



**HAL**  
open science

## How People Understand Voting Rules

Antoinette Baujard, Roberto Brunetti, Isabelle Lebon, Simone Marsilio

► **To cite this version:**

Antoinette Baujard, Roberto Brunetti, Isabelle Lebon, Simone Marsilio. How People Understand Voting Rules. 2024. hal-04653702

**HAL Id: hal-04653702**

**<https://hal.science/hal-04653702v1>**

Preprint submitted on 19 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# How People Understand Voting Rules

Antoinette Baujard\*    Roberto Brunetti†    Isabelle Lebon‡    Simone Marsilio§¶

July 11, 2024

## Abstract

If individuals are to be empowered in their selection or use of a voting rule, it is necessary that they understand it. This paper analyzes people’s understanding of two voting rules: evaluative voting and majority judgment. We first distinguish three components of understanding in this context: how to fill in the ballot; how votes are aggregated; and how to vote strategically. To measure each component, we draw on results from a lab experiment on incentivized voting where participants are exogenously assigned single-peaked preferences and answer comprehension questions on the rules employed. We find that most participants understand how to fill in the ballot and how votes are aggregated; however, participants’ understanding of vote aggregation under majority judgment is lower and more heterogeneous. While some participants correctly understand its aggregation property, a sizable group fails to grasp it. We also observe no difference in voting behavior between evaluative voting and majority judgment: the data confirm the theoretical prediction that under evaluative voting there will be a greater incidence of strategic voting through the use of extreme grades, but contradict the prediction that under majority judgment voters will vote more sincerely. Finally, we find that with majority judgment, the better voters understand how votes are aggregated, the more they make use of extreme grades.

**Keywords:** voting rules; understanding; evaluative voting; majority judgment; laboratory experiment

**JEL codes:** A13, C92, D71, D72, O35

---

\*Université Jean Monnet Saint-Etienne, CNRS, Université Lumière Lyon 2, emlyon business school, GATE, 42023, Lyon, France

†Université Lumière Lyon 2, CNRS, Université Jean Monnet Saint-Etienne, emlyon business school, GATE, 69007, Lyon, France, and Univ Rennes, CNRS, CREM-UMR6211, F-35000 Rennes, France. Email: [roberto.brunetti@cnrs.fr](mailto:roberto.brunetti@cnrs.fr)

‡Normandie Univ., CREM, UMR CNRS 6211, Caen, France and TEPP-CNRS, Caen, France

§Università Vita-Salute San Raffaele, Milano

¶We are grateful to the participants at the Social Choice and Welfare meeting (Paris), the GATE on the Docks workshop (Lyon), and the Amsterdam/Saint-Etienne Workshop on Social Choice (Saint-Etienne) for their insightful comments and discussions, as well as to Jean-François Laslier for his comments on the first draft of this paper. We would also like to thank Sylvain Bouveret, Gérard Chatagnon, Herrade Igersheim, Rida Laraki, Rémy Poulain, Elven Priour, and Quentin Thevenet for their help and support at various stages of the design conception. This research was jointly funded by the Research Grant “Voter Autrement 2022” from the University Jean Monnet, by CREM Caen, and by the French National Research Agency (ANR) under the Citizens project “ANR-22-CE26-0019-01”. For the purpose of Open Access, a CC-BY public copyright licence (<https://creativecommons.org/licenses/by/4.0/>) has been applied by the authors to the present document and will be applied to all subsequent versions up to the Author Accepted Manuscript arising from this submission.

# 1 Introduction

The social choice literature presents a wide range of alternative voting rules intended to counter the shortcomings of traditional ones. Yet while these rules may have desirable theoretical properties (Sen, 1995; Brams and Fishburn, 2002), their validity in practice depends on assumptions about voters’ behavior. One crucial assumption is that voters properly understand the rules in question, which may not always be true, especially where they involve unusual balloting systems and vote aggregation methods (Laslier, 2011). Unexpected issues can therefore arise: for instance, with rules where voters assign grades to every candidate, voting inconsistencies increase when the range of available grades widens (Baujard et al., 2018). Moreover, the complexity of a rule can be a barrier to its usability (Bourgeois-Gironde and Ferreira, 2023)<sup>1</sup> as well as its capacity to meet other desirable properties such as voter satisfaction (Blais et al., 2021). In the end, then, understanding is a normative property *per se*, as it concerns the degree to which citizens are empowered to use a given rule to reach their objectives. Since voter confusion can only be observed in the real world, this issue calls for empirical study. To our knowledge, however, only a few papers have sought to tackle this issue, and in particular we do not know of any attempt to offer a systematic method to measure citizens’ understanding of voting rules.

This paper aims to fill this gap by scrutinizing individuals’ understanding of two alternative voting rules: evaluative voting<sup>2</sup> and majority judgment. These two rules offer an ideal testing ground for exploring voters’ understanding for at least three reasons. First, both rules feature one-round elections with multi-nominal ballots—*i.e.*, voters assign grades to every candidate—yet with two differences: in evaluative voting, grades are numerical and aggregated by their sum, while in majority judgment, grades are verbal and aggregated by the median; we can therefore leverage the subtle yet relevant differences between these rules to study people’s understanding of their specific characteristics. Second, evaluative voting and majority judgment are increasingly the subject of study among academics (since Hillinger, 2006; Balinski and Laraki, 2007), but there is limited empirical evidence on how individuals use these rules. Third, even though both rules are already used in practice,<sup>3</sup> many people lack experience with them. Therefore, we can study people’s understanding absent any previous knowledge of the rule or any *status quo* bias.

To do this, we run an incentivized laboratory experiment based on Van Der Straeten et al. (2010), where people are assigned single-peaked preferences and vote for five candidates. In our case, participants vote using either evaluating voting or majority judgment depending on the treatment to which they are assigned. The laboratory conditions offer time and incentives for individuals to encounter the two voting

---

<sup>1</sup>In a more general context, Oprea (2020) considers rules as algorithms to be implemented by humans and finds that most subjects in a lab-experiment are willing to pay to avoid implementing complex rules.

<sup>2</sup>Evaluative voting is also known as utilitarian voting, range voting, score voting, or point voting. The expression *vote par note* (grade voting) is also used in French.

<sup>3</sup>For instance, an equivalent of three-level evaluative voting is used in Latvia’s parliamentary elections, where voters can give a negative, null, or positive grade to each candidate in the chosen party list (Laslier et al., 2015). A variation of evaluative voting in the form of cumulative voting is used in Luxembourg (Laslier et al., 2015) and the German states of Bremen and Hamburg (Bowler et al., 2018), where voters can assign a fixed total of points to different candidates. Majority judgment was used in France for the local elections of the French political party “La République en Marche”, and for the primary elections of the left-wing coalition (*primaire populaire*) in 2022.

rules, learn them, and practice them. We conform to the common practice in experimental economics of reinforcing the experimental instructions with examples and context-rich practice rounds (Freeman et al., 2018). This allows us to scrutinize participants’ understanding of voting rules in an ideal learning condition and at different stages. Most importantly, we can study people’s understanding of the aspects of specific voting rules by observing their voting behavior and incentivized responses to targeted questions. This approach overcomes the metacognitive issue of individuals’ illusion of understanding (Rozenblit and Keil, 2002; Sloman and Fernbach, 2017), which can bias their *self-assessed* and *general* knowledge of voting rules.

We distinguish and measure three components of understanding: First, we investigate whether people understand how to fill in a ballot, which involves participants’ knowledge of what grades mean and how to use them to provide support to the desired candidate(s). This component is measured by the number of inconsistent votes (Igersheim et al., 2016; Baujard et al., 2018; Darmann and Klamler, 2023), *i.e.*, choices such that participants rate a more-preferred candidate lower than a less-preferred one. Second, we observe peoples’ understanding of how votes are aggregated, which concerns the aggregation methods (sum/mean vs. median) used in each voting rule. This component is measured using a set of incentivized questions where participants compute the results of mock elections with evaluative voting and majority judgment, and where they also answer some true/false questions. Finally, we study whether people vote strategically, which is related to participants’ knowledge of how to vote to maximize their payoffs. To do this, we study the number of extreme grades assigned to candidates by voters, which signals a departure from sincere voting, and look at the difference when voting with one or the other rule.

We present four main results: First, only a few voters vote inconsistently with both the multinomial rules under test, suggesting that people understand how to fill in a ballot and can easily express their preferences with grades. Second, participants’ understanding of how votes are aggregated is, on average, lower and more heterogeneous with majority judgment than evaluative voting. This difference arises from participants’ difficulties in understanding the definition of the median and its implications on the vote outcome. Third, and in contrast to theoretical predictions, we show that participants vote similarly with evaluative voting and majority judgment: indeed, most voters only use extreme grades, which is in line with the prediction that there would be a greater incidence of strategic voting in the case of evaluative voting (Núñez and Laslier, 2014) but at odds with the prediction of more sincere voting with majority judgment (Balinski and Laraki, 2020). Finally, we provide evidence that those participants who grasp the aggregation property of the majority judgment rule assign extreme grades even more often than participants who do not grasp it.

