Verdicts, Elections, and Counterterrorism: When Groups Take Unofficial Votes

Journal:	*Academy of Management Discoveries*
Manuscript ID	AMD-2021-0099.R3
Manuscript Type:	Discoveries-through-Prose
Keywords:	Group Decision-making & Dynamics < Decision Making and Behavioral Economics, Judgment < Decision Making and Behavioral Economics, Participation < Interpersonal & Team Processes

Abstract: When groups make decisions, they often take preliminary, unofficial votes (straw polls) to see which option the members initially prefer. Does their method of voting matter? We ran an experiment where 93 groups simulated counterterrorism pursuit teams and had to determine which of three suspects represented the greatest threat. Groups that used multivoting in their unofficial votes—where members had multiple votes to distribute across options—were 50% more likely to choose the correct suspect than groups that used plurality or ranked-choice voting. This was because the members of the multivoting groups were more likely to identify the correct suspect even before their team discussion. In this article, we look at various voting methods used for presidential elections, Hugo Award winners, and American Idol contestants. We show how different voting methods can lead to entirely different results and provide initial evidence that multivoting may be a superior method for group decision-making.

Discoveries-through-Prose

Verdicts, Elections, and Counterterrorism: When Teams Take Unofficial Votes

Michael D. Johnson
University of Washington
544 Paccar Hall
Seattle, WA 98195
USA
(206) 616-2756
mdj3@uw.edu

Eli Awtrey
University of Cincinnati
3352 Carl H. Lindner Hall
2906 Woodside Drive
Cincinnati, OH 45221
USA
(513) 556-4582
awtreyei@ucmail.uc.edu

Wei Jee Ong
National University of Singapore
15 Kent Ridge Drive, #08-39
Singapore 119245
Singapore
+65 66015768
weijee@nus.edu.sg

VERDICTS, ELECTIONS, AND COUNTERTERRORISM: WHEN TEAMS TAKE UNOFFICIAL VOTES

ABSTRACT

When groups make decisions, they often take preliminary, unofficial votes (straw polls) to see which option the members initially prefer. Does their method of voting matter? We ran an experiment where 93 groups simulated counterterrorism pursuit teams and had to determine which of three suspects represented the greatest threat. Groups that used multivoting in their unofficial votes—where members had multiple votes to distribute across options—were 50% more likely to choose the correct suspect than groups that used plurality or ranked-choice voting. This was because the members of the multivoting groups were more likely to identify the correct suspect even before their team discussion. In this article, we look at various voting methods used for presidential elections, Hugo Award winners, and *American Idol* contestants. We show how different voting methods can lead to entirely different results and provide initial evidence that multivoting may be a superior method for group decision-making.

Keywords: Teams; group decision-making; voting; social decision schemes; social choice

"Eleven guilty, one not guilty. Well, now we know where we are."
—Jury foreman, *12 Angry Men*

In the opening scenes of the film *12 Angry Men*, a jury retires for deliberation on a charge of premeditated murder, which under their sentencing guidelines would mean the death penalty. The jury foreman's initial move is to take an unofficial vote of the jury. He begins by asking who believes the defendant is guilty, and Jurors 1 through 7 and 9 through 12 all raise their hands—although some of them do so hesitantly, waiting to see how others vote. Only Juror 8 votes not guilty.

American juries require a unanimous verdict, and this jury does not have it. As they discuss the case, Juror 8 says that he doesn't know whether the defendant is guilty but thinks they should discuss it because a man's life is on the line. Juror 8's initial dissenting vote, though, ultimately reverses the opinions of the other jury members, and in the end, they vote to acquit.

Taking preliminary unofficial votes, especially among juries, is pretty common; one study found that two-thirds of all juries take unofficial votes within the first thirty minutes of their deliberations (Devine et al., 2004). But do they help? Can taking an unofficial vote, like in *12 Angry Men*, facilitate good group decision-making?

It’s complicated. Although fictionalized dramas might have us believe that “a show of hands” is advisable, *12 Angry Men* obscures the range of options available to groups as they signal their initial perspectives. To explore the nuances surrounding unofficial votes and their outcomes on group decision making, we conducted an experiment with different types of preliminary unofficial votes. One is plurality voting, where voters can only choose one option. Plurality voting is used in the vast majority of political elections and is commonly used outside of politics as well. Another method is ranked-choice voting, where voters indicate their preferences from best to worst. Growing in popularity, ranked-choice voting is used to determine the winner in some political elections and the occasional non-political contest. Multivoting is where voters are given multiple votes that they can allocate across options. It is the least common of all.

Maybe multivoting should be more common, though. Why? Because we found that when groups used multivoting, their decisions were far better than groups that used plurality or ranked-choice voting. And the reason that their decisions were so much better was that the members made better guesses at the right answer before the group discussed the issue. The story of what led to this discovery includes Hugo Award science fiction writers, *American Idol* contestants, and United Nations Secretary-General elections. It also involves a 200-year-old French manuscript, counterterrorism, and a disputed presidential election.

SOCIAL DECISION SCHEMES AND SOCIAL CHOICE THEORIES

Researchers from two disciplines—psychology and economics—pioneered the modern investigation into how groups translate individual member preferences into overall group decisions. Leading the psychologists was Jim Davis of the University of Illinois, who developed a theory called *social decision schemes* (Davis, 1973). Davis and the other psychologists—including Norbert Kerr, Pat Laughlin, Garold Stasser, and others—conducted study after study on how groups make decisions (Stasser, 1999). They found that they could fairly accurately predict jury verdicts and other group decisions on issues as far-ranging as how to control pollution to how to resolve the Rwandan civil war.

