 1, with only 1.6% explained by Function 2.

Function 1 had high standardized discriminant function coefficients (i.e., a measure of unique variance for predictors) for WWET cardinal scores (.80),WWET plurality scores (-.84) and WWET RCV scores (.68). Additionally, Function 1 had high structure matrix coefficients (i.e., a measure of individual, non-unique, contributions) for Who Won the Election Task choice scores for cardinal (.97), plurality (-.85), and RCV (-.71).

Together, these results indicate that Function 1 distinguished participants who chose the cardinal option and those who chose the plurality option, with RCV choices being somewhere in between these two distributions. However, the structure matrix coefficients indicate that Ranked-Choice voting choices tended to be more similar to Plurality choices, than cardinal (see Table B3).

Table E3. Discriminant Functions for Who Won the Election Task Choices (Time Point 1) Predicting Voting System Preferences (Time Point 1)

	Standardized Discriminant Function Coefficients		Structure Matrix Coefficients	
Predictor Variables	*Function 1*	*Function 2*	*Function 1*	*Function 2*
WWET Cardinal	.80	1.41	.97	.21
WWET Plurality	-.84	2.77	-.85	.33
WWET RCV	.68	-1.39	-.71	.15

The overall hit rate for successfully predicting voting system preferences from choices in the Who Won the Election Task was 41.90%. A Huberty's *I* of .13 was found, indicating a small effect size in the Who Won the Election Task's ability to predict voting system preferences (Huberty, & Lowman, 2000). The hit rate for three groups was 59.80% for plurality ($z = 5.76$, $p < .001$), 27.10% for RCV ($z = -1.56$, $p = .119$) and 49.21% ($z = 2.74$, $p < .01$) for cardinal voting. From the Who Won the Election Task choices we could correctly classify participants who preferred plurality and cardinal voting systems, but the predictions for RCV participants was at chance levels.

Time Point 2 (After Formally Learning About Voting Systems)

We found that the 3 predictors of representing competing voting mechanisms for determining electoral outcomes (plurality, ranked-choice, and cardinal voting choices in the Who Won the Election Task) were able to significantly distinguish participants' voting system preferences, $\lambda = .81$, $\chi^2(6) = 64.99$, $p < .001$. The second function was not found to be significant, $\lambda = 1.00$, $\chi^2(2) = 1.24$, $p = .538$. Of the variance explained, 98.3% is explained by Function 1, with only 1.7% explained by Function 2.

Function 1 had high standardized discriminant function coefficients (i.e., a measure of unique variance for predictors) for WWET plurality scores (1.13), and lower standardized discriminant function coefficients for WWET cardinal scores (-.44) and WWET RCV scores (-.59). Additionally, Function 1 had high structure matrix coefficients (i.e., a measure of individual, non-unique, contributions) for Who Won the Election Task choice scores for cardinal (-.92), plurality (.94), and RCV (-.59).

Together, these results indicate that Function 1 distinguished participants who chose the cardinal option and those who chose the plurality option, with RCV choices being somewhere in

between these two distributions. However, the structure matrix coefficients indicate that Ranked-Choice voting choices tended to be more similar to Plurality choices, than cardinal (see Table B4).

Table E4. Discriminant Functions for Who Won the Election Task Choices (Time Point 2) Predicting Voting System Preferences (Time Point 2)

	Standardized Discriminant Function Coefficients		Structure Matrix Coefficients	
Predictor Variables	*Function 1*	*Function 2*	*Function 1*	*Function 2*
WWET Cardinal	-.44	.64	-.92	-.01
WWET Plurality	1.13	-1.33	.94	.32
WWET RCV	-.59	.64	.79	.61

The overall hit rate for successfully predicting voting system preferences from choices in the Who Won the Election Task was 47.50%. A Huberty's *I* of .22 was found, indicating a medium effect size in the Who Won the Election Task's ability to predict voting system preferences (Huberty, & Lowman, 2000). The hit rate for three groups was 61.32% for plurality ($z = 6.20$, $p < .001$), 30.34% for RCV ($z = -.68$, $p = .497$) and 62.32% ($z = 5.18$, $p < .001$) for cardinal voting. From the Who Won the Election Task choices we could correctly classify participants who preferred plurality and cardinal voting systems, but the predictions for RCV participants was at chance levels.

Study 3

In Study 3 participants were randomized to either "raw+summary data" or "raw data only" version of the Who Won the Election Task. Because participants made similar WWET choices across both conditions I run the LDA collapsing across condition (which boosted the statistical power of the analysis).

LDA Collapsing across condition assignment

We found that the 3 predictors of representing competing voting mechanisms for determining electoral outcomes (plurality, ranked-choice, and cardinal voting choices in the Who Won the Election Task) were not able to significantly distinguish participants' voting system preferences, $\lambda = .96$, $\chi^2(6) = 9.16$, $p = .165$. The structure matrix coefficients and standardized discriminant function coefficients can be seen in Table E5.