**Related literature** Our paper relates to three strands of literature. First, our paper speaks to the social choice literature that studies alternative voting methods, such as evaluative voting and majority judgment. This literature tends to focus on the theoretical properties of these voting rules (Hillinger, 2006; Balinski and Laraki, 2007; Núñez and Laslier, 2014; Macé, 2018) and tests their applicability and empirical implications (Baujard and Igersheim, 2009; Baujard et al., 2014, 2018, 2021; Balinski and Laraki, 2011, 2012; Darmann et al., 2017, 2019; Igersheim et al., 2022; Darmann and Klamler, 2023),

and only a handful of papers probe voters’ understanding of the rules. Few papers measure self-reported understanding through questions such as “How well do you think you understand evaluative voting?”, to which most respondents answer positively (see [Baujard and Igersheim, 2009](#) for evaluative and approval voting, [Kimball and Kropf, 2016](#) for cumulative voting, and [Donovan et al., 2019, 2022](#), for ranked-choice voting). This approach has two major limitations. First, the definition of what is meant by “understanding” is left to respondents’ interpretations; second, participants who claim to understand the rule might not realize that they do not understand it, or might not want to admit it.<sup>4</sup>

A second method relies on assessing voters’ errors when casting a ballot. For instance, in the case of ranked-choice voting, this implies checking for the presence of undervotes and overvotes (see, *e.g.*, [Cormack, 2023](#)).<sup>5</sup> In the case of multi-nominal rules, single-name ballots can signal voters’ misunderstanding, especially since in most countries voting requires a single name.<sup>6</sup> However, single-name ballots can also derive from the characteristics of the political supply (where there are no similar candidates that voters would co-approve) or from an intrinsic preference for single-name voting rules (which could be reflected in a *status quo* bias). Hence, it is challenging to disentangle respondents’ misunderstanding from other factors. Finally, some studies focus on inconsistent votes, *i.e.*, votes such that voters prefer X and not Y but give a strictly higher evaluation to candidate Y than to X ([Igersheim et al., 2016](#); [Baujard et al., 2018](#); [Darmann and Klamlar, 2023](#)). In these studies, inconsistent votes are generally rare but become more common with broader grading scales (from 3.41% with a small centered scale to 8.9% with the [0–20] scale) ([Baujard et al., 2018](#)). [Balinski and Laraki \(2011\)](#) have conducted a similar analysis on inconsistent votes under majority judgment; however, the latter study only assessed the difference between the aggregate results with majority voting and the second round of the actual presidential election in Orsay, a suburban town near Paris.

A third method relies on the relationship between the complexity of a voting rule and voters’ behavior in the laboratory ([Herzberg and Wilson, 1988](#); [Van Der Straeten et al., 2010](#); [Dellis and Kröger, 2023](#)).<sup>7</sup> In particular, [Van Der Straeten et al. \(2010\)](#) provide the experimental design we build on here, and study voting behavior with one-round plurality, two-round majority, approval voting, and single-transferable voting with Hare transfers. They find that fewer individuals vote strategically with more complex voting rules: specifically two-round majority and single-transferable voting. However, in these papers, the complexity of voting rules is assumed rather than investigated, leaving aside how complex each voting rule is from the voters’ perspective.

Our contribution to the literature on the properties of voting rules is threefold. First, our experiment features *incentivized* questions that test people’s understanding of the aggregation properties of the rules, which is an understudied topic. Second, by considering inconsistent and strategic voting as in the

---

<sup>4</sup>Responses can be influenced by the experimenter demand effect ([Zizzo, 2010](#)) so that participants might conform to what is viewed as appropriate behavior. Relatedly, participants might want to protect their self-image.

<sup>5</sup>In ranked-choice voting voters must rank every candidate: hence, undervotes imply ranking fewer than the existing candidates, while overvotes imply assigning a rank to the same candidate(s) more than once.

<sup>6</sup>In *in-situ* experiments with approval voting conducted in France in 2012, between 18.62% and 22.43% of voters surveyed approved only one candidate, depending on which scale of evaluative rules was under experiment ([Baujard et al., 2012](#)).

<sup>7</sup>In parallel to this literature, several studies in political science have investigated whether people vote correctly. This concept varies depending on individuals’ priorities—some may prioritize expression over instrumental voting to maximize their payoffs—and also hinges on the accuracy of available information ([Lau and Redlawsk, 1997](#); [Blais et al., 2016](#)).

literature mentioned above, our paper combines three components of understanding in one design and tests their interrelation. Third, to our knowledge, our paper is the first to test majority judgment in the lab.

Our paper also contributes to a second strand of literature, namely the behavioral social choice literature initiated by [Regenwetter \(2006\)](#). This literature investigates the divergences between social choice theory and the empirical evidence, as in the case of the Condorcet paradox of cyclical majorities ([Regenwetter et al., 2003, 2008](#)). In general, considering how voters use voting rules in the field can help to moderate some pessimistic theoretical predictions ([Regenwetter and Tsetlin, 2004](#); [Darmann et al., 2019](#); [Lachat and Laslier, 2024](#)) and to scrutinize critical normative issues that are concealed by a standard theoretical approach ([Baujard et al., 2014](#); [Darmann et al., 2017](#); [Darmann and Klamler, 2023](#)). In this paper, our data support the claim that evaluative voting can lead to strategic voting in the form of extreme grades being given to all candidates ([Núñez and Laslier, 2014](#)). At the same time, the same behavior is observed with majority judgment, which contrasts with what has been claimed concerning its lower manipulability ([Balinski and Laraki, 2011, 2020](#)).

Third, we contribute to the experimental methodological literature that studies subjects' understanding of instructions in economic experiments. Early studies inquired whether participants' confusion drove some common experimental results. For instance, [Andreoni \(1995\)](#) and [Houser and Kurzban \(2002\)](#) find that in a standard public good game, a participant's choice not to free-ride is partly driven by their confusion about the experimental task. This raises the question of whether participants understand the instructions as well as the implications of this possible confusion. Some papers focus on which types of instructions reduce confusion ([Bigoni and Dragone, 2012](#); [Alekseev et al., 2017](#); [Freeman et al., 2018](#); [Ramalingam et al., 2018](#)), others on which types of participants are likely to experience more confusion ([Bigoni and Dragone, 2012](#); [Pulickal and Chakravarty, 2023](#)). Although we do not focus on *how* to deliver instructions, nor on *who* understands them best, our paper brings a fresh perspective on *what* participants may fail to understand in the context of voting rules. In particular, we observe participants' confusion over the median and the implications of this in the case of majority judgment. This result acts as a warning for experiments where the instructions include statistical notions as simple as the median.<sup>8</sup>

The remainder of the paper is organized as follows. Section 2 presents the voting rules tested, namely evaluative voting and majority judgment, and offers a comprehensive definition of "understanding". Section 3 presents the experiment, the design, and the subjects. In section 4, we set out our results. Section 5 concludes. Detailed figures and the experiment description are provided in the appendices. Data are available on a remote depository under a Creative Commons license.

---

<sup>8</sup>For example, many studies offer statistics, including averages, ratios, and quintiles, in order to reduce peoples' misperceptions about society, but the effect is often muted. In a comprehensive literature review, [Marino et al. \(2024\)](#) argue that one possible explanation is that some people might fail to understand the information provided as their analytical reasoning skills are limited.

## 2 Understanding voting rules

### 2.1 Tested voting rules

We restrict our attention to voting rules which only elect one candidate, by contrast with voting rules, such as proportional rules, which are designed to elect a committee. Evaluative voting and majority judgment share the common feature of being a one-round voting rule based on multi-nominal ballots, *i.e.*, where voters assess every candidate with a grade. However, they differ concerning the grades available to voters and their aggregation method.

**Evaluative voting.** With evaluative voting, each voter assigns every candidate a grade on a pre-determined cardinal numerical scale (in our case, between  $-1$  and  $+4$ ), and the winner is the candidate who obtains the highest sum of grades (Hillinger, 2006; Smaoui and Lepelley, 2013; Macé, 2018). This is equivalent to saying that the winner is the one who obtains the highest average grade.

**Majority judgment.** With majority judgment, each voter assigns a verbal grade (in our case, from “To reject” to “Very good”, for a total of six grades) to each candidate, and the winner is the candidate who obtains the highest median grade, also called the “majority grade”. However, even with six different grades, two or more candidates often receive the same majority grade (the pigeonhole principle), and a tie-breaking rule is needed. Unlike other best-median rules, the majority judgment is characterized by a specific tie-breaking rule (Balinski and Laraki, 2007, 2011; Fabre, 2021). This rule can be implemented in two ways, whose explanations correspond to two treatments in our design.

For a large population of voters, we can apply the *majority gauge method*. In the case of ties, the tie is broken by considering that the largest block of voters who disagree with the endorsement decides, *i.e.*, the largest group of voters who think the candidate is worth more (or less) than the endorsement. In the *iterative deletion method* (the one suited for a low number of voters), if two candidates share the same majority grade, a single majority grade is removed from each candidate until the tie is resolved and the candidates can be ranked.

### 2.2 The components of understanding

**How to fill in a ballot** The first level of understanding concerns whether voters know how to fill in a ballot, which depends on what balloting information is required by each voting rule: a single name, a selection of a subset of candidates, an assessment of every candidate, and so on. In elections requiring a single name to elect a committee or a single winner, the low ratio of nulls (a part of which may be voluntary) suggests that difficulties are rarely encountered. By contrast, when voters independently assess each candidate with a numerical or verbal grade, they may encounter problems in interpreting grades. Indeed, the ordinal meaning of grades is not necessarily similar for everybody, and a given grade may convey dissimilar cardinal meanings (Baujard et al., 2021). Yet the aforementioned evidence shows that most voters cast multi-nominal ballots and assign grades consistent with their preferences with both evaluative voting and majority judgment (Balinski and Laraki, 2011; Igersheim et al., 2017; Baujard

et al., 2018). Additionally, as long as the grading scale and the number of candidates are the same with both voting rules, as it will be in our experiment, there is no reason why we should observe more mistakes with one rule than the other. Thus, we propose the following:

**Hypothesis 1:** Most participants understand how to fill in a ballot with both evaluative voting and majority judgment.