Different rules, or "social decision schemes," governed group decisions in different situations. One social decision scheme is *truth wins,* meaning that the group made the right decision if at least one member's opinion was correct. Another is *majority wins*, meaning that whatever opinion was shared by the most members ended up being the group's decision. They also discovered the *leniency bias*, which is the tendency of juries to acquit when jurors are split in their opinions. The groups were murmurations of birds, and the psychologists could predict which way the flocks would turn.

The economists pursued their own theory of group decision-making called *social choice theory* (Gaertner, 2009). It started a few decades before the psychologists began their work, when a Scottish economist named Duncan Black discovered a 200-year-old manuscript in the dusty archives of a London library (Black, 1948). The manuscript was written by the 18th century French mathematician and aristocrat Marie Jean Antoine Nicolas de Caritat, the Marquis of Condorcet. It described a specific method for counting ranked-choice votes that appeared to be superior to any method in use in the 20th century. Condorcet discovered that by pitting each option against all the other options in ranked-choice voting, he could most accurately predict which one the group truly preferred (McLean, 1990).

Black was profoundly inspired by Condorcet. Like Davis's psychological theory of social decision schemes, Black's economic theory of social choice was about how to predict group decisions from knowing individual member preferences. But these two theories operated on parallel tracks for years. As noted by Laughlin (2011, p. 64), "social choice theory and social decision scheme theory address the same fundamental issues, but there has been relatively little contact between them."

Social decision schemes and social choice theory differed in important ways. Social choice theory was less focused on jury verdicts and more on political elections, with an eye toward determining the true preference of an electorate: Which candidate do the voters really prefer? A second difference between the two theories was in how they were tested. Black, and the economists that followed him, did not conduct studies like Davis; they developed mathematical proofs. The psychologists and the economists differed in a third way as well: how people voted. Unlike the simple guilty/not guilty votes of Davis's jurors, the economists had more complex ways that people could vote.

We took insights from both perspectives in our research. We conducted an experiment with real participants like the psychologists did. And we used complex voting methods like the economists did. But we did it in a different context: counterterrorism pursuit teams.

COUNTERTERRORISM PURSUIT TEAMS

After the 9/11 attacks, the United States ramped up its efforts to combat terrorism. The Department of Homeland Security was formed to coordinate the efforts of various intelligence agencies. In 2010, the National Counterterrorism Center announced that it was putting together "pursuit teams," whose sole purpose would be to connect the dots of intelligence across agencies and run down potential threats (Schmitt, 2010).

We simulated these pursuit teams in our laboratory by creating groups of undergraduate students whose job was to combine intelligence in a task called "Threat Target" (Thompson & Flammang, 2014; see the Appendix for more details about the task). Groups participating in this task are given intelligence about three terrorism suspects, and they need to decide which one represents the greatest threat. The intelligence is distributed across the group members. No one has all the information about any one suspect, so the group members must share intelligence to identify the most threatening one. The CIA had previously vetted the task and identified one suspect who was clearly more of a threat than the others. To confirm this assessment, we asked a retired British Naval Intelligence officer to vet it as well. After subjecting the task to the same techniques he had used in tracking down war criminals in the Balkans, he agreed with the CIA's assessment on which suspect was the greatest threat.

Like real pursuit teams where the members have intelligence from different sources, each of the members of our Threat Target groups received different information from the others. Then the groups took a preliminary, unofficial vote to see what each member thought about the suspects. But we manipulated the way they voted using three different voting methods. In one condition, group members voted for just one suspect (plurality voting). In another condition, group members made a rank order of the suspects from most to least threatening (ranked-choice voting). In the third condition, the group members distributed ten votes across the three suspects in any way they wished (multivoting).

The groups then looked at the results of their unofficial votes and discussed the suspects. If they combined their information well, they could identify the one suspect who was clearly more of a threat than the others. Then they returned a verdict. (See the Appendix for detailed information about how we conducted the study.)

VOTING METHODS

Imagine you are the leader of a counterterrorism pursuit team. The team is comprised of members from the Central Intelligence Agency, the Federal Bureau of Investigation, the National Security Agency, the Department of Justice, and intelligence agencies from three branches of the armed forces. Your current task is to choose which of two terrorism suspects to pursue, codenamed Jackal and Badger.

Before you discuss the unique intelligence from the different agencies, you take an unofficial vote of the seven team members to see which suspect they initially think is most threatening. The members from Air Force, Army, and Navy Intelligence, along with the member from the NSA, vote for Badger. The members from the CIA, FBI, and Department of Justice vote for Jackal. In this vote, your team thinks that Badger is more threatening than Jackal by a count of four to three.

INSERT TABLE 1 HERE

As the team discusses the intelligence they have on these suspects, though, the NSA agent says, “I know we are just talking about these two suspects, but we really should consider Eagle as well. The NSA has strong evidence about Eagle.”

You trust this NSA agent, and take another unofficial vote, including Eagle as a third suspect. The representatives from the NSA and Air Force Intelligence, who originally voted for Badger, shift their votes to Eagle. Every other team member sticks with their original vote. Now the vote results in three blocs (Bloc 1: Army and Navy Intelligence; Bloc 2: CIA, FBI, and DOJ; Bloc 3: Air Force Intelligence and NSA). The votes are three for Jackal (Bloc 2), two for Badger (Bloc 1), and two for Eagle (Bloc 3). By including Eagle as a third option, your unofficial vote now implies that Jackal is more threatening than Badger, even though moments ago the first vote showed the opposite to be true!