Table E5. Study 3: Discriminant Functions for Who Won the Election Task Choices Predicting Voting System Preferences

	Standardized Discriminant Function Coefficients		Structure Matrix Coefficients	
Predictor Variables	*Function 1*	*Function 2*	*Function 1*	*Function 2*
WWET Cardinal	.51	.60	-.63	.62
WWET Plurality	-1.44	-2.81	.71	-.70
WWET RCV	2.70	2.73	.87	-.49

The overall hit rate for successfully predicting voting system preferences from choices in the

Who Won the Election Task was 34.42%. A Huberty's *I* of .02 was found, indicating a null effect size in the Who Won the Election Task's ability to predict voting system preferences (Huberty, & Lowman, 2000). The hit rate for three groups was 65.91% for plurality ($z = 4.67$, $p < .001$), 20.51% for RCV ($z = -2.35$, $p = .019$) and 31.18% for cardinal voting ($z = -.37$, $p = .711$). From the Who Won the Election Task choices, we could correctly classify participants who preferred a plurality voting system, but the predictions for participants who preferred cardinal were at chance levels, and RCV was worse than chance.

LDA by condition

Table E6. LDA Hit Rate by condition for Study 3

	Voting_System		
Condition	Cardinal	Plurality	RCV
Raw+Summary Data	41.03	51.61*	25.00
Raw Data Only	31.48	46.15	38.24

Note: *$p < .05$; **$p < .01$.

Study 4

We found that the 3 predictors of representing competing voting mechanisms for determining electoral outcomes (plurality, ranked-choice, and cardinal voting choices in the Who Won the Election Task) were able to significantly distinguish participants' voting system preferences, $\lambda = .91$, $\chi^2(6) = 14.77$, $p = .022$. The second function was not found to be significant, $\lambda = .99$, $\chi^2(2) = 2.36$, $p = .308$. Of the variance explained, 84.5% is explained by Function 1, with only 15.5% explained by Function 2. The structure matrix coefficients and standardized discriminant function coefficients can be seen in Table E7.

Table E7. Study 4: Discriminant Functions for Who Won the Election Task Choices Predicting Voting System Preferences

	Standardized Discriminant Function Coefficients		Structure Matrix Coefficients	
Predictor Variables	*Function 1*	*Function 2*	*Function 1*	*Function 2*
WWET Cardinal	.08	1.65	-.74	.67
WWET Plurality	-1.20	-1.17	.72	-.44
WWET RCV	2.11	2.27	.91	-.27

Appendix F. Learning Comprehension Results

Study 2

Note: Time 1 = After Basic Introduction & Time 2 = After teaching intervention.

Table F1. Learning Comprehension Scores Descriptive Statistics and Pairwise Comparisons

	Time 1 *M* (*SD*)			Time 2 *M* (*SD*)			Effect Size: Time 2 - Time 1 (Cohen's d)		
Teaching Intervention	*Cardinal*	*Plurality*	*RCV*	*Cardinal*	*Plurality*	*RCV*	*Cardinal*	*Plurality*	*RCV*
Retrieval Practice	2.22 (1.53)	3.51 (.83)	1.30 (1.20)	3.18 (1.32)	3.60 (.87)	1.77 (1.39)	.68	.11	.37
Re-Study	2.35 (1.58)	3.65 (.77)	1.47 (1.36)	2.92 (1.44)	3.69 (.78)	1.76 (1.40)	.38	.05	.20
Discovery Learning	2.65 (1.42)	3.69 (.77)	1.32 (1.33)	3.00 (1.44)	3.80 (.55)	1.69 (1.52)	.24	.16	.26
Control	2.06 (1.68)	3.66 (.84)	1.32 (1.33)	2.17 (1.72)	3.65 (.84)	1.40 (1.40)	.07	-.01	.05

Table F2. Learning Comprehension Scores Group Level Descriptive Statistics and Pairwise Comparisons

Teaching Intervention	Time 1 Total *M* (*SD*)	Time 2 Total *M* (*SD*)	Effect Size: Time 2 - Time 1 (Cohen's d)
Retrieval Practice	7.02 (2.39)	8.56 (2.67)	.61
Re-Study	7.47 (2.45)	8.37 (2.51)	.36
Discovery Learning	7.67 (2.27)	8.49 (2.54)	.34
Control	7.05 (2.57)	7.22 (2.84)	.07
Total	7.29 (2.43)	8.17 (2.69)	.34

Appendix G. Fairness Ratings

Study 2

Table G1. Descriptive Statistics of Raw Fairness Ratings

	Time 1 (After voting system introduction) *M* (*SD*)			Time 2 (After teaching intervention) *M* (*SD*)		
Teaching Intervention	*Cardinal*	*Plurality*	*RCV*	*Cardinal*	*Plurality*	*RCV*
Retrieval Practice	53.09 (24.10)	63.36 (22.62)	67.92 (20.6)	50.09 (22.44)	60.85 (22.56)	65.13 (18.57)
Re-Study	52.31 (23.23)	60.01 (23.97)	70.36 (21.30)	52.71 (22.43)	59.49 (22.62)	67.06 (18.60)
Discovery Learning	50.40 (24.80)	62.80 (24.00)	67.20 (21.60)	43.20 (20.50)	62.00 (23.40)	63.90 (20.60)
Control	56.03 (23.52)	58.44 (21.87)	70.03 (19.36)	58.13 (22.92)	58.06 (21.84)	68.86 (18.37)