**How votes are aggregated.** The second level of understanding concerns the ability of voters to know how ballots are counted and aggregated to obtain the election outcome with a given voting rule. This level of understanding helps guarantee the legitimacy of an election outcome, which will be acceptable to voters not only if they approve the rule employed, but also if they can compute the result by themselves and reproduce the same outcome as the officials. This is why scrutinizing the ballots and counting the cast votes is a public and transparent activity in many countries. However, with complex voting rules the results cannot be computed at a decentralized level or with pen and paper,<sup>9</sup> and the issue of legitimacy becomes all the more important. Past experiments have provided data on voters’ use of more or less complex rules. However, as Laslier (2011) regrets, “Unfortunately, we do not learn from these experiments whether [participants] understand the way ballots are counted. This is not a problem for rules using simple counting schemes, but it is a problem for complex evaluative or ranking ballots.”

Evaluative voting and majority judgment differ in their aggregation methods. Therefore, we study whether participants understand the implications of the use of the sum/average and the median. Regarding majority judgment, there are three potential sources of confusion. The first obstacle is the definition of the median itself: some individuals might not be used to computing medians in their daily lives. Second, as a “best median” rule, majority judgment does not respect the intuitive features of additive rules. In particular, when we compute the sum or the average of grades, assigning a higher number of better grades to one candidate than another always improves the outcome for the first candidate; however, this monotonicity property does not hold when candidates are ranked on the basis of their median grades in the case of a variable population. This violation of the monotonicity property in the context of variable population is a counter-intuitive feature of the median that is likely to create confusion. Third, the tie-breaking rule could be a potential (and major) reason for voters’ misunderstanding. This paper does not tackle people’s understanding of the tie-breaking rule, as this could complicate the comparison between majority judgment and evaluative voting. Notably, these three obstacles are absent in evaluative voting: computing sums or averages is a daily task for most people; the rule respects the monotonicity property; and no tie-breaking rule is needed. Therefore, we propose the following:

**Hypothesis 2:** More individuals understand how votes are aggregated with evaluative voting than with majority judgment.

---

<sup>9</sup>For instance, the single-transferable vote used in Ireland requires the centralization of the ballots, while majority judgment and most participatory budgeting allocation rules require an algorithm to compute the outcome.



**How to vote strategically.** The third level of understanding concerns voters’ behaviors with respect to rules. First, rules have various properties.<sup>10</sup> Second, each voting rule calls for different voting behavior that is adapted to the specifics of the rule. However, if individuals do not understand rules correctly, they may behave homogeneously irrespective of the rule they are presented with. In the lab, we can assess whether voters vote correctly, that is, if they “maximize one’s payoffs in lab elections with monetary incentives” (Blais et al., 2016). Understanding how to vote can also be interpreted as the ability to vote strategically with the rule one is presented with. For evaluative voting, Núñez and Laslier (2014) show that strategic voters should overstate their preferences, using only extreme grades (either the lowest or the highest) and discarding intermediate grades. However, previous evidence shows that a significant proportion of voters use intermediate grades (Igersheim et al., 2016, 2022), although they do so less frequently in laboratory elections (Igersheim et al., 2016). By contrast, majority judgment is strategy-proof in grading, given that the aggregation method relies on the median grade (Balinski and Laraki, 2011, 2020). Indeed, assigning an extreme or slightly lower (higher) grade to a candidate does not alter its majority grade as long as both grades are above or below the median. Hence, voters in this case ought to vote more sincerely than with the highly manipulable evaluative voting. Therefore, we propose the following:

**Hypothesis 3:** Individuals assign extreme grades more frequently when using evaluative voting than majority judgment.

As measuring people’s understanding is hardly straightforward, and because we are not aware of a specific experimental protocol targeting such an aim, we designed a protocol in which participants vote in elections with monetary incentives under evaluative voting and majority judgment, and are then given some exercises and questions to answer.

## 3 The experiment

### 3.1 Experiment design

The experiment comprises three parts, as illustrated in Table 1: The first part is designed to teach participants four voting rules (the official two-round system, used in France in the general elections; approval voting; evaluative voting; and majority judgment) and test them in a given context. Depending on the treatment to which participants are assigned, the second part is devoted to testing evaluative voting or majority judgment in fictitious elections with controlled preferences. Finally, the third part involves a survey to gauge participants’ comprehension and satisfaction with the voting rules to which

---

<sup>10</sup>Laslier (2011) refers to the ability of voters to identify the normative properties of a voting rule, such as whether a rule will favor consensual candidates or restore a tribune to minor candidates. It would be interesting to compare which properties of the rules that participants properly or wrongly identify, to the benchmark of the scientific knowledge on their respective properties. However, the definition of a benchmark requires both the existence of numerous publications and a scientific consensus on the theoretical and empirical properties of these rules, which is still a work in progress for majority judgment and evaluative voting. For this reason, we do not tackle this issue here.

they have been introduced. Some socio-demographic questions follow. The instructions, translated from French to English, are presented in Appendix B. In the following paragraphs, we present each part of the experiment in detail.

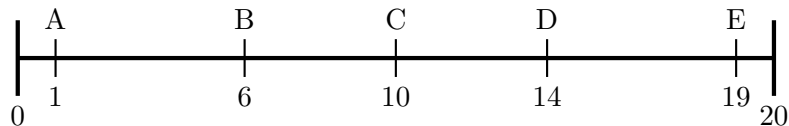
Table 1: Sequence of the experiment

<b>Part I</b>	<ul style="list-style-type: none"> <li>- Learning of two-round system, approval voting, evaluative voting, and majority judgment</li> <li>- Vote in specific context</li> <li>- Satisfaction questions</li> </ul>
<b>Part II</b>	<ul style="list-style-type: none"> <li>- First series of elections - Uniform preference distribution</li> <li>- Second series of elections - Trimodal preference distribution</li> <li>- Third series of elections - Bimodal preference distribution</li> </ul>
<b>Part III</b>	<ul style="list-style-type: none"> <li>- Comprehension questions</li> <li>- Satisfaction questions</li> <li>- Socio-demographic questionnaire</li> </ul>

**Part I.** In the first part, participants read the explanation of each voting rule and cast votes on the color of the condominium’s common walls. Initially, they use the two-round system, which they are most familiar with due to its use in French presidential elections (indeed, only participants familiar with the French system/who had already voted with it were eligible for the experiment). Then, participants try approval voting,<sup>11</sup> evaluative voting and majority judgment. After learning about each voting system, participants answer some questions about their satisfaction with the different voting systems.

**Part II.** In the second part of the experiment, we use a standard design that has already been employed in many publications to study the impacts of voting rules, which candidates they favor, and voters’ behavior (*e.g.* Van Der Straeten et al., 2010; Igersheim et al., 2016). Twenty-one participants vote for five candidates, labeled A, B, C, D, and E, presented on a 0 to 20 scale axis (see figure 1). The candidates are symmetrically positioned so that there is an extreme left candidate (A) located at 1, a moderate left (B) located at 6, a centrist (C) located at 10, a moderate right (D) located at 14, and an extreme right (E) located at 19. Participants are informed that the candidates’ positions remain constant throughout the experiment.

Figure 1: Position of the five candidates

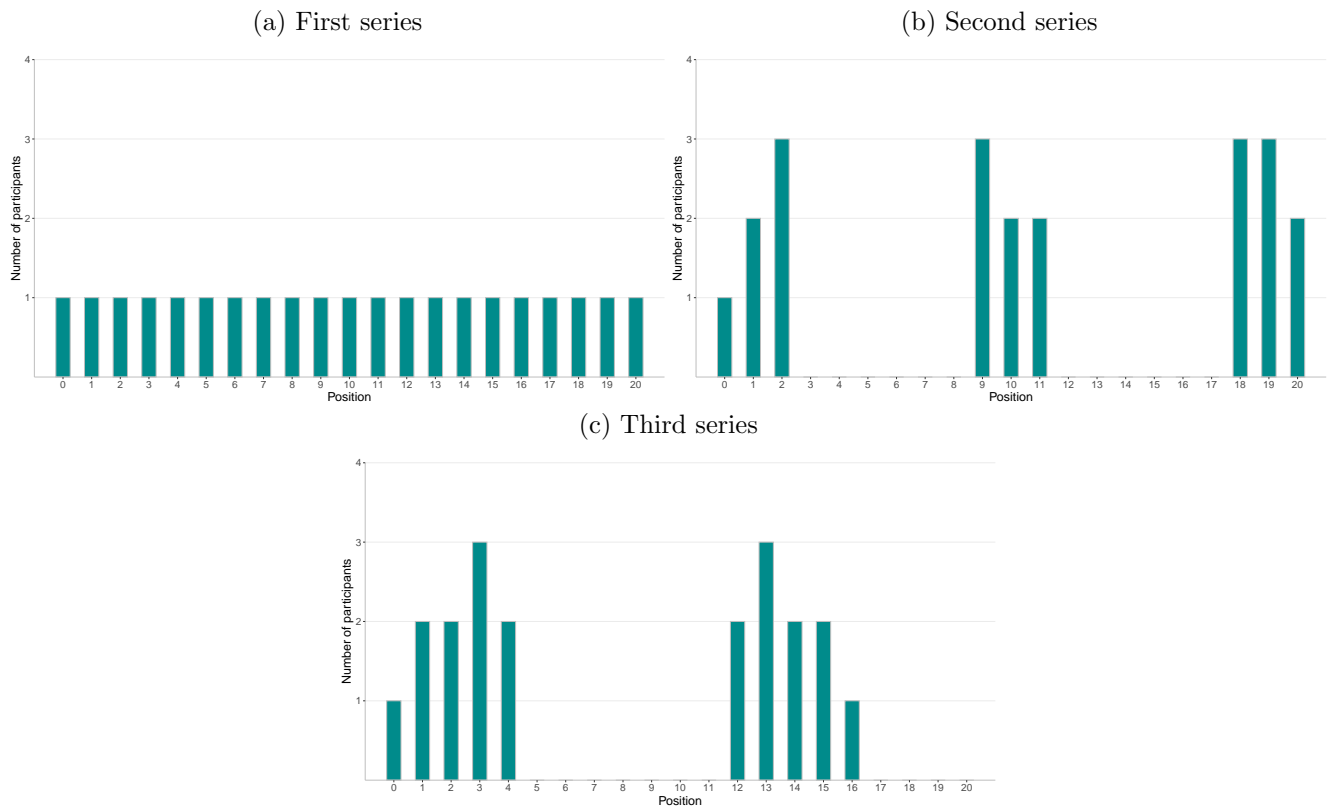


Participants vote across three series of five elections each. At the start of each series, each participant