INSERT TABLE 2 HERE

Plurality voting

This phenomenon is known as the spoiler effect, and it can occur when *plurality voting* is used with more than two options (Gaertner, 2009). Spoiler candidates siphon votes away from leading candidates, and can change the group's apparent preference, as Eagle did in the example above. Because plurality voting allows people to vote for just one option, it forces voters to "go all in," like a poker player making a big bet.

Plurality voting does not require a majority vote for one option to win over another. In the 2000 US presidential election, neither George W. Bush nor Al Gore won over 50% of the votes in the state of Florida. In fact, George W. Bush (48.847%) won Florida by just 537 more votes than Al Gore (48.838%), and in that election, third-party candidate Ralph Nader won almost 100,000 votes in that state. Had he not been in the election as a spoiler candidate, it is likely that enough of his supporters would have voted for Gore that the outcome of the entire election would have changed.

One-third of the teams in our Threat Target experiment used plurality voting in their preliminary, unofficial votes. Remember that their unofficial vote was taken prior to their discussion, so it only *informed* their final verdict; it did not *determine* it. When they returned their final verdict, only 31% chose the suspect that the CIA had identified as the most threatening. This is about as much as one would expect by chance alone ($\frac{1}{3}$, as there were three suspects).

To figure out why plurality voting teams did so poorly, we looked at the results of their unofficial votes. This would tell whether the teams were incorrect before their discussions, and if so, by how much. It was possible, in fact, that in these teams the members had initially identified the correct suspect and then switched to an incorrect one following their discussions. But in only 6% of these teams did a majority of the team members identify the correct suspect. Chance alone would lead to approximately 11% of teams having a majority of their members choose the correct suspect in their unofficial vote.[1] These teams started worse than chance and ended about as good as chance.

This is bad news for groups that use plurality voting. Making team members "go all in" on one choice with plurality voting was about as good as deciding based on the flip of a coin.

Ranked-choice voting

To fix the problems associated with "go all in" plurality voting, the economists investigated ranked-choice voting, where voters indicate their order of preference for each option. This is how the Oscars are determined, with each member of the Academy ranking their choices from best to worst. Ric Robertson, the former COO of the Academy of Motion Picture Arts and Sciences, said about the voting process, "If you do it by plurality, you could have a winner with just 11 percent of the vote" (Seelye, 2018), and that ranked-choice voting prevents this kind of result from occurring.

The World Science Fiction Society (WSFS) also uses ranked-choice voting when handing out its annual awards. They give the Hugo Awards to the people who have published the best science fiction or fantasy work of the year. The WSFS invites their members to nominate people for the awards. Using a process called "E Pluribus Hugo," the WSFS creates a slate of the top six nominees as the final ballot in each award category. Voters then rank the final six nominees in each category from one to six, best to worst. E Pluribus Hugo was created because

---

[1] The probability of a 4-person team choosing the correct suspect in their final decision is not the same as having a majority of individual members choose the correct suspect. The probability of a team choosing the correct suspect in their final decision is $\frac{1}{3}$, as there are three suspects. The probability of a majority of members choosing the correct suspect in the unofficial vote, however, is the sum of the probability of all four members choosing the correct suspect ($\frac{1}{3}^4 = 1.23\%$) plus the probability of exactly three members choosing the correct suspect ($\left(\frac{1}{3}^3 \times \frac{2}{3}\right) \times 4 = 9.88\%$), or 11.11%.

two factions of voters—called "Sad Puppies" and "Rabid Puppies"—coordinated their votes to control the winners of the awards under the old plurality voting system.[2]

Back to your pursuit team. Imagine that now, like the WSFS, you decide to use ranked-choice voting, where each team member ranks the three suspects from most to least threatening. Now the vote results in three blocs (Bloc 1: Army and Navy Intelligence; Bloc 2: CIA, FBI, and DOJ; Bloc 3: Air Force Intelligence and NSA).

INSERT TABLE 3 HERE

This is where Condorcet's method from his lost and found manuscript can help you determine whom your team thinks is the greatest threat. Condorcet showed that a group's preference in ranked-choice voting could be determined by running a series of head-to-head contests between each option (Gaertner, 2009). With your team's unofficial vote, you do this three times: you compare Badger to Jackal, Jackal to Eagle, and Eagle to Badger. In this method, it doesn't matter what a voter ranks first; it only matters that they rank one option over another. For instance, if a team member ranks Badger second and Jackal third, they think that Badger is more threatening than Jackal.

And in fact, your team does think Badger is more threatening than Jackal by a vote of 4-3. They also think that Eagle is more threatening than Jackal by 4-3, and they think Badger is more threatening than Eagle by 5-2. Using Condorcet's method, the team as whole ranks the suspects as Badger > Eagle > Jackal. The unofficial vote using plurality voting for these three suspects showed Jackal to be their *first* choice, but under ranked-choice voting Jackal is their *last* choice!

[2] Both factions espoused right-wing, anti-diversity agendas, and had accused the WSFS of favoring authors who weren't straight, white males. The name "Sad Puppies" was a joke in bad taste, based on an animal welfare ad featuring musician Sarah McLachlan. The faction's name suggested the puppies in the ad were sad not because they had been abused, but because the WSFS awards were going to undeserving authors. "Rabid Puppies" was a more militant faction that split off "Sad Puppies."

The Nobel-prizewinning economists Eric Maskin and Amartya Sen (2016) showed that if Condorcet's method had been used in the Republican primary leading up to the 2016 U.S. election, Donald Trump would not have become President of the United States. Examining the exit polls, they found that people voted generally in three blocs.