Table G2. Descriptive Statistics and Effect Size of Absolute Difference Scores of Fairness Ratings Between Time 1 and Time 2

	Absolute Difference Scores *M* (*SD*)				Absolute Difference Scores Cohen's *d*			
Teaching Intervention	*Cardinal*	*Plurality*	*RCV*	*Avg.*	*Cardinal*	*Plurality*	*RCV*	*Avg.*
Retrieval Practice	16.50 (13.80)	12.90 (13.80)	13.80 (13.50)	14.40 (9.10)	1.20	.93	1.02	1.58
Re-Study	13.80 (13.20)	14.40 (16.00)	16.40 (14.20)	14.90 (11.10)	1.04	.90	1.15	1.34
Discovery Learning	18.90 (15.60)	14.80 (14.60)	16.00 (14.80)	16.60 (8.93)	1.21	1.02	1.08	1.86
Control	11.30 (11.80)	11.20 (11.80)	10.40 (11.00)	11.00 (8.20)	.95	.94	.95	1.34
Avg.	15.10 (13.90)	13.30 (14.10)	14.10 (13.60)	–	1.09	.94	1.04	–

Appendix H. Individual Difference Measures

H1. American Voting Questions

1. Right now most elections in the United States, for local, state, and federal office, use the Plurality voting system. Do you feel like laws should be changed to switch to either a Ranked-Choice or Cardinal voting system instead?
    A. Keep using the Plurality voting system.
    B. Change the laws to switch to using a Cardinal or Ranked Choice voting system.
    C. No preference

2. Is it possible in the United States for Candidate A to get more total votes than Candidate B, but for Candidate B to win the electoral college, so that Candidate B becomes president instead of Candidate A?
    A. Yes, this is possible and has happened
    B. Yes, it is possible but it has not happened in modern history
    C. No, this is not possible
    D. I'm not sure.

3. Thinking for a moment about the way in which the president is elected in this country, which would you prefer?
    A. Amend the Constitution or change the current system so the candidate who receives the most total votes nationwide wins the election.
    B. Keep the current system, in which the candidate who wins the most votes in the Electoral College wins the election.
    C. No preference.

    Note: Question #3 is from the 2020 Pew Research Center's American Trends Panel. https://www.pewresearch.org/wp-content/uploads/2020/03/Electoral-College-topline.pdf

4. Were you eligible to vote in the 2020 United States Presidential election?
    A. Yes
    B. No

5. Did you vote in the 2020 United States Presidential election?
    A. Yes
    B. No

6. How important is voting?
    [Likert-scale 1-5 responses: 1=Not very important, 5=Very Important]

Citation: Caddick, Z.A. & Rottman, B.M., (2020). A Brief American Voting Questionnaire

H2. Cognitive Reflection Test

Instructions: On the next few pages are several problems that vary in difficulty. Try to answer as many as you can.

1. A bat and a ball cost $1.10 in total. The bat costs a dollar more than the ball. How much does the ball cost? _____ cents. [*correct answer*: 5]

2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? _____ minutes. [*correct answer*: 5]

3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? _____ days. [*correct answer*: 47]

Scoring details: CRT items are scored by the summation of all correct answers (0-3).

Citation: Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*, 25–42.

H3. Subjective Numeracy Scale

[Part 1]
Instructions: For each of the following questions, please select the circle that best reflects how good you are at doing the following things. [Likert-scale 1-6 responses: 1=Not at all good, 6=Extremely good]

1. How good are you at working with fractions?

2. How good are you at working with percentages?

3. How good are you at calculating a 15% tip?

4. How good are you at figuring out how much a shirt will cost if it is 25% off?

[Part 2]
Instructions: For each of the following questions, please select the circle that best reflects your answer.

5. When reading the newspaper, how helpful do you find tables and graphs that are parts of a story? [Likert-scale 1-6 responses: 1=Not at all helpful, 6=Extremely helpful]

6. When people tell you the chance of something happening, do you prefer that they use words ("it rarely happens") or numbers ("there's a 1% chance")? [Likert-scale 1-6 responses: 1=Always prefer words, 6=Always prefer numbers]

7. When you hear a weather forecast, do you prefer predictions using percentages (e.g., “there will be a 20% chance of rain today”) or predictions using only words (e.g., “there is a small chance of rain today”)? [Likert-scale 1-6 responses: 1=Always prefer percentages, 6=Always prefer words]

8. How often do you find numerical information to be useful? [Likert-scale 1-6 responses: 1=Never, 6=Very often]

Scoring details: Item 7 is reverse scored. Items 1-4 comprise Subjective Numeracy Scale subscale for ability. Items 5-8 comprise Subjective Numeracy Scale subscale for preference. Scoring is calculated by averaging all items.