<sup>11</sup>Approval voting is a special case of evaluative voting rules, with the specific grading scale [0, 1]. Under approval voting, either voters approve or do not approve every candidate, and the ranking of candidates corresponds to the ordering of their number of approvals (Brams and Fishburn, 1978).

is allocated a position on the  $[0,20]$  axis, which remains fixed throughout the series. Their position represents their preferences. The distribution of participants' positions varies across the three-election series: In the first series, preferences are uniformly distributed, meaning that each participant is assigned one position between 0 and 20 with equal probability (Figure 2a). In the second series, the distribution is trimodal, with the preference parameter taking a value in either the  $[0-2]$ , the  $[9-11]$ , or the  $[18-20]$  interval (Figure 2b). This preference distribution resembles the French political spectrum, especially since 2022, and several other countries' political spectrum. Finally, the last series has a bimodal distribution, with preferences taking a value in the  $[0-5]$  or the  $[12-16]$  intervals (figure 2c). This preference distribution mirrors the bi-partisan political spectrum, for instance in the US case. Having three election series allows us to test voting behavior in a stylized case (first series) and two more realistic cases (second and third series) and assess whether the preference distribution impacts voting behavior and the election outcome (with the latter being a research question relegated to another paper).

Figure 2: Voters preferences distribution



Participants are made aware that their monetary reward depends linearly on the difference between their position and the elected candidate's position. The maximum gain, set at 20 tokens (10 euros), is achieved when there is no difference (a distance of zero) and decreases by one euro for each point of difference. Participants are informed of their potential payoff after each election, but they further know that their *actual* payoff depends only on the last election in each 5-elections series.

**Part III.** The last part consists of six questions on evaluative voting and six on majority judgment displayed to every participant, regardless of the treatment they are in. In the first three questions, we ask participants to solve some exercises to determine which of two candidates won an election based on some votes expressed through grades. There are specific circumstances in which calculating the election result for a profile works differently for the two rules, and must therefore be understood differently. One particular feature among these differences inspires the exercise we propose. A rule respects the monotonicity property if the outcome in favor of a candidate is reinforced after a change, either because the intensity of the support by a voter improves in the fixed population case, or because there are additional voters in favor of this candidate. Because evaluative voting is additive, any improvement in the support of a candidate will improve their voting outcome, and never lower their ranking. Conversely, in the variable population case, majority judgment may violate the monotonicity property since the ranking of candidates is not derived from the (sum or) average grades, but from the median grades. We therefore specifically observe this feature, by comparing participants' answers to the second and the third question. As participants will process exercises successively, first regarding evaluative voting and then majority judgment, they might remember their answers and a similar answer would not be meaningful. Hence, in order to avoid that they too easily fall into this trap, we start from a different profile in question 2 for evaluative voting and for majority judgment.

In the third question about evaluative voting, we add two voters to the nine voters from the previous question (see Table 2). Candidate B wins (with 9 points, against A who only gets 8 points). The two new voters assign high grades to both candidates: +3 to the previous winner B, and +2 to A. The candidate B receiving two +3 grades is favored and wins the new election (now with 15 points against 12 for A). Indeed, more intense support for a winner consolidates his position under evaluative voting.

Questions two and three are similar for majority judgment. However, the logic of vote aggregation with majority judgment is different. In the second exercise Candidate B wins, because B's median grade "Fairly good" is better than A's median grade "Acceptable". With two new voters, we need to examine which candidate is supported by a group of six people, likely to be composed differently than the group of five from the previous majority. And indeed, even though candidate B receives new better grades than A does, he eventually loses in the new election. Indeed, in our case, the new grades are all higher than the previous median grades ("Very good" for the previous winner and "Good" for the previous loser), lifting A's median grade but not B's. After population change, A hence wins with "Good" against B whose median grade is still "Fairly good". Questions two and three are then aimed at assessing whether participants understand how a voting outcome is differently impacted by a higher number of better grades depending on the rule.

Table 2: Second and third questions

(a) Evaluative voting						(b) Majority judgment							
Grades	-1	0	1	2	3	4	Grades	Reject	Insuff.	Accept.	Fairly Good	Good	V. Good
Cand. A	XXXX	X		XXX	XX	X	Cand. A	X	XXX	X		XXX	XXX
Cand. B	XX	XX	XX	X	XXX	X	Cand. B	XX		XX	XX		XXXXX

**Note:** Each black cross indicates a vote in the second question. Red crosses represent the additional grades given in the third question. The original profiles in the second question are meant to differ for evaluative voting and majority judgment to avoid participants’ confusion between the two exercises; however, given that the number of votes is the same, the difficulty of computation of their voting outcome is equivalent.

In the last three questions, participants choose whether a statement is true or false. These statements are focused on the final evaluation obtained for each candidate and how they are computed—specifically, whether they are calculated as the mean or the median of the grades that voters assign. Therefore, we assess whether participants can disentangle the mean and median definitions and associate them with the correct voting rule.

These comprehension questions are incentivized, meaning that participants’ payoffs depend on the number of correct answers they provide. The aim is to increase participants’ concentration and reduce the likelihood of random answers. Each correct answer to the first three questions is rewarded with 2 points, a wrong answer receives 0 points, and answering “I do not know” is worth 0.5 points. For the last three questions, participants earn 2 points if all answers are correct, 0 points if at least one answer is incorrect, and 0.5 points if participants always choose not to answer.

**Treatments** Two treatments vary the voting system used in the elections in the second part: In the *Evaluative voting* treatment (EV, one-third of the sessions), participants vote by assigning a numerical grade between  $-1$  and  $+4$  to each candidate. Additionally, in the *Majority judgment* treatment (MJ, two-third of sessions), participants vote by assigning a verbal grade between “To be rejected”, “Insufficient”, “Acceptable”, “Fairly good”, “Good”, and “Very good” to each candidate. Within the MJ treatment, two sub-treatments provide different explanations in the first part of the experiment concerning which candidate is elected when more than one candidate receives the same median grade. The explanation of the tie-breaking rule relies on the majority gauge for half of the participants and on the iterative removal process for the other half (see section 2.1 for a description of the two methods). This paper does not examine participants’ understanding of the tie-breaking rule. However, since this is arguably a challenging feature of the majority judgment, we conceive these two separate treatments as a means to measure and control for the potential impact on participants of explaining the tie-breaking rule relative to their levels of understanding. As results are robust in these two sub-treatments, and for the sake of simplicity, we pool the two treatments in the analysis.

## 3.2 Procedures

The experimental sessions were conducted in the experimental lab of the *Groupe d'Analyse et de Théorie Economique* (GATE-LAB), located in Ecully, next to Lyon, France, in 2023. We conducted nine sessions with 21 participants per session, for 189 participants. On average, participants took 1 hour and 6 minutes to complete the experiment and gained 20.03 euros, including a 5 euros show-up fee.

Women represent 53% of our sample. 91% of the participants are students, and this composition is reflected in the age composition of our sample, as 88% of participants are between 18 and 24 years old. Within the student group, 21.71% study economics, 33.71% management, and 30.29% engineering. Further, 54.29% of students come from the “Ecole Management” (a management higher education institution) and 27.43% from the “Ecole Centrale” (an engineering higher education institution), two prestigious institutions where students are required to have a high level of mathematics. For a comparison of descriptive statistics across treatments, see table A.2.1 in the Appendix.

## 4 Results

### 4.1 How to fill in a ballot

A first *non-incentivized* measure of participants’ understanding of how to fill in a ballot involves a question in the first part of the experiment on how a participant can provide maximal support to a candidate when using grades or judgments. Of course, the correct answer is assigning the maximal grade (+4) or judgment (*Very good*) to the candidate. We find that 94.71% of participants answer the question correctly for evaluative voting and 95.24% answer the question correctly for majority judgment.

A second *incentivized* measure of this understanding component is to consider the consistency of participants’ votes, similarly as in [Baujard et al. \(2018\)](#) and [Darmann and Klamler \(2023\)](#). We consider the vote participants give to all pairs of candidates in the last election of each series, where the election results impact voters’ payoff. A vote is defined as inconsistent when a participant assigns a higher grade to a candidate more distant from his position than another candidate.<sup>12</sup> We find that 94.86% of votes in the EV treatment and 96.32% of votes in the MJ treatment are consistent. These proportions are virtually the same in the first election of the series (94.87% in the EV treatment and 96.32% in the MJ treatment).<sup>13</sup> This suggests that the high level of participants’ consistency is immediate from the beginning and remains constant throughout the experiment.