INSERT TABLE 4 HERE

Ted Cruz and John Kasich were both preferred over Trump (60%–40%), and Kasich was preferred over Cruz (65–35%). With ranked-choice voting, John Kasich would have won the nomination; Trump only won because of the spoiler effect. To prevent a spoiler from winning again, the state of Maine used ranked-choice voting in its presidential primaries in 2020 (Astor, 2019). Other states, and several countries around the world, are now using ranked-choice voting in various elections for the same reason.

One-third of the groups in our Threat Target experiment used ranked-choice voting in their unofficial vote. But they did no better than the groups that used plurality voting. Only 32% of them identified the correct suspect in their final verdict. Like the groups that voted using the plurality method, this was about as good as you would expect by chance. Bad news for ranked-choice voting, too.

As with the plurality voting groups, we looked at the results of the unofficial votes of the ranked-choice voting groups, to see whether they were more or less likely to identify the correct suspect before their discussions. Similar to the plurality voting groups, only 7% of these groups—worse than what would be expected by chance—had a majority of their members rank the correct suspect as the most threatening. In our experiment, ranked-choice voting showed no advantage over plurality voting: the groups started worse than chance and ended as good as chance. This was a surprising result, considering the advantages of ranked-choice voting over plurality voting. We suspect that this was because the advantage of ranked-choice voting is likely stronger in situations where the members are expressing their preferences (as in political voting), rather than when they are trying to find an objectively correct answer.

Multivoting

A third but less common voting method is gaining favor in many companies and also in reality TV shows. When voting for company directors, shareholders often receive a number of votes corresponding to the number of shares they hold and may split their votes across candidates in any combination. Similarly, the later rounds of voting on the show *American Idol* are determined by fan vote. Each fan has ten votes that they can allocate across multiple contestants. They can place all ten of their votes on their favorite or split their vote across two or more contestants. If fans split their vote, they can do so in any way they choose: six votes to one contestant and four to another; seven votes to one contestant, two votes to another, and one vote to a third contestant; etc. The votes are then tallied to determine the winners.

This voting method—where voters have multiple votes to allocate across options—is called *multivoting*. It is often used in Agile teams: cross-functional groups which can rapidly adapt through the use of sprints, scrums, and daily standup meetings. Multivoting is sometimes called “dot voting” in these teams, because it often involves giving team members small circular stickers and having them distribute them across multiple options on a whiteboard. Members have much more leeway in expressing their preferences than they do in either plurality or ranked-choice voting. As the team members look at the results, they begin to get a consensus over which option the team as a whole prefers.

Despite its use in Agile teams, not much research has been done on multivoting. Researchers have, however, looked at two related techniques. Under *cumulative voting*, voters get as many votes as there are seats to fill in an election (Richie, 1994). For example, if a company was electing board members for three positions, shareholders would have three votes to allocate. They could go all in on one candidate or split their votes across two or more candidates. Under *quadratic voting*, voters can buy votes: the more they spend, the more votes they get. But the price increases with each vote they buy. For example, if the first vote cost $1, the second would cost $4, the third would cost $9, and so on (it is called quadratic voting because the price increase follows a quadratic function). This method has recently been proposed for corporate governance, where shareholders would not vote according to the number of shares they hold, but by the number of votes they buy (Posner & Weyl, 2014).

But neither of these methods seem appropriate for unofficial voting prior to team discussions. Cumulative voting permits voters to express a small degree of nuance, but not as much as they can with ten (or more) votes. Quadratic voting is not appropriate for teams in organizations, as no one would expect to have to pay to vote on their team's decision; plus, why pay for votes when it is only an unofficial vote?

So, let's take one more unofficial vote with your pursuit team: Instead of having everyone on your team choose only one suspect (plurality), or asking them to rank order the three suspects, you use multivoting and allow them to split their vote across the suspects in any way they wish. Team members may think two of the suspects are equally threatening, which they can show by giving them equal shares of their vote. Or they may think that one suspect is far more threatening than the other two, which they can show by giving a much larger share of their vote to that suspect.

Like *American* Idol voting, you give each member of your pursuit team 10 votes to allocate across the suspects. This new vote of your team comes out like this, in the same three blocs.

INSERT TABLE 5 HERE

The Army and Navy members in the first bloc think that Badger and Eagle are equally threatening (4 votes to each of these suspects). They also think that these suspects are twice as threatening as Jackal (2 votes). The CIA, FBI, and DOJ members in the second bloc think Jackal is slightly more threatening (4 votes) than Badger and Eagle (3 votes each). The Air Force and NSA members in the third bloc think that Eagle (6 votes) is three times as threatening as either Badger or Jackal (2 votes each).

To determine the results, add up the total number of votes for each suspect. Badger received 21 votes (eight from the first bloc, nine from the second bloc, and four from the third bloc). Jackal received 20 votes, and Eagle received 29 votes. Unlike the plurality and ranked-choice votes, the results now point to Eagle as the most threatening suspect. Thus, the three voting methods resulted in different "winners:" Jackal in plurality voting, Badger in ranked-choice voting, and Eagle in multivoting. Different voting methods yielded entirely different results.

One-third of the groups in our Threat Target experiment used multivoting in their unofficial votes. 45% of them identified the correct suspect in their final verdict—a 50% improvement over the other two voting methods. Why did the multivoting groups do so much better than the groups that used plurality or ranked-choice voting?

The answer may lie in the unofficial vote itself. All the groups improved after discussion, but the multivoting groups started better. Only 6% of the plurality voting groups and 7% of the ranked-choice voting groups identified the correct suspect in their unofficial votes, but 30% of the multivoting groups did. It is like they were competing in a race, but the multivoting groups got a head start on the others. And in fact, we tested this statistically through mediation analysis (see Appendix for details). This test showed that the accuracy of the unofficial vote was the mechanism that explained why voting type affected the groups' final decision accuracy.