Citation: Fagerlin, A., Zikmund-Fisher, B.J., Ubel, P.A., Jankovic, A., Derry, H.A., & Smith, D.M. (2007). Measuring numeracy without a math test: Development of the Subjective Numeracy Scale (SNS). *Medical Decision Making*, *27*, 672-680.

H4. Oxford Utilitarianism Scale

Instructions: Indicate how much you agree or disagree with each of the following statements (1 = strongly disagree, 4 = neither agree nor disagree, 7 = strongly agree).

1. From a moral perspective, people should care about the well-being of all human beings on the planet equally; they should not favor the well-being of people who are especially close to them either physically or emotionally.

2. From a moral point of view, we should feel obliged to give one of our kidneys to a person with kidney failure since we don't need two kidneys to survive, but really only one to be healthy.

3. If the only way to save another person's life during an emergency is to sacrifice one's own leg, then one is morally required to make this sacrifice.

4. It is just as wrong to fail to help someone as it is to actively harm them yourself.

5. It is morally wrong to keep money that one doesn't really need if one can donate it to causes that provide effective help to those who will benefit a great deal.

6. It is morally right to harm an innocent person if harming them is a necessary means to helping several other innocent people.

7. If the only way to ensure the overall well-being and happiness of the people is through the use of political oppression for a short, limited period, then political oppression should be used.

8. It is permissible to torture an innocent person if this would be necessary to provide information to prevent a bomb going off that would kill hundreds of people.

9. Sometimes it is morally necessary for innocent people to die as collateral damage—if more people are saved overall.

Scoring details: Items 1-5 comprise Utilitarianism sub-scale for impartial beneficence and items 6-9 comprise sub-scale for instrumental harm. Scoring is calculated by averaging all items.

Citation: Kahane, G., Everett, J. A., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2018). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychological Review, 125*(2), 131-164.

H5. Moral Foundations Questionnaire

[Part 1]
Instructions: When you decide whether something is right or wrong, to what extent are the following considerations relevant to your thinking? Please rate each statement using this scale:

0 = not at all relevant (This consideration has nothing to do with my judgments of right and wrong)
1 = not very relevant
2 = slightly relevant
3 = somewhat relevant
4 = very relevant
5 = extremely relevant (This is one of the most important factors when I judge right and wrong)

1. Whether or not someone suffered emotionally
2. Whether or not some people were treated differently than others
3. Whether or not someone's action showed love for his or her country
4. Whether or not someone showed a lack of respect for authority
5. Whether or not someone violated standards of purity and decency
6. Whether or not someone was good at math
7. Whether or not someone cared for someone weak or vulnerable
8. Whether or not someone acted unfairly
9. Whether or not someone did something to betray his or her group
10. Whether or not someone conformed to the traditions of society
11. Whether or not someone did something disgusting
12. Whether or not someone was cruel
13. Whether or not someone was denied his or her rights
14. Whether or not someone showed a lack of loyalty
15. Whether or not an action caused chaos or disorder
16. Whether or not someone acted in a way that God would approve of

[Part 2]
Instructions: Please read the following sentences and indicate your agreement or disagreement:

1 = Strongly disagree
2 = Slightly disagree
3 = Slightly disagree
4 = Slightly agree
5 = Moderately agree
6 = Strongly agree

17. Compassion for those who are suffering is the most crucial virtue.
18. When the government makes laws, the number one principle should be ensuring that everyone is treated fairly.

19. I am proud of my country's history.
20. Respect for authority is something all children need to learn.
21. People should not do things that are disgusting, even if no one is harmed.
22. It is better to do good than to do bad.
23. One of the worst things a person could do is hurt a defenseless animal.
24. Justice is the most important requirement for a society.
25. People should be loyal to their family members, even when they have done something wrong.
26. Men and women each have different roles to play in society.
27. I would call some acts wrong on the grounds that they are unnatural.
28. It can never be right to kill a human being.
29. I think it's morally wrong that rich children inherit a lot of money while poor children inherit nothing.
30. It is more important to be a team player than to express oneself.
31. If I were a soldier and disagreed with my commanding officer's orders, I would obey anyway because that is my duty.
32. Chastity is an important and valuable virtue.

Scoring details: Items belong to one of 5 sub-scales, where the score for each sub-scale is a summation. Sub-scale scoring is the summation of individual-item responses. Item/sub-scale pairings:

Harm/Care: 1, 7, 12, 17, 23, 28
Fairness/Reciprocity: 2, 8, 13, 18, 24, 29
In-group/Loyalty: 3, 9, 14, 19, 25, 30
Authority/Respect: 4, 10, 15, 20, 26, 31
Purity/Sanctity: 5, 11, 16, 21, 27, 32

Citation: Graham, Haidt, & Nosek, (2008). Moral Foundations Questionnaire: 30-Item Full Version. www.moralfoundations.org

H6. Need for Cognitive Closure

Instructions: Read each of the following statements and decide how much you agree with each according to your beliefs and experiences.