If we consider the total number of inconsistent votes per participant (see Figure 3), we find that 65.08% of the sample always votes consistently in the EV treatment, and the proportion is 73.02% in the MJ treatment. Of the remaining participants, some vote inconsistently only once (20.63% in the EV treatment, 12.70% in the MJ treatment) and some more than once (14.29% in both treatments). The difference between treatments is not statistically significant (Fisher’s exact test:  $p = 0.348$ ) and is

---

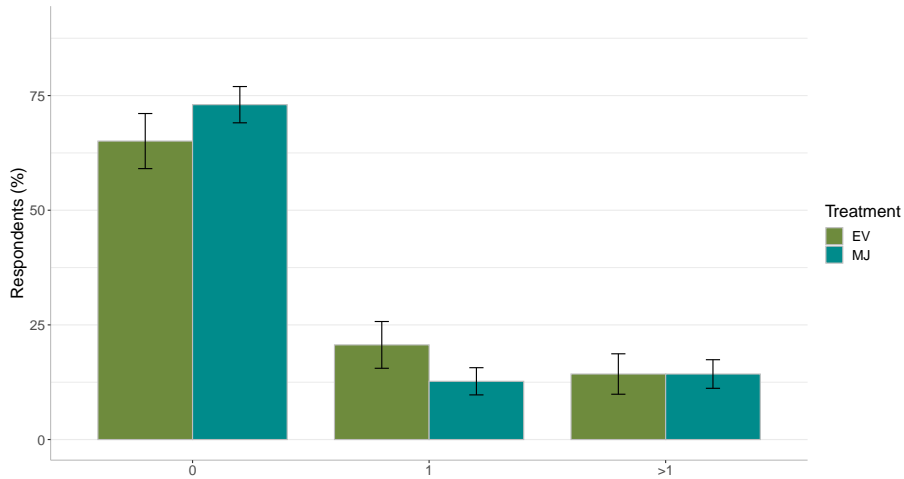
<sup>12</sup>In the case in which two candidates are equidistant (the situation of participants in positions 8, 10, and 12), a vote is defined as inconsistent if the participant gives different grades to the two candidates.

<sup>13</sup>Only four participants gave the same grade to every candidate, specifically, the same verbal grade in the MJ treatment.

mainly driven by participants voting inconsistently only once, which might result from a minor lapse in attention. Therefore, our results corroborate Hypothesis 1, and we can make the following statement:

**Result 1:** Most participants understand how to fill in a ballot with both evaluative voting and majority judgment. There are few inconsistencies, and we cannot infer a difference in inconsistent votes between the two voting rules.

Figure 3: Number of inconsistent votes, by treatment



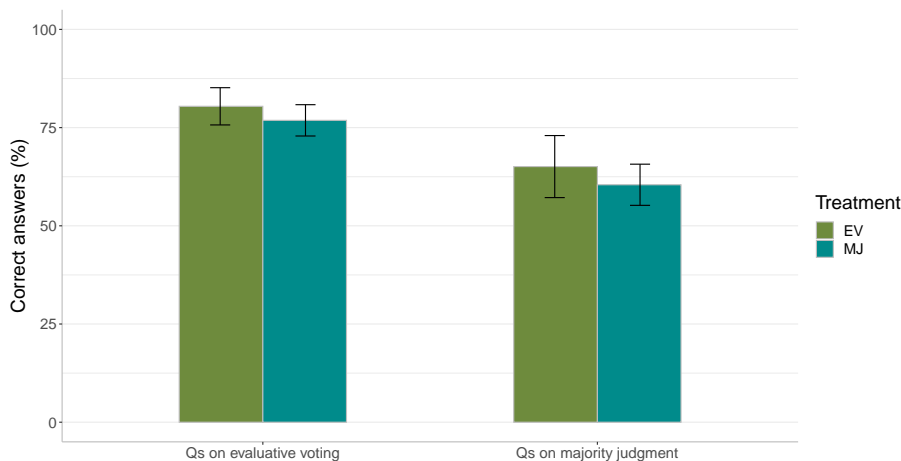
**Note:** Standard errors indicated by the bars.

## 4.2 How votes are aggregated

The second level of understanding pertains to how votes are aggregated. Figure 4 depicts the average percentage of correct answers to questions on this aspect for evaluative voting and majority judgment in each treatment (see Appendix B.8 for questions on evaluative voting and Appendix B.9 for questions on majority judgment). The average share of correct answers to questions on evaluative voting is 80.42% in the EV treatment and 76.85% in the MJ treatment, whereas for majority judgment questions it is 65.08% in the EV treatment and 60.45% in the MJ treatment.<sup>14</sup> Hence, evaluative voting exhibits a higher level of understanding than majority judgment across all treatments, with a statistically significant difference (in both treatments, two-sided paired t-test:  $p < 0.01$ ). Moreover, the comprehension of majority judgment does not improve in the MJ treatment in which participants have had the opportunity to experience this voting rule extensively in part II. If anything, these participants' comprehension is even worse, but the difference is not significant (two-sided t-test:  $p = 0.3229$ ). By contrast, participants in the EV treatment score slightly better in the questions about evaluative voting, but the difference is again not significant (two-sided t-test:  $p = 0.2814$ ).

<sup>14</sup>In both cases, the remaining answers are nearly evenly distributed between incorrect responses and participants opting not to answer. Interestingly, participants who claimed to know MJ before the experiment (27 in the last five sessions) scored even worse (56% of correct answers) than other participants.

Figure 4: Average share of correct answers to the comprehension questions



**Note:** Confidence intervals at 95% indicated by the bars.

Another notable difference in Figure 4 is the larger confidence intervals in the share of correct answers for questions on majority judgment compared to those on evaluative voting. Indeed, the standard deviation of correct answers to the majority judgment questions is significantly higher than in the evaluative voting questions (F-test:  $p < 0.01$ ). Figure A.1.1 in the Appendix shows the distribution of correct answers (and Figure A.1.2 the distribution of wrong answers). For both the questions on evaluative voting and on majority judgment and across the two treatments, the distribution mode is 100% correct answers. However, only 6 participants (3.17%) answer correctly less than 50% of the time to questions on evaluative voting, while this is the case for 52 participants (27.51%) for majority judgment questions. In other words, there is greater polarization in the comprehension of majority judgment. To summarize:

**Result 2:** Most people understand how votes are aggregated in evaluative voting and majority judgment. However, participants' understanding of majority judgment is, on average, lower and more heterogeneous compared to their understanding of evaluative voting. Experience with the voting rules does not improve their understanding.

The next step in the analysis is to understand which questions were difficult for participants. In the first three questions on the results of an election where participants can see the voters' votes for two candidates, participants perform remarkably well when answering about evaluative voting (94% of correct answers, on average). In contrast, they perform less well answering questions on majority judgment (74.25% of correct answers, on average). In particular, subjects are much more likely to be wrong in their response to the third question on majority judgment (61.38% correct answers) than in the case of evaluative voting (94.71% correct answers). As explained in section 3.1, the third question is aimed at evaluating participants' understanding of the monotonicity property when the set of voters varies. This indicates that subjects believe both evaluative voting and majority judgment rules to be



monotonic, whereas in fact monotonicity is a feature only of evaluative voting.

Participants score generally worse in the True and False questions (see table A.3.1 in the Appendix). In the questions about evaluative voting, the average share of correct answers is 62.08%, while the share is 49.74% for questions on majority judgment. The worst score for questions on evaluative voting pertains to the statement “The grade corresponds to the average score given to this candidate by all voters” (47.62% correct answers). However, participants’ confusion likely arises from the fact that, in the experiment, the grade is presented as the sum rather than the average, except once at the beginning of the experiment. The decision to use the average in the statement was influenced by the need to compare this statement to the equivalent one in the case of majority judgment. The worst score for majority judgment pertains to the last question, where only 43.39% of participants confirm that the majority grade given to a candidate “Corresponds to the judgment given by the voter who divides in two the set of voters ranked from the most favorable to the most unfavorable”. This suggests that participants often fail to understand the definition of the median, confirming our Hypothesis 2:

**Result 2bis:** The most difficult barrier to understanding how votes are counted in majority judgment is its definition of median and its implication in terms of vote aggregation.