INSERT FIGURE 1 HERE

This is surprising because research on social decision schemes centered around how the group's *discussion* affects their decision (Kerr & Tindale, 2004). Our head start discovery for multivoting groups might mean that the members thought more deeply about the decision *before* their discussion, making it more likely that team members would initially identify the correct suspect.

This suggests that when groups use multivoting, their members engage in more premeeting elaboration (Loyd et al., 2013). Premeeting elaboration is a term used to describe people's preparation for group discussions, in terms of how much they process information before meeting with the other team members. It helps groups make more accurate decisions: in a similar task to ours where teams had to identify the correct murder suspect, Loyd et al. found that groups whose members did more premeeting elaboration were far more likely to make the right decision. Because multivoting forces people to make more nuanced distinctions between the options presented to them, they must process their information more deeply.

HELPING TEAMS MAKE BETTER DECISIONS

The film *12 Angry Men* ends with every juror ready to return a not guilty verdict—save one. Juror #3 stubbornly holds to his original opinion that the defendant is guilty, (before, finally, changing his vote to 'not guilty,' following a personal revelation), prompting another juror to note, "It takes courage to stand alone." This ironic statement—ironic because Juror #8 stood alone at the beginning in his not guilty opinion—highlights a challenge with taking unofficial votes: Taking a contrary position to the rest of the group can come with social pressure to conform to what everyone else thinks.

But why should we expect dissenters to take such heroic, courageous stands? Wouldn't it be better if members could take contrary positions without such pressure? We think that social pressure is greatly diminished when using multivoting, partly because people can better express their own ambivalence in their vote, and partly because other members will also express ambivalence. Multivoting creates a situation where dissent is normalized. In *12 Angry Men*, the jury was fortunate to have a courageous dissenter. Had they used multivoting, he would not have had to be so courageous, as other jurors may have also expressed ambivalence in their initial votes.

Multivoting showed positive results in our study, but it may not be the only way that teams can get a head start on their decision-making. If other methods can be developed that help team members form better initial judgments, teams using these methods would have an advantage over other teams. For example, groups might start with a concerted effort to gather more information before they discuss the decision. They would start with better ideas and finish with better decisions.

This depends, however, on the quality and distribution of information that the group members have before their discussion. Sometimes, as in "hidden profile" tasks, every member has information that points to a wrong answer. If they combine their information, they can arrive at the correct answer, but past research has found that this is very difficult to do. Multivoting is not likely to solve this problem: even though members will think more deeply about their preferences, none of them is likely to allocate the most votes to the correct answer.

The advantages of multivoting may also depend on the expertise of the group members. In our study, we used college students to simulate counterterrorism pursuit teams. Real pursuit teams comprised of expert intelligence experts may find multivoting to be more or less effective in identifying the most threatening suspect. They likely engage in extensive information processing prior to group discussions, which might lessen the benefit of multivoting. On the other hand, experts are sometimes biased, and multivoting may provide a systematic consideration of their preferences that lessens the bias.

We expect the advantages of multivoting to lie mostly in its use as a voting method that is followed by group discussion, rather than as a final vote. Although many of the examples we used in this article are about final votes (like political elections), our experiment did not have groups multivote to make their final decision; they reached a consensus instead. It is possible that because multivoting groups were more likely to identify the correct suspect in their unofficial vote, multivoting *could* be effective as a final vote if discussion is not possible.

Final votes are often more about assessing a group's preference, rather than finding a correct answer. When viewers vote for *American Idol* contestants, they are expressing which contestants they like, not which one is objectively the best. We are unsure whether multivoting is the best method to use in these kinds of votes. It may be, because as they process more information, voters are more likely to reveal their true preferences. Indeed, it is possible that this extends to individual decision-makers who will not engage in group discussion but need to publicly express an initial opinion to an audience. Ranked-choice votes may also outperform plurality votes in these contexts as well. As the saying goes, "the jury is still out" on this.

Multivoting is not a panacea for the challenges faced by groups when they make decisions. It comes with a heavy cognitive load: it's much easier to choose one option—or even put the options in a rank order—than it is to allocate many votes across many options. Some groups may find this load to be too high, especially when they are faced with many options. Marketing research has found that customers are less likely to buy products when they have too many options; it's just too difficult to determine which one is best. Imagine trying to allocate 100 votes across 15 options, for example: making fine-grained distinctions across so many options may be nigh impossible. In such cases, groups may want to reduce the option set prior to voting.

How many votes should voters get when multivoting? There is no clear answer, but here is a formula that can serve as a guide: $v > \frac{o(o-1)}{2}$, where v is the number of votes each voter has to allocate and o is the number of options they can allocate their votes to. If $v < \frac{o(o-1)}{2}$, voters cannot even express ranked-choice preferences, much less express more nuance in their preferences. To illustrate, consider a group voting on three options. If $v = \frac{o(o-1)}{2}$, each voter has three votes ($3 = \frac{3(3-1)}{2}$). A voter can express a true ranked-choice preference by allocating two votes to one option, one vote to a second option, and no votes to a third option. If $v < \frac{o(o-1)}{2}$, voters have at most two votes, meaning they would be forced to give the same number of votes to two of the options (either 1 vote or zero votes to two of the options). Similarly, a group voting on five options must have a minimum of 10 votes ($10 = \frac{5(5-1)}{2}$) to express a ranked-choice preference (4, 3, 2, 1, 0). More than 10 votes would allow voters to express more nuance than ranked-choice voting.