1=Strongly disagree, 2=Moderately disagree, 3=Slightly disagree, 4=Slightly agree, 5= Moderately agree, 6=Strongly agree

---

1. We think that having clear rules and order at work is essential for success.
2. Even after I've made up my mind about something, We am always eager to consider a different opinion.
3. We don't like situations that are uncertain.
4. We dislike questions which could be answered in many different ways.
5. We like to have friends who are unpredictable.
6. We find that a well ordered life with regular hours suits my temperament.
7. We enjoy the uncertainty of going into a new situation without knowing what might happen.
8. When dining out, We like to go to places where We have been before so that We know what to expect.
9. We feel uncomfortable when We don't understand the reason why an event occurred in my life.
10. We feel irritated when one person disagrees with what everyone else in a group believes.
11. We hate to change my plans at the last minute.
12. We would describe myself as indecisive.
13. When We go shopping, We have difficulty deciding exactly what it is We want.
14. When faced with a problem We usually see the one best solution very quickly
15. When We am confused about an important issue, We feel very upset.
16. We tend to put off making important decisions until the last possible moment.
17. We usually make important decisions quickly and confidently.
18. We have never been late for an appointment or work.
19. We think it is fun to change my plans at the last moment.
20. My personal space is usually messy and disorganized.
21. In most social conflicts, We can easily see which side is right and which is wrong.
22. We have never known someone We did not like.
23. We tend to struggle with most decisions.
24. We believe orderliness and organization are among the most important characteristics of a good student.
25. When considering most conflict situations, We can usually see how both sides could be right.
26. We don't like to be with people who are capable of unexpected actions.
27. We prefer to socialize with familiar friends because We know what to expect from them.
28. We think that We would learn best in a class that lacks clearly stated objectives and requirements.
29. When thinking about a problem, We consider as many different opinions on the issue as possible.
30. We don't like to go into a situation without knowing what We can expect from it.
31. We like to know what people are thinking all the time.
32. We dislike it when a person's statement could mean many different things.

33. It's annoying to listen to someone who cannot seem to make up his or her mind.
34. We find that establishing a consistent routine enables me to enjoy life more.
35. We enjoy having a clear and structured mode of life.
36. We prefer interacting with people whose opinions are very different from my own.
37. We like to have a plan for everything and a place for everything.
38. We feel uncomfortable when someone's meaning or intention is unclear to me.
39. We believe that one should never engage in leisure activities.
40. When trying to solve a problem We often see so many possible options that it's confusing.
41. We always see many possible solutions to problems We face.
42. I'd rather know bad news than stay in a state of uncertainty.
43. We feel that there is no such thing as an honest mistake.
44. We do not usually consult many different options before forming my own view.
45. We dislike unpredictable situations.
46. We have never hurt another person's feelings.
47. We dislike the routine aspects of my work (studies).

Scoring details:

1. Reverse-score items 2, 5, 7, 12, 13, 16, 19, 20, 23, 25, 28, 29, 36, 40, 41, and 47.
2. Sum items 18, 22, 39, 43, and 46 to form a lie score.
3. Remove the subject if the lie score is greater than 15.
4. Sum all items except for the above listed lie items to calculate the need for closure score.
5. Use the top and bottom quartiles to determine high and low need for closure subjects.
6. If factors are required, use the following scoring system:
   Order: 1, 6, 11, 20, 24, 28, 34, 35, 37, 47
   Predictability: 5, 7, 8, 19, 26, 27, 30, 45
   Decisiveness: 12, 13, 14, 16, 17, 23, 40
   Ambiguity: 3, 9, 15, 21, 31, 32, 33, 38, 42
   Closed Mindedness: 2, 4, 10, 25, 29, 36, 41, 44

Citation: Kruglanski, A. W., Atash, M. N., De Grada, E., Mannetti, L., & Pierro, A. (2013). Need for Closure Scale (NFC). Measurement instrument database for the social science. *Journal of Personality and Social Psychology*, *73*, 1005-1016.

H7. Political Orientation Questions

1. Generally speaking, do you usually think of yourself as a Democrat, Republican, an Independent, or what?
    A. Democrat
    B. Republican
    C. Independent
    D. Other [fill in the blank]
    E. No Preference

[if either Democrat or Republican is selected for item 1, item 2 is answered (otherwise it is skipped)]

2. Would you consider yourself a strong [Republican/Democrat]?
    A. Strong
    B. Not very strong

[if either Independent, Other, or No Preference is selected for item 1, item 3 is answered (otherwise it is skipped)]

3. Do you think of yourself as closer to the Republican Party or the Democratic Party?
    A. Closer to Republican
    B. Neither
    C. Closer to Democratic

4. We hear a lot of talk these days about liberals and conservatives. Here is a seven-point scale on which the political views that people might hold are arranged from extremely liberal to extremely conservative. Where would you place yourself on this scale?
    1. Extremely liberal
    2. Liberal
    3. Slightly liberal
    4. Moderate; middle of the road
    5. Slightly conservative
    6. Conservative
    7. Extremely conservative

Appendix I. Metacognition Results

Scatterplots were created showing the relationship between participants' judgment of knowledge ratings for each voting system and corresponding comprehension scores. Studies 3 and 4 were broken down by condition assignment.