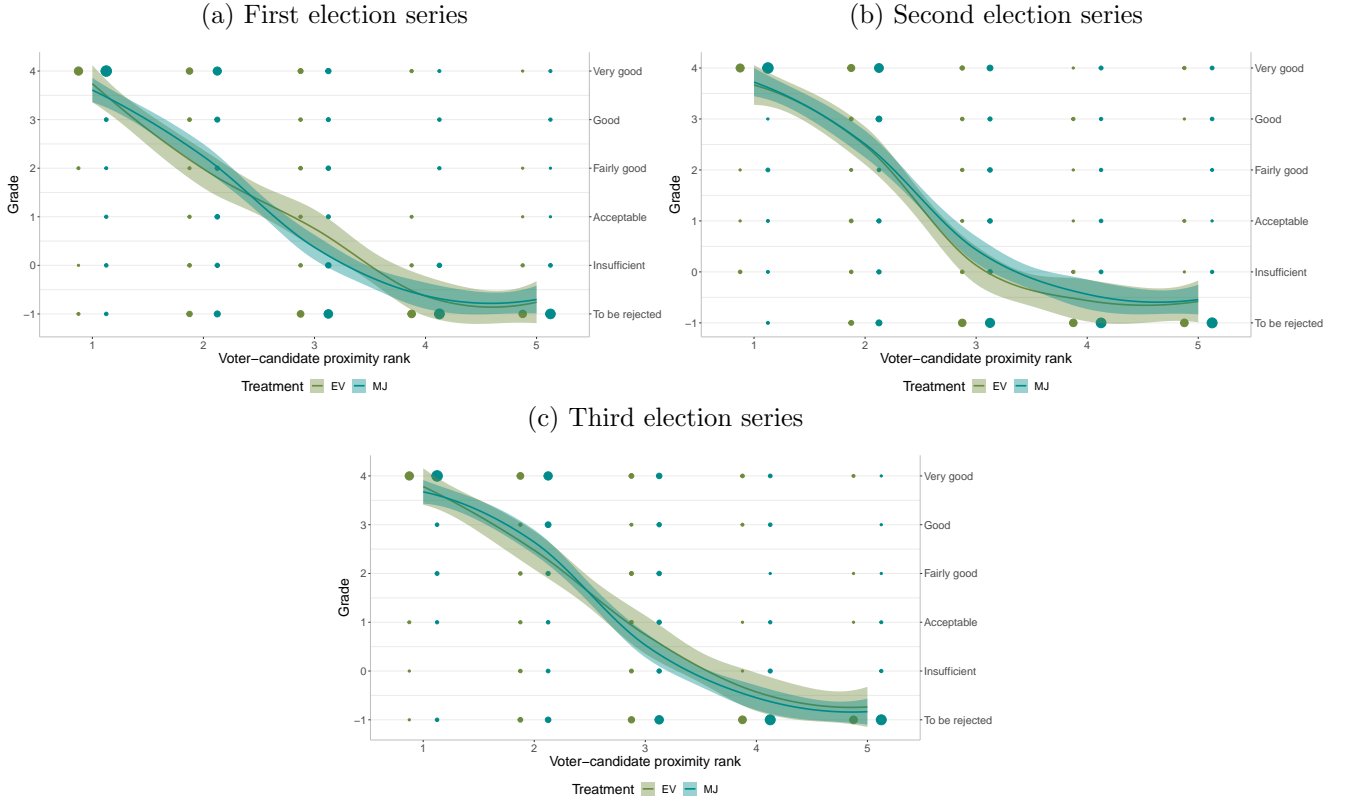
### 4.3 How to vote strategically

To investigate participants’ voting behavior, we first define a voter–candidate proximity rank (1 to 5) based on the distance between a participant and each candidate.<sup>15</sup> Figure 5 shows the relationship between the voter–candidate proximity rank and the grades participants assigned to each candidate in the fifth election of each series. We observe no differences in voting behavior across treatments. First, participants assign the highest grade to the closest candidate so that, by pooling the three election series, 92.06% of votes are +4 in the EV treatment and 89.15% are “Very good” in the MJ treatment. Second, participants gradually decrease grades for candidates ranked second and third while providing intermediate grades. Finally, the majority of grades assigned to the fourth and fifth furthest candidates are the minimum, either  $-1$  (83.07% to the fourth and 91.01% to the fifth candidate) or *To be rejected* (80.95% to the fourth and 89.42% to the fifth candidate).

---

<sup>15</sup>If, for one participant, two candidates are equidistant, they display the same rank.

Figure 5: Voting behavior



**Note:** We consider the vote in only the fifth election in each election series. The voter-candidate proximity rank takes the value 1 if the candidate is the closest to a participant’s position, 2 if the candidate is the second closest to the participant’s position, etc. The lines represent the fitted values of local regressions. The shaded areas indicate confidence intervals at 95%.

Next, we count the number of extreme votes per election and assess participants’ learning throughout the election series. In Table 3, we show the share of extreme grades in the EV and MJ treatments in each election series and repetition in each series. In the EV treatment, the share of extreme grades evolves from 66.98% to 86.03% in the first series, from 77.78% to 85.08% in the second, and from 85.08% to 86.35% in the third. Similarly, in the MJ treatment, the share increases from 63.65% to 79.68% in the first series, from 76.51% to 81.90% in the second series, and from 80.95% to 82.22% in the third. Thus, with both voting rules, the number of extreme votes is already high from the first election in each series and increases as elections progress.

Table 3: Extreme votes and learning

	First series					Second series					Third series				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
EV	66.98	75.87	82.22	84.44	86.03	77.78	81.90	83.49	86.03	85.08	85.08	85.71	84.13	85.71	86.35
MJ	63.65	69.21	73.49	76.83	79.68	76.51	78.89	80.79	81.27	81.90	80.95	80.48	81.27	81.11	82.22

**Note:** Each value represents the share of votes equal to +4 or -1 in the EV treatment, and “Very good” or “To be rejected” in the MJ treatment.

Finally, since participants’ voting behavior likely depends on the other participants’ voting behavior in their session, a robust test to examine differences in voting behavior across treatments involves regressing the number of times participants assign an extreme grade in the election on the treatment variable while clustering standard errors at the session level. We run this regression for the last election of every election series, and the three election series pooled to maximize statistical power. We also control for participants’ position fixed effects, participants’ share of correct answers to the questions on evaluative voting and majority judgment, and participants’ gender. The results of this analysis are presented in Table 4, indicating no significant effect of the treatment on the choice. Besides the lack of statistical significance, the treatment effect is also low compared to the share of extreme grades in the EV treatment. Indeed, adding the treatment effect to this share gives us an average number of extreme grades between 4.05 and 4.15 in the MJ treatment. Summing up, our data do not provide support for more sincere voting with majority judgment (Hypothesis 3), hence:

**Result 3:** Participants vote the same way in the EV and MJ treatments, with most participants assigning only extreme grades.

Table 4: Regression analysis - Extreme votes

	1st series	2nd series	3rd series	All series pooled
	(1)	(2)	(3)	(4)
MJ	-0.25 (0.2)	-0.099 (0.1)	-0.15 (0.2)	-0.17 (0.1)
Constant	3.53*** (0.8)	2.51*** (0.7)	3.81*** (0.6)	3.32*** (0.4)
Average extreme votes (EV)	4.30	4.25	4.32	4.29
Observations	189	189	189	567
R-squared	0.28	0.16	0.12	0.15

**Note:** OLS regression. The dependent variable is the number of times a participant assigns an extreme grade in an election. This ranges between 0 and 5 in columns (1) to (3) and between 0 and 15 in column (4). We control for participants’ position fixed effects, gender, and share of correct answers to the questions on evaluative voting and majority judgment. Standard errors clustered at the session level in parentheses. \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

*A priori*, there is no straightforward reason why voters should not vote sincerely with majority judgment. Result 3 might either capture participants’ bad understanding of majority judgment, that they follow a simple costless heuristics, or that they face an opportunity for manipulation. For the latter point, the possibility to manipulate the majority judgment is discussed by [Felsenthal and Machover \(2008\)](#) and [Laslier \(2019\)](#): If there is a tie between two candidates, a voter can influence the outcome by exaggerating his preferences. This means assigning a grade above the median to the favored candidate and below the median to the other. Of course, this does not require using extreme grades, as long as the grades are above or below the median. However, using extreme grades increases the likelihood of a grade being above or below the median, provided voters do not have fixed beliefs about other voters’ support for the candidates. To disentangle these three interpretations, in the next subsection we assess whether participants using extreme grades perform well in other components of understanding, particularly in

questions about aggregation by the median.

#### 4.4 Correlation between the levels of understanding

We now study whether the three levels of understanding are correlated, shedding light also on participants' voting behavior. To do this, we use three variables at the participant level: the number of inconsistent choices in the last election of the three series, the share of correct answers to questions on either evaluative voting or majority judgment, and the number of extreme votes the candidate gave to each candidate. Figure A.4.1 in the Appendix presents the correlation matrix where we include these three variables and split the sample by treatment. The strongest correlation pertains to participants giving extreme grades, who tend to vote consistently in both the EV and MJ treatments. Moreover, in the EV treatment, there is no significant nor strong correlation between participants' understanding of the aggregation properties of evaluative voting and their consistency and extreme voting. This result is likely driven by the extremely low variation of the share of correct answers (*i.e.*, most participants understand well, see Figure A.1.1). By contrast, in the MJ treatment, participants answering correctly to the questions about vote aggregation by median are less likely to vote inconsistently and more likely to assign extreme grades.

The relationship between participants' understanding of vote aggregation and their voting behavior is further analyzed in Table 5, where we regress the number of extreme grades per election in all election series on participants' share of correct answers to the questions on the majority judgment and evaluative voting. A deeper understanding of vote aggregation under evaluative voting does not increase extreme grading. However, as already noted, the low variation in people's understanding of how votes are aggregated likely explains the absence of correlation. By contrast, in the MJ treatment, participants who understand vote aggregation under majority judgment assign extreme grades more frequently ( $p < 0.01$ ).

**Result 4:** When voting with majority judgment, participants who understand its vote aggregation method use extreme grades more frequently.

Result 4 is in some sense counterintuitive, as we would expect voters who understand the properties of the median to use intermediate grades more often, as they understand that overstating their preferences is not useful as a means to impact the election outcome. A possible explanation of participants' use of extreme grades under majority judgment is that these more sophisticated participants perceive majority judgment to be manipulable. Considering that there are some manipulable situations (as discussed above), these sophisticated voters understand that extreme grading is sometimes an optimal strategy, and, in general, a costless heuristic. When voters do not understand well vote aggregation under majority judgment, the fact that they use mostly extreme grades reveals that they consider that majority judgment shares the same properties of evaluative voting. Hence, the sophisticated voters understand the third component while the others do not.

Finally, we need to highlight a limit of our analysis: Assigning only extreme grades to fictitious candidates with assigned preferences in the lab has no cost in terms of preference expression, while in

real-life elections voters might want to express their nuanced opinions (Igersheim et al., 2016).