Of course, multivoting is not the only complex voting system that allows voters to reveal more nuance in their preferences. Another possibility is *approval voting*, where voters indicate which candidates are acceptable to them and which are not. Approval voting is used in unofficial voting for the United Nations Secretary-General, where each permanent member of the UN Security Council receives an unofficial ballot with names of potential candidates. For each candidate, members indicate whether they encourage, discourage, or have no preference for their candidacy. The coordinating body tallies the votes, removes candidates that receive mostly "discourage" votes, and conducts another unofficial vote. When António Guterres, the current UN Secretary-General, was voted into office in 2016, the UN ran six such unofficial votes before nominating him. We encourage future research on other voting methods like this one; perhaps they will yield even better results than multivoting in some contexts.

Voting matters—for jury verdicts, political elections, and in our experiment, for counterterrorism pursuit teams. How teams vote also matters; voting methods can impact the fate of criminal defendants and presidential elections. Complex voting methods like multivoting are not familiar, and as a result, are rarely used—but it is their complexity that gives them the edge.

REFERENCES

Astor, M. 2019. Maine lessens risk of third-party spoilers by letting voters rank choices. ***New York Times****, September 8, 2019, Late Edition – Final, Section A*: 16.

Black, D. 1948. On the rationale of group decision-making. ***Journal of Political Economy***, *56*: 23–34.

Davis, J. H. 1973. Group decision and social interaction: A theory of social decision schemes. ***Psychological Review***, *80*: 97–125.

Devine, D. J., Olafson, K. M., Jarvis, L. L., Bott, J. P., Clayton, L. D., & Wolfe, J. M. T. 2004. Explaining jury verdicts: Is leniency bias for real? ***Journal of Applied Social Psychology***, *34*: 2069–2098.

Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. 2009. Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS estimation. ***Structural Equation Modeling: A Multidisciplinary Journal***, 16: 625-641.

Gaertner, W. 2009. ***A primer in social choice theory*** (Rev. ed., LSE Perspectives in Economic Analysis). Oxford; New York: Oxford University Press.

Kerr, N. L., & Tindale, R. S. 2004. Group performance and decision making. ***Annual Review of Psychology***, *55*: 623–655.

Laughlin, P. R. 2011. Social choice theory, social decision scheme theory, and group decision-making. ***Group Processes & Intergroup Relations***, *14*: 63–79.

Loyd, D. L., Wang, C. S., Phillips, K. W., & Lount, R. B. (2013). Social category diversity promotes premeeting elaboration: The role of relationship focus. ***Organization Science,*** *24(3):* 757–772.

Maskin, E., & Sen, A. 2016. How to let the majority rule. ***New York Times***: 7.

McLean, I. 1990. The Borda and Condorcet principles: Three medieval applications. ***Social Choice and Welfare****,7*(2): 99-108.

Posner, E. A., & Weyl, E. G. 2014. Quadratic voting as efficient corporate governance. ***University of Chicago Law Review***, *81*(1): 251–272.

Richie, R. 1994. Cumulative voting: Benefits and drawbacks. ***National Civic Review***, *83*(2): 214.

Schmitt, E. 2010. New teams connect dots of terror plots. ***New York Times****, January 29, 2010, Section A:* 3.
Seelye, K. Q. 2018. As Australia and the Oscars go, so goes Maine? ***New York Times****, June 12, 2018, Section A:* 14.
Selig, J. P., & Preacher, K. J. 2008. ***Monte Carlo method for assessing mediation: An interactive tool for creating confidence intervals for indirect effects*** [Computer software]. Available from http://quantpsy.org/.
Stasser, G. (1999). A primer of social decision scheme theory: Models of group influence, competitive model-testing, and prospective modeling. ***Organizational Behavior and Human Decision Processes***, *80*: 3–20.
Thompson, L. L., & Flammang, M. P. 2014. Threat target. ***Kellogg Team and Group Center***: Kellogg School of Management.
Zhao, X., Lynch, J. G., Jr., & Chen, Q. 2010. Reconsidering Baron and Kenny: Myths and truths about mediation analysis. ***Journal of Consumer Research***, *37*: 197–206.

Table 1. Results of first plurality vote for most threatening suspect.

ARMY NAVY AIRFORCE NSA	CIA FBI JUSTICE
Badger	Jackal

Table 2. Results of second plurality vote for most threatening suspect.

ARMY NAVY	CIA FBI JUSTICE	AIRFORCE NSA
Badger	Jackal	Eagle

Table 3. Results of ranked-choice vote for the most threatening suspect.

ARMY NAVY	CIA FBI JUSTICE	AIRFORCE NSA
Badger	Jackal	Eagle
Eagle	Badger	Badger
Jackal	Eagle	Jackal

Table 4. Estimates of exit poll results in the 2016 US Republican presidential primary.