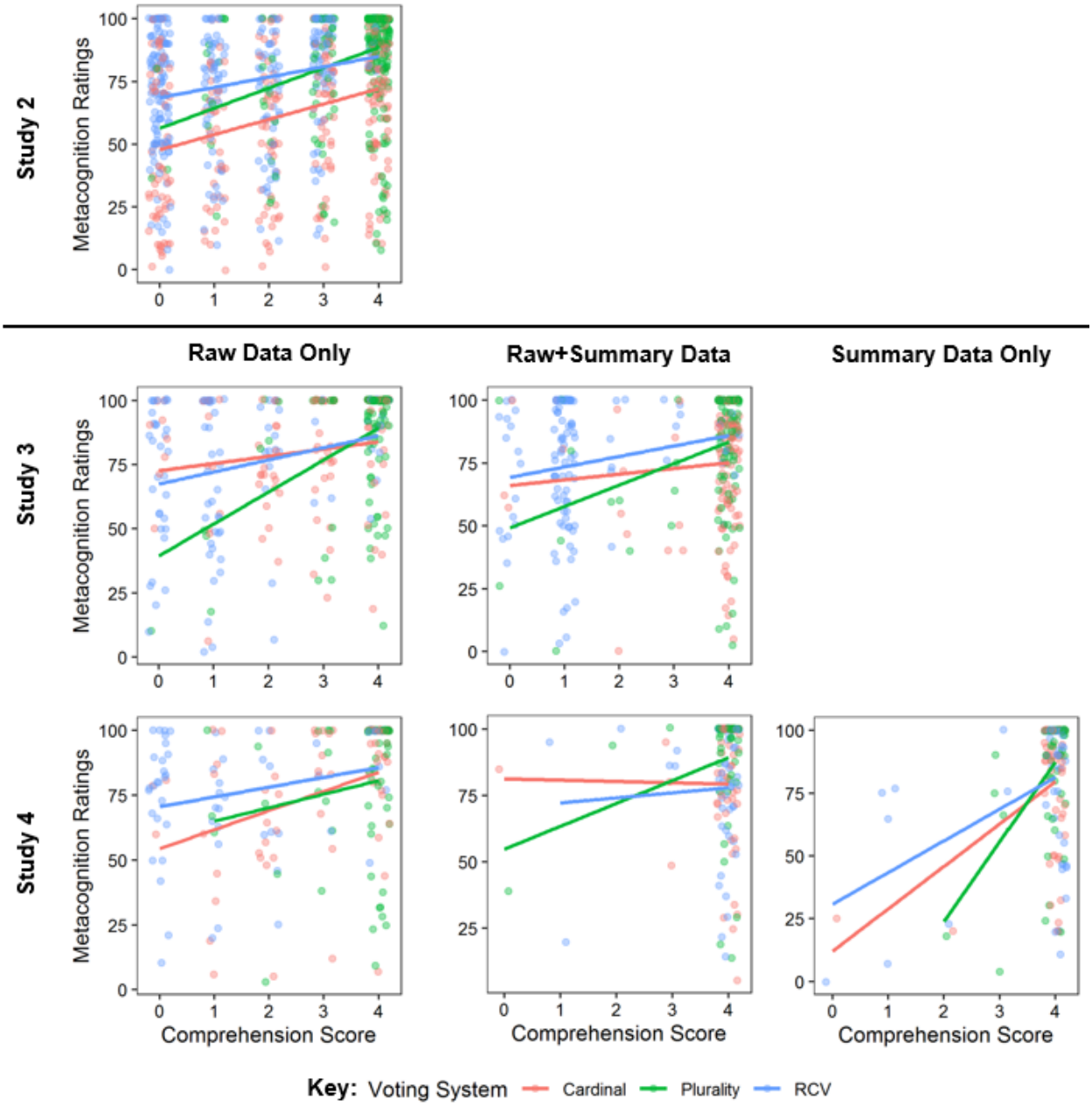


Figure I1. Scatterplot of Judgment of Knowledge Ratings and Comprehension Scores. Note: A best fit line was added to detail the nature of the relationship between the two variables.

Bibliography

Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., ... & Tuan, H. L. (2004). Inquiry in science education: International perspectives. *Science Education*, *88*(3), 397-419.

Adamek, R. J. (1994). A review: Public opinion and Roe v. Wade: Measurement difficulties. *The Public Opinion Quarterly*, *58*(3), 409-418.

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research*, *87*(3), 659-701.

Alexander, P. A., Fives, H., Buehl, M. M., & Mulhern, J. (2002). Teaching as persuasion. *Teaching and Teacher Education*, *18*(7), 795-813.

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery based instruction enhance learning? *Journal of Educational Psychology*, *103*, 1–18.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in cognitive sciences*, *15*(3), 122-131.

Arrow, K. J. (1951). Social Choice and Individual Values.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*(3), 379–384.

Bear, A., & Rand, D. G. (2016). Intuition, deliberation, and the evolution of cooperation. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(4), 936–941.

Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2008). How are preferences revealed?. *Journal of Public Economics*, *92*(8-9), 1787-1794.