Table 5: Regression analysis - Extreme votes and understanding

	Treatment: EV	Treatment: MJ
	(1)	(2)
% correct answers EV	0.0029 (0.01)	0.0081 (0.005)
% correct answers MJ	0.0039 (0.005)	0.0083*** (0.002)
Average extreme votes	4.29	4.06
Observations	189	378
R-squared	0.18	0.18

**Note:** OLS regression. The dependent variable is the number of times a participant assigns an extreme grade in an election. This ranges between 0 and 5. We control for participants’ position fixed effects and gender. Standard errors clustered at the session level in parentheses. \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

## 5 Conclusion

Philosophers of science have pointed out that the epistemic values of scientific assertions should be scrutinized in practice and not just in theory. As Ankeny et al. (2011) have remarked, in describing this concern with practice in the context of philosophy of science, “Our concern is not only about how pre-existing knowledge gets applied to practical ends but also about how knowledge itself is fundamentally shaped by its intended uses.” Science may conceive tools, devices, or public decisions in isolation of any operational constraints on their application, but taking practice seriously requires scrutinizing how contextual information may *also* affect the validity of scientific assertions. In particular, when science seeks to establish normative propositions, it is all the more important to check whether the context and the stage of application may affect them, and how, as illustrated by Baujard (2022) in the domain of the evaluation of public policies. This paper adapts this latter approach to the case of voting rules, highlighting the assumption that voters understand the rules as a crucial factor that needs to be tested empirically. To this end, we leverage an experimental design where participants learn and practice voting rules several times to analyze people’s understanding of the evaluative voting and majority judgment voting rules. We distinguish three levels of understanding: how to fill in a ballot, how votes are aggregated, and how to vote strategically. We provide experimental measures for each of these and analyze how these levels of understanding differ for the two rules.

The experimental data we gathered allow us to highlight four main results. First, participants vote consistently with both tested voting rules. Second, most voters understand how votes are counted. However, participants’ understanding of majority judgment is, on average, lower and more heterogeneous than their understanding of evaluative voting because of their difficulties with defining the median and grasping its implications for the voting outcome. Third, we show that participants vote similarly with

evaluative voting and majority judgment. The data confirm the theoretical prediction that under evaluative voting there would be more strategic voting through the use of extreme grades, but it contradicts the prediction of more sincere voting under majority judgment. Fourth, those participants who understand the system of vote aggregation with majority judgment are more likely to assign only extreme grades. This suggests that majority judgment is perceived to be manipulable, but also refers to the possibility of its being manipulated by those “super voters” who best understand how median rules operate. By contrast, the fact that participants who did not understand vote aggregation under majority judgment still used only extreme grades suggests their confusion about the properties of majority judgment.

Overall, we highlight that there are sources of confusion when voters vote under majority judgment. For three reasons, we believe these measures of misunderstanding represent a lower bound. First, our sample of participants is mostly composed of bright students. While this selection bias is often used as a critique of lab experiments, in our case it reinforces the claim that more representative samples would exhibit higher levels of misunderstanding. Second, the participants in our sample were in an ideal situation to learn voting rules, practice them several times, and to vote with clear monetary incentives. This situation is difficult to replicate in real life, as people vote only rarely and might pay little attention to information given about new rules. Third, the current protocol did not measure any understanding of the tie-breaking rule. Should a protocol include the tie-breaking rule, which is a major component of majority judgment, the measure of its misunderstanding would most likely be even higher.

There are numerous potential avenues for future investigation. First, our methodology can be easily applied to other voting rules to test the relationship between voters’ understanding of these rules’ properties and their voting behavior. Second, a dimension not tackled here is which normative properties individuals associate with different rules (*e.g.*, which candidates will be favored by each voting rule). This line of inquiry is promising as it could help explain voters’ preferences for different voting rules and their possible reluctance to adopt new ones. Third, it would be possible to include a benchmark treatment with limited information on the rules before individuals get to vote, in order to study the causal impact of voters’ understanding on their behavior. Finally, we remark that the whole of our society is pervaded by rules of all kinds: Some rules are designed facilitate the assessment or analysis of states of affairs, others to aid in individual decisions, and others still to frame collective decision-making. We observe that there is a need for more research focused on the layman’s understanding of rules.

## References

- Alekseev, A., Charness, G., and Gneezy, U. (2017). Experimental Methods: When and Why Contextual Instructions are Important. *Journal of Economic Behavior & Organization*, 134:48–59.
- Andreoni, J. (1995). Cooperation in Public-Goods Experiments: Kindness or Confusion? *The American Economic Review*, 85(4):891–904.
- Ankeny, R., Chang, H., Boumans, M., and Boon, M. (2011). Introduction: Philosophy of Science in Practice. *European Journal for Philosophy of Science*, 1(3):303.
- Balinski, M. and Laraki, R. (2007). A Theory of Measuring, Electing, and Ranking. *Proceedings of the National Academy of Sciences*, 104(21):8720–8725.
- Balinski, M. and Laraki, R. (2011). Election by Majority Judgment: Experimental Evidence. In Dolez, B., Grofman, B., and Laurent, A., editors, *In Situ and Laboratory Experiments on Electoral Law Reform: French Presidential Elections*, Studies in Public Choice, pages 13–54. Springer, New York, NY.
- Balinski, M. and Laraki, R. (2012). Jugement majoritaire versus vote majoritaire. (via les présidentielles 2011-2012). *Revue française d'économie*, 27(4):11–44.
- Balinski, M. and Laraki, R. (2020). Majority Judgment vs. Majority Rule. *Social Choice and Welfare*, 54(2):429–461.
- Baujard, A. (2022). Ethics and Technique in Welfare Economics: How Welfarism Evolves in the Making. *Revue économique*, 73(6):1039–1053.
- Baujard, A., Gavrel, F., Igersheim, H., Laslier, J.-F., and Lebon, I. (2012). Compte-rendu de l'expérimentation des votes par approbation et par note lors des élections présidentielles françaises le 22 avril 2012 à Saint-Etienne, Strasbourg et Louvigny. *Working Papers*.
- Baujard, A., Gavrel, F., Igersheim, H., Laslier, J.-F., and Lebon, I. (2018). How voters use grade scales in evaluative voting. *European Journal of Political Economy*, 55:14–28.
- Baujard, A. and Igersheim, H. (2009). Expérimentation du vote par note et du vote par approbation le 22 avril 2007. Premiers résultats. *Revue économique*, 60(1):189–201.
- Baujard, A., Igersheim, H., and Lebon, I. (2021). Some regrettable grading scale effects under different versions of evaluative voting. *Social Choice and Welfare*, 56(4):803–834.
- Baujard, A., Igersheim, H., Lebon, I., Gavrel, F., and Laslier, J.-F. (2014). Who's favored by evaluative voting? An experiment conducted during the 2012 French presidential election. *Electoral Studies*, 34:131–145.
- Bigoni, M. and Dragone, D. (2012). Effective and efficient experimental instructions. *Economics Letters*, 117(2):460–463.

- Blais, A., Plescia, C., and Sevi, S. (2021). Do people want smarter ballots? *Research & Politics*, 8(4).
- Blais, A., St-Vincent, S. L., Pilet, J.-B., and Treibich, R. (2016). Voting correctly in lab elections with monetary incentives: The impact of district magnitude. *Party Politics*, 22(4):544–551.
- Bourgeois-Gironde, S. and Ferreira, J. V. (2023). The expressive power of voting rules. *Social Choice and Welfare*, 62:233–273.
- Bowler, S., McElroy, G., and Müller, S. (2018). Voter preferences and party loyalty under cumulative voting: Political behaviour after electoral reform in Bremen and Hamburg. *Electoral Studies*, 51:93–102.
- Brams, S. J. and Fishburn, P. C. (1978). Approval Voting. *American Political Science Review*, 72(3):831–847.
- Brams, S. J. and Fishburn, P. C. (2002). Chapter 4 Voting procedures. In *Handbook of Social Choice and Welfare*, volume 1 of *Handbook of Social Choice and Welfare*, pages 173–236. Elsevier.
- Cormack, L. (2023). More Choices, More Problems? Ranked Choice Voting Errors in New York City. *American Politics Research*, 52(3):306–319.
- Darmann, A., Grundner, J., and Klamler, C. (2017). Election outcomes under different ways to announce preferences: an analysis of the 2015 parliament election in the Austrian federal state of Styria. *Public Choice*, 173(1):201–216.
- Darmann, A., Grundner, J., and Klamler, C. (2019). Evaluative voting or classical voting rules: Does it make a difference? Empirical evidence for consensus among voting rules. *European Journal of Political Economy*, 59:345–353.
- Darmann, A. and Klamler, C. (2023). Does the rule matter? A comparison of preference elicitation methods and voting rules based on data from an Austrian regional parliamentary election in 2019. *Public Choice*, 197(1):63–87.
- Dellis, A. and Kröger, S. (2023). How Do People Vote Under Instant Runoff Voting? An Experiment on Complexity and Voting Behavior. Available at SSRN: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4602773](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4602773).
- Donovan, T., Tolbert, C., and Gracey, K. (2019). Self-Reported Understanding of Ranked-Choice Voting. *Social Science Quarterly*, 100(5):1768–1776.
- Donovan, T., Tolbert, C., and Harper, S. (2022). Demographic differences in understanding and utilization of ranked choice voting. *Social Science Quarterly*, 103(7):1539–1550.
- Fabre, A. (2021). Tie-breaking the Highest Median: Alternatives to the Majority Judgment. *Social Choice and Welfare*, 56(1):101–124.