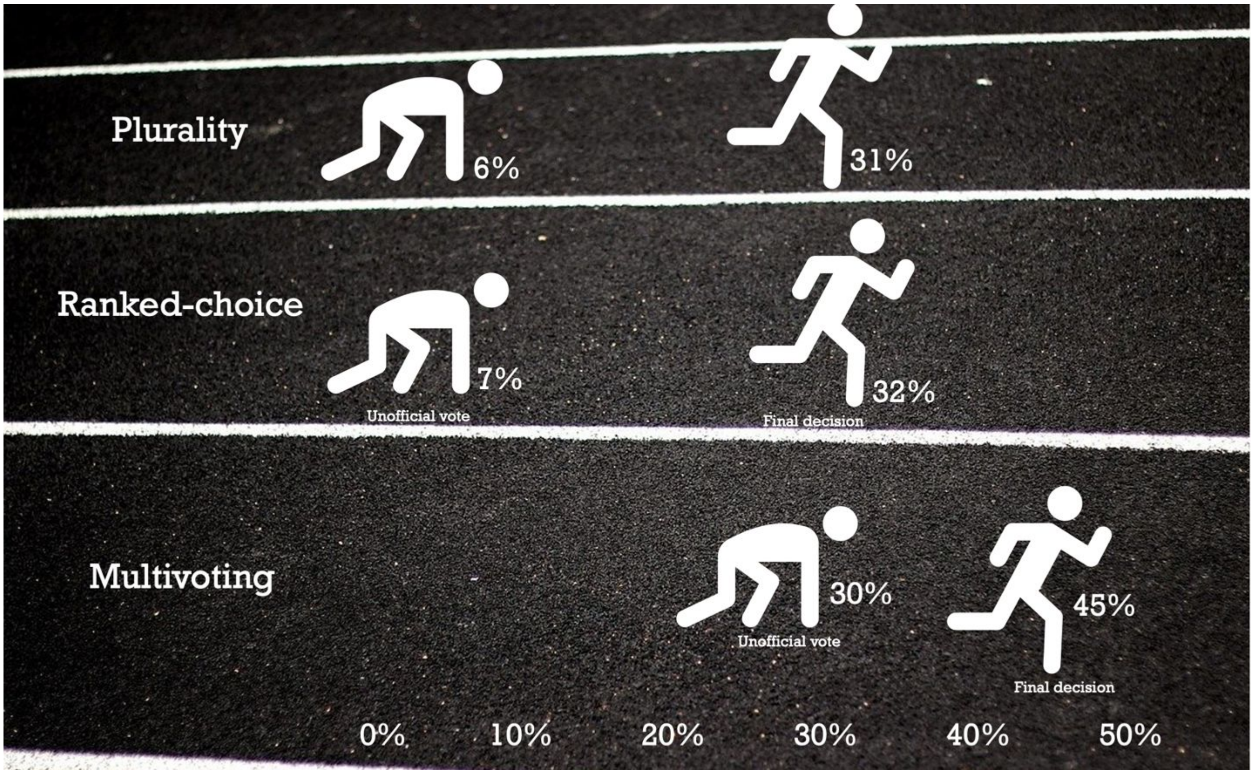
40% of voters	35% of voters	25% of voters
Donald Trump	Ted Cruz	John Kasich
John Kasich	John Kasich	Ted Cruz
Ted Cruz	Donald Trump	Donald Trump

Table 5. Results of multivote for the most threatening suspect.

ARMY NAVY	CIA FBI JUSTICE	AIRFORCE NSA
Badger: 4 votes	Jackal: 4 votes	Eagle: 6 votes
Eagle: 4 votes	Badger: 3 votes	Badger: 2 votes
Jackal: 2 votes	Eagle: 3 votes	Jackal: 2 votes

(The rank orders here are the same as in the previous vote.)

Figure 1. Results of various voting methods in the Threat Target experiment.



Unofficial vote percentages refer to groups where the majority of group members identified the correct suspect. Final vote percentages refer to groups who ultimately identified the correct suspect.

APPENDIX

Participants

Four hundred eighty-three undergraduate students (M_{age} = 20.80, SD_{age} = 3.59, 44.4% female, 53.2% White) from a large American university took part in this study for course credit. The participants signed up for a lab session at their discretion; when they arrived at the lab, they were randomly assigned to four-person teams (except for five teams of three), resulting in 120 teams. No individuals or teams were removed from the analysis for any reason.

Task

The teams completed a modified version of "Threat Target" (Thompson & Flammang, 2004). Threat Target is a team task where participants imagine they are a CIA analysis team. The teams were told that intelligence has been gathered on potential terrorist threats and they are tasked with determining which of three suspects (A, B, or C) represents the greatest terrorist threat. Having no more than three options ensured the task was not overly complex and participants were not under excessive cognitive load. For example, with three options, voters have six ways they can rank-order preferences. Adding one option would increase this to 24 ways. Similarly, while our multivoting condition with three options yielded 66 ways to distribute votes, adding one more option would increase this to 286 ways. After consenting to participate and answering a few preliminary questions on a survey, participants individually read general information about the task, and then received information on the three suspects.

This is a hidden-profile task, such that all participants received some of the same information about the suspects and also received some specific information about the suspects that their teammates did not receive. Hidden profile tasks require group members to combine their individually held information in order to reach the correct decision. Threat Target was previously pilot-tested by others with several government anti-terrorism groups (Thompson &

Flammang, 2014). In every case where the groups were given the full information about the suspects, the groups identified one particular suspect (B) as representing the greatest threat.

To increase our confidence that there was indeed a correct answer for this task, we enlisted the help of a former British naval intelligence officer who had previously worked in teams identifying threats and hunting war criminals. He subjected this task to the same scrutiny and using the same materials he used when serving on a team that pursued war criminals, providing us with a verbal report of how he used these materials to analyze the available information about the suspects. His conclusion agreed with the anti-terrorism groups that suspect B represented the greatest threat. Thus, we continued with confidence that the task indeed had a single correct answer.

While the information in some hidden profile tasks is distributed in such a way that a suboptimal decision is usually made, the intent with this experiment was to start instead from a point of ambiguity. Thus, we redesigned the information distribution such that participants on their own were equally likely to select any of the three candidates.