Bianco, W. T., Lynch, M. S., Miller, G. J., & Sened, I. (2006). "A Theory Waiting to Be Discovered and Used": A Reanalysis of Canonical Experiments on Majority-Rule Decision Making. *The Journal of Politics*, *68*(4), 838-851.

Blais, A., Plescia, C., & Sevi, S. (2021). Choosing to vote as usual. *Available at SSRN 3784822*.

Blake, A. (2021, September 2). *Are Republicans catching the car on abortion?* Washington Post. https://www.washingtonpost.com/politics/2021/09/02/republicans-abortion-texas-law/

Bond, R., & Messing, S. (2015). Quantifying Social Media's Political Space: Estimating Ideology from Publicly Revealed Preferences on Facebook. *American Political Science*

Review, *109*(1), 62–78. https://doi.org/10.1017/S0003055414000525

Boston Globe Editorial Board (2020, November 4). A voting setback: Massachusetts rejects ranked choice. The Boston Globe. https://www.bostonglobe.com/2020/11/04/opinion/voting-setback-massachusetts-rejects-ranked-choice/

Brenan, M. (2020). Amid pandemic, confidence in key U.S. institutions surges. https://news.gallup.com/poll/317135/amid-pandemic-confidence-key-institutions-surges.aspx

Carey, J., Helmke, G., Nyhan, B., Sanders, M., Stokes, S., & Yamaya, S. (2020). The Effect of Electoral Inversions on Democratic Legitimacy: Evidence from the United States. [pre-print]

Chater, N., & Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science?. *Trends in Cognitive Sciences*, *7*(1), 19-22.

Chu, Y. H., Huang, K. P., Lagos, M., & Mattes, R. (2020). A lost decade for third-wave democracies?. *Journal of Democracy*, *31*(2), 166-181.

Clay, Z., & Zuberbühler, K. (2009). Food-associated calling sequences in bonobos. *Animal Behaviour*, *77*(6), 1387-1396.

Condorcet, N. (1785). Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix.

Cook, E. A., Jelen, T. G., & Wilcox, C. (1993). Measuring public attitudes on abortion: methodological and substantive considerations. *Family Planning Perspectives*, 118-145.

Dakin, S. C. (2001). Information limit on the spatial integration of local orientation signals. *JOSA A*, *18*(5), 1016-1026.

De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of educational research*, *68*(2), 179-201.

Donovan, T., Tolbert, C., & Gracey, K. (2019). Self-reported understanding of ranked-choice voting. *Social Science Quarterly*, *100*(5), 1768-1776.

Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one's own ignorance. In *Advances in experimental social psychology* (Vol. 44, pp. 247-296). Academic Press.

Duverger, M. (1959). *Political parties: Their organization and activity in the modern state*. Metheun & Co. Ltd..

Fagerlin, A., Zikmund-Fisher, B.J., Ubel, P.A., Jankovic, A., Derry, H.A., & Smith, D.M. (2007). Measuring numeracy without a math test: Development of the Subjective Numeracy Scale (SNS). *Medical Decision Making*, *27*, 672–680.

FairVote. (2022). *Electoral Systems*. https://fairvote.org/resources/electoral-systems/#electoral_systems_rcv_vs_approval_voting

FairVote (2022). Ranked Choice Voting Information. https://fairvote.org/our-reforms/ranked-choice-voting-information/

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, *34*(10), 906–911.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*, 25–42.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, *82*, 300–329.

Gallup, Inc. (2022, October 25). *Party Affiliation | Gallup Historical Trends*. Gallup.com. https://news.gallup.com/poll/15370/Party-Affiliation.aspx

Gibbard, Allan (1973). Manipulation of voting schemes: A general result. *Econometrica*. *41*(4), 587–601. doi:10.2307/1914083

Graham, Haidt, & Nosek, (2008). Moral Foundations Questionnaire: 30-Item Full Version. www.moralfoundations.org

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), 814.

Haidt, J., Graham, J., & Joseph, C. (2009). Above and below left-right: Ideological narratives and moral foundations. *Psychological Inquiry*, *20*(2–3), 110–119.

Hallsson, B. G., Siebner, H. R., & Hulme, O. J. (2018). Fairness, fast and slow: A review of dual process models of fairness. *Neuroscience and Biobehavioral Reviews*, *89*(June 2017), 49–60.

Herz, Nathaniel. (2022, September 1). *After Peltola win in Alaska, a debate erupts over ranked-choice voting*. The Washington Post. https://www.washingtonpost.com/politics/2022/09/01/peltola-palin-ranked-choice-voting-alaska/

Huberty, C. J., & Lowman, L. L. (2000). Group overlap as a basis for effect size. *Educational and Psychological Measurement*, *60*(4), 543–563.

Jones, J. M. (2021, November 20). *Support for third U.S. political party at high point*. Gallup.com. Retrieved November 5, 2022, from https://news.gallup.com/poll/329639/support-third-political-party-high-point.aspx

Kahane, G., Everett, J. A., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2018). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychological Review*, *125*(2), 131–164.