- Felsenthal, D. S. and Machover, M. (2008). The Majority Judgement Voting Procedure: A Critical Evaluation. *Homo oeconomicus*, 25 (3/4):319–334.
- Freeman, D. J., Kimbrough, E. O., Petersen, G. M., and Tong, H. T. (2018). Instructions. *Journal of the Economic Science Association*, 4(2):165–179.
- Herzberg, R. Q. and Wilson, R. K. (1988). Results on Sophisticated Voting in an Experimental Setting. *The Journal of Politics*, 50(2):471–486.
- Hillinger, C. (2006). The Case for Utilitarian Voting. *Munich Discussion Paper, No. 2005-11*, Ludwig-Maximilians-Universität München, Volkswirtschaftliche Fakultät, München, <https://doi.org/10.5282/ubm/epub.653>.
- Houser, D. and Kurzban, R. (2002). Revisiting Kindness and Confusion in Public Goods Experiments. *American Economic Review*, 92(4):1062–1069.
- Igersheim, H., Baujard, A., Gavrel, F., Laslier, J.-F., and Lebon, I. (2016). Individual Behavior Under Evaluative Voting: A Comparison Between Laboratory and In Situ Experiments. In Blais, A., Laslier, J.-F., and Van Der Straeten, K., editors, *Voting Experiments*, pages 257–269. Springer International Publishing, Cham.
- Igersheim, H., Baujard, A., and Laslier, J.-F. (2017). La question du vote. Expérimentations en laboratoire et In Situ. *L'Actualité économique*, 92(1-2):151–189.
- Igersheim, H., Durand, F., Hamlin, A., and Laslier, J.-F. (2022). Comparing voting methods: 2016 US presidential election. *European Journal of Political Economy*, 71:102057.
- Kimball, D. C. and Kropf, M. (2016). Voter Competence with Cumulative Voting\*. *Social Science Quarterly*, 97(3):619–635.
- Lachat, R. and Laslier, J.-F. (2024). Alternatives to plurality rule for single-winner elections: When do they make a difference? *European Journal of Political Economy*, 81:102505.
- Laslier, J.-F. (2011). Lessons from In Situ Experiments during French Elections. In Dolez, B., Grofman, B., and Laurent, A., editors, *In Situ and Laboratory Experiments on Electoral Law Reform: French Presidential Elections*, Studies in Public Choice, pages 91–104. Springer, New York, NY.
- Laslier, J.-F. (2019). The strange “Majority Judgment”. *Revue économique*, 70(4):569–588.
- Laslier, J.-F., Blais, A., Bol, D., Golder, S. N., Harfst, P., Stephenson, L. B., and Van der Straeten, K. (2015). The EuroVotePlus experiment. *European Union Politics*, 16(4):601–615.
- Lau, R. R. and Redlawsk, D. P. (1997). Voting Correctly. *American Political Science Review*, 91(3):585–598.

- Macé, A. (2018). Voting with evaluations: Characterizations of evaluative voting and range voting. *Journal of Mathematical Economics*, 79:10–17.
- Marino, M., Iacono, R., and Mollerstrom, J. (2024). (Mis-)Perceptions, information, and political polarization: A survey and a systematic literature review. *European Journal of Political Economy*, 85:102578.
- Núñez, M. and Laslier, J. F. (2014). Preference intensity representation: strategic overstating in large elections. *Social Choice and Welfare*, 42(2):313–340.
- Oprea, R. (2020). What Makes a Rule Complex? *American Economic Review*, 110(12):3913–3951.
- Pulickal, A. and Chakravarty, S. (2023). Subject confusion and task non-completion: Methodological insights from an artefactual field experiment with adolescents in India. *Journal of Behavioral and Experimental Economics*, 103:101986.
- Ramalingam, A., Morales, A. J., and Walker, J. M. (2018). Varying experimental instructions to improve comprehension: Punishment in public goods games. *Journal of Behavioral and Experimental Economics*, 73:66–73.
- Regenwetter, M., editor (2006). *Behavioral social choice: probabilistic models, statistical inference, and applications*. Cambridge University Press, Cambridge ; New York.
- Regenwetter, M., Grofman, B., Popova, A., Messner, W., Davis-Stober, C. P., and Cavagnaro, D. R. (2008). Behavioural social choice: a status report. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1518):833–843.
- Regenwetter, M., Marley, A., and Grofman, B. (2003). General concepts of value restriction and preference majority. *Social Choice and Welfare*, 21(1):149–173.
- Regenwetter, M. and Tsetlin, I. (2004). Approval voting and positional voting methods: Inference, relationship, examples. *Social Choice and Welfare*, 22(3):539–566.
- Rozenblit, L. and Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, 26(5):521–562. Publisher: John Wiley & Sons, Ltd.
- Sen, A. (1995). How to Judge Voting Schemes. *The Journal of Economic Perspectives*, 9(1):91–98.
- Slooman, S. and Fernbach, P. (2017). *The Knowledge Illusion: Why We Never Think Alone*. Riverhead Books, New York, NY.
- Smaoui, H. and Lepelley, D. (2013). Le système de vote par note à trois niveaux : étude d’un nouveau mode de scrutin. *Revue d’économie politique*, 123(6):827–850.
- Van Der Straeten, K., Laslier, J.-F., Sauger, N., and Blais, A. (2010). Strategic, sincere, and heuristic voting under four election rules: an experimental study. *Social Choice and Welfare*, 35(3):435–472.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, 13(1):75–98.

# Appendices

## Table of Contents

---

<b>A</b>	<b>Additional figures and tables</b>	<b>28</b>
A.1	Figures on understanding . . . . .	28
A.2	Summary statistics . . . . .	30
A.3	Answers to True/False questions . . . . .	31
A.4	Correlation among levels of understanding . . . . .	32
<b>B</b>	<b>Experiment details</b>	<b>33</b>
B.1	Session presentation . . . . .	33
B.2	Contextualized experience . . . . .	34
B.3	Other voting methods and contextualized application . . . . .	35
B.4	Introduction to decontextualized experimentat . . . . .	42
B.5	The vote with majority judgment . . . . .	45
B.6	The vote with evaluative voting . . . . .	46
B.7	Comprehension - Presentation . . . . .	47
B.8	Comprehension - Evaluative voting . . . . .	47
B.9	Comprehension - Majority judgment . . . . .	48
B.10	Comprehension - Approval voting . . . . .	49
B.11	Satisfaction and assessment of voting methods . . . . .	50
B.12	Identity . . . . .	50
B.13	Remuneration . . . . .	51
B.14	Last screen . . . . .	51

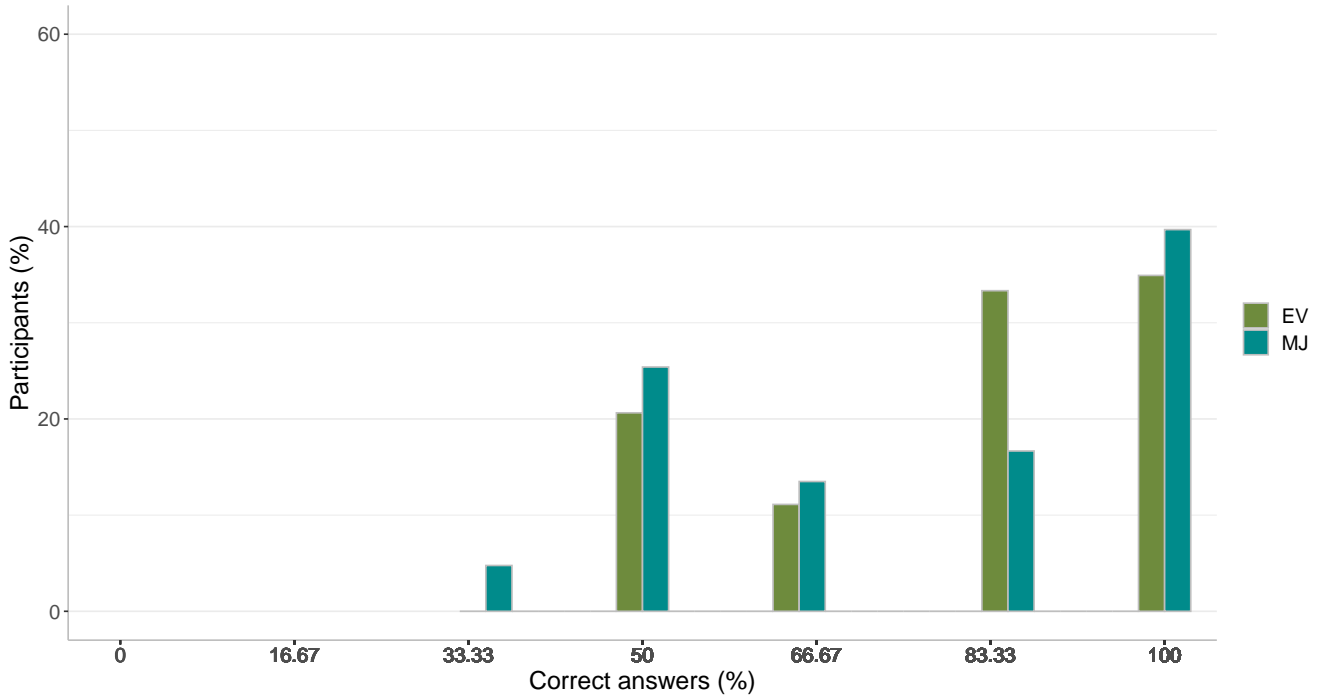
---

# A Additional figures and tables

## A.1 Figures on understanding

Figure A.1.1: Distribution of share of correct answers

(a) Questions on evaluative voting



(b) Questions on majority judgment