Manipulation of unofficial vote voting condition

We randomly assigned teams to one of three experimental unofficial vote conditions, or a control condition with no unofficial voting. In the experimental conditions, after reading the scenario and suspect information, team members independently and secretly voted on the threat potential of the three suspects. They did this by indicating their choices on slips of paper. In the plurality voting condition, participants chose one of the three suspects by indicating an X by that suspect's name. In the ranked voting condition, participants indicated a rank order of the suspects' threat potential by rank ordering them from 1 (highest threat) to 3 (lowest threat). In the multivoting condition, participants were told that they had 10 votes to allocate across the three suspects. They did so by writing a number of votes next to each suspect's name. Thus, a participant could give all ten votes to one suspect or spread their votes across multiple suspects.

Following the unofficial vote, we escorted participants into a private room for team discussion. At the beginning of the discussion, we gave the unofficial vote voting slips to the teams so they could see their teammates' initial preferences (although no names were on these slips, participants often shared with their teammates which slip was theirs as they explained why they voted the way they did). We gave teams 15 minutes to come to a consensus decision about which suspect they collectively believed was the greatest threat. We did not give any instructions to the teams about how they should discuss the scenario or come to their final decision. In the control condition, teams simply began discussing without an unofficial vote.

Measures

We measured *decision quality* with a binary variable indicating whether the final team decision was correct (1) or incorrect (0). As a mediator, we examined the unofficial vote results from each team according to their condition to determine if most of the team members thought the correct suspect was the most threatening before their team discussion. For the plurality condition, we measured this as most members choosing the correct suspect in their unofficial vote. For the ranked-choice condition, we measured this as most members choosing the correct suspect as their top-ranked threat in the unofficial vote. For the multivoting condition, we measured this as whether the members assigned more votes to the correct suspect than to either of the other suspects. This resulted in a binary variable called *majority unofficial vote correct.*

Analytical approach

We modeled the data with a path analysis conducted in R 4.1.3 using the *lavaan* package. Because all variables were arranged as binary dummy codes, careful consideration of the estimation procedure is required. While the standard procedures in both *lavaan* and Mplus use a diagonally weighted least squares estimator for binary endogenous variables, there is some concern about the accuracy of this approach with smaller sample sizes (under 200; Forero, Maydeu-Olivares & Gallardo-Pujol, 2009). Thus, we used an unweighted least squares estimator with robust standard errors to account for the inherent non-normality of binary variables. (The default diagonally weighted least squares estimators had a similar pattern of results.)

Results

The results of the path analysis are presented in Table A1. We first examined the impact of voting condition on the results of the unofficial vote (majority unofficial vote correct, for all teams in experimental conditions). Relative to the plurality voting condition, the ranked-choice voting condition teams were not significantly more likely to identify the correct suspect in their initial unofficial votes (b = 0.069, p = .891). The multivoting condition, however, significantly improved the likelihood of a majority correct unofficial vote (b = 1.018, p = .015).[3]

INSERT TABLE A1 HERE

Next, we examined whether the results of the unofficial votes mediated the effect of voting condition on decision quality. Teams that had a majority of members get the unofficial vote correct were more likely to make the correct decision as a team (b = 0.766, p < .001). A bootstrap test of the indirect effect showed that majority correct unofficial vote was a statistically significant mediator between the multivoting condition and decision quality (95% CI [0.068,1.492]). Bootstrapped standard errors were also evaluated and show a similar pattern of results. We also conducted a separate analysis which modeled the direct effect of conditions

[3] In addition, the same model was run using ranked-choice voting rather than plurality voting as the base case. This showed that the differential effect of the multivoting condition relative to ranked voting was significant (b = 0.950, p = .026), while the differential effect of the plurality condition relative to ranked voting was not statistically significant (with the same results as shown in Table 1).

on the dependent variable without the mediator. We found that the direct relationship between multivoting condition and decision quality was not statistically significant ($b = 0.375$, $p = .242$), this is an example of an indirect-only mediation (Zhao et al., 2010).

Supplemental Analyses

We further tested for differences in the dependent variable across conditions, after including teams in the control condition. These teams were not included in the main analyses as they did not engage in any unofficial voting, and therefore could not provide data on the mediator of initial vote accuracy. Similarly, a one-way ANOVA revealed no significant differences in decision quality across the four conditions ($F(3, 116) = .62$, $p = .605$).

Table A1: Results of path analysis

Path	*b*	SE	*z*	*p*
Mediator: Majority correct unofficial vote				
Ranked condition	0.069	0.501	0.137	.891
Multivoting condition	1.018	0.419	2.432	.015
DV: Decision quality				
Majority correct unofficial vote	0.766	0.135	5.684	<.001
Ranked condition	-0.028	0.400	-0.069	.945
Multivoting condition	-0.405	0.385	-1.054	.292

N = 93 teams
(Note: For all condition comparisons, the base condition is plurality voting.)

Michael D. Johnson (mdj3@uw.edu) is the Boeing Company Endowed Professor of Business Management in the Foster School of Business at the University of Washington. He received his Ph.D. in Organizational Behavior/Human Resource Management from Michigan State University. His research focuses on groups and teams, emotions, and social identities.

Eli Awtrey (eli.awtrey@uc.edu) is an assistant professor of management at the Carl H. Lindner College of Business, University of Cincinnati. He received his Ph.D. from the University of Washington's Foster School of Business. His current research investigates teams and positive organizational behavior, with a focus on how interpersonal relationships in a group influence processes and outcomes.

Wei Jee Ong (weijee@nus.edu.sg) is an assistant professor of management and organisation at the National University of Singapore Business School. He received his Ph.D. from the University of Washington's Foster School of Business. His research examines the constraints individuals face at work and how they navigate them, particularly in the domains of leadership, job design, and ethics.