Kaiser Family Foundation. (2019, August 22). *KFF Poll: Public Opinion and Knowledge on Reproductive Health Policy*. KFF. https://www.kff.org/womens-health-policy/poll-finding/kff-poll-public-opinion-and-knowledge-on-reproductive-health-policy/

Kameda, T., Toyokawa, W., & Tindale, R. S. (2022). Information aggregation and collective intelligence beyond the wisdom of crowds. *Nature Reviews Psychology*, 1-13.

Kimball, D., & Anthony, J. (2021). Public Perceptions of Alternative Voting Systems: Results from a National Survey Experiment. *Available at SSRN 3854047*.

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology*, *53*(4), 636.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480.

Lachenbruch, P. A., & Mickey, M. R. (1968). Estimation of error rates in discriminant analysis. *Technometrics*, *10*, 1–11.

Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of educational research*, *86*(3), 681-718.

List, C. (2013, December 18). Social Choice Theory (Stanford Encyclopedia of Philosophy). Stanford Encyclopedia of Philosophy. Retrieved July 8, 2022, from https://plato.stanford.edu/entries/social-choice/

Morgan, M., & Glennerster, A. (1991). Efficiency of locating centres of dot-clusters by human observers. *Vision Research*, *31*(12), 2075–2083.

Moser, A.. (2022, August 30). Polling Shows Alaskan Voters Understand Ranked Choice Voting. *Alaskans for Better Elections*. Retrieved November 13, 2022, from https://alaskansforbetterelections.com/polling-shows-alaskan-voters-understand-ranked-choice-voting/

Mr. Beat. (2019, October 18). *Why the Way We Vote Is Terrible*. YouTube. https://www.youtube.com/watch?v=7mYPACssLbI

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., … Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, *14*, 47–61.

Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2015). On the reception and detection of pseudo-profound bullshit. *Judgment and Decision making*, *10*(6),

549-563.

Pennycook, G., Cheyne, J. A., Seli, P., Koehler, D. J., & Fugelsang, J. A. (2012). Analytic cognitive style predicts religious and paranormal belief. *Cognition*, *123*(3), 335-346.

Pennycook, G., & Rand, D. G. (2019). Cognitive reflection and the 2016 US Presidential election. *Personality and Social Psychology Bulletin*, *45*(2), 224-239.

Pyc, M. A., & Rawson, K. A. (2011). Costs and benefits of dropout schedules of test-restudy practice: Implications for student learning. *Applied Cognitive Psychology*, *25*, 87–95.

Rangevoting.org. (n.d.). Voting: The wisdom of bees. Retrieved June 13, 2022, from https://www.rangevoting.org/ApisMellifera.html

Redlawsk, D. P. (2002). Hot cognition or cool consideration? Testing the effects of motivated reasoning on political decision making. *Journal of Politics*, *64*(4), 1021-1044.

Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, *26*(5), 521-562.

Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychological bulletin*, *140*(6), 1432–1463.

Satterthwaite, Mark Allen (1975). Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, *10*(2), 187–217. doi:10.1016/0022-0531(75)90050-2.

Seeley, T. D. (2009). *The wisdom of the hive: the social physiology of honey bee colonies*. Harvard University Press.

Seeley, T. D., Camazine, S., & Sneyd, J. (1991). Collective decision-making in honey bees: how colonies choose among nectar sources. *Behavioral Ecology and Sociobiology*, *28*(4), 277-290.

Sen, A. (2018). *Collective choice and social welfare*. Harvard University Press.

Schofield, N., Grofman, B., & Feld, S. L. (1988). The core and the stability of group choice in spatial voting games. *American Political Science Review*, *82*(1), 195-211.

Shaw, A., Choshen-Hillel, S., & Caruso, E. M. (2016). The development of inequity aversion: Understanding when (and why) people give others the bigger piece of the pie. *Psychological Science*, *27*(10), 1352-1359.

Svinicki, M. D. (1998). A theoretical foundation for discovery learning. *Advances in Physiology Education*, *275*(6), S4.

Szafir, D. A., Haroz, S., Gleicher, M., & Franconeri, S. (2016). Four types of ensemble coding in data visualizations. *Journal of vision*, *16*(5):11, 1-19.

Toplak, M. E., West, R.F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, *39*(7),1275–1289

The Center for Election Science. (2021, September 23). *Approval voting*. https://electionscience.org/library/approval-voting/

Undem, T. (n.d.). *Why so many polls get American attitudes on abortion wrong*. Vox. Retrieved November 6, 2022, from https://www.vox.com/a/abortion-decision-statistics-opinions/abortion-polling-mistakes

Van Der Straeten, K., Sauger, N., Laslier, J. F., & Blais, A. (2013). Sorting out mechanical and psychological effects in candidate elections: An appraisal with experimental data. *British Journal of Political Science*, *43*(4), 937-944.

Wike, R., & Schumacher, S. (2020a). 2. Attitudes toward elected officials, voting and the state. https://www.pewresearch.org/global/2020/02/27/satisfaction-with-democracy/

Wike, R., & Schumacher, S. (2020b). 3. Satisfaction with democracy. https://www.pewresearch.org/global/2020/02/27/satisfaction-with-democracy/

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, *147*(4), 399–435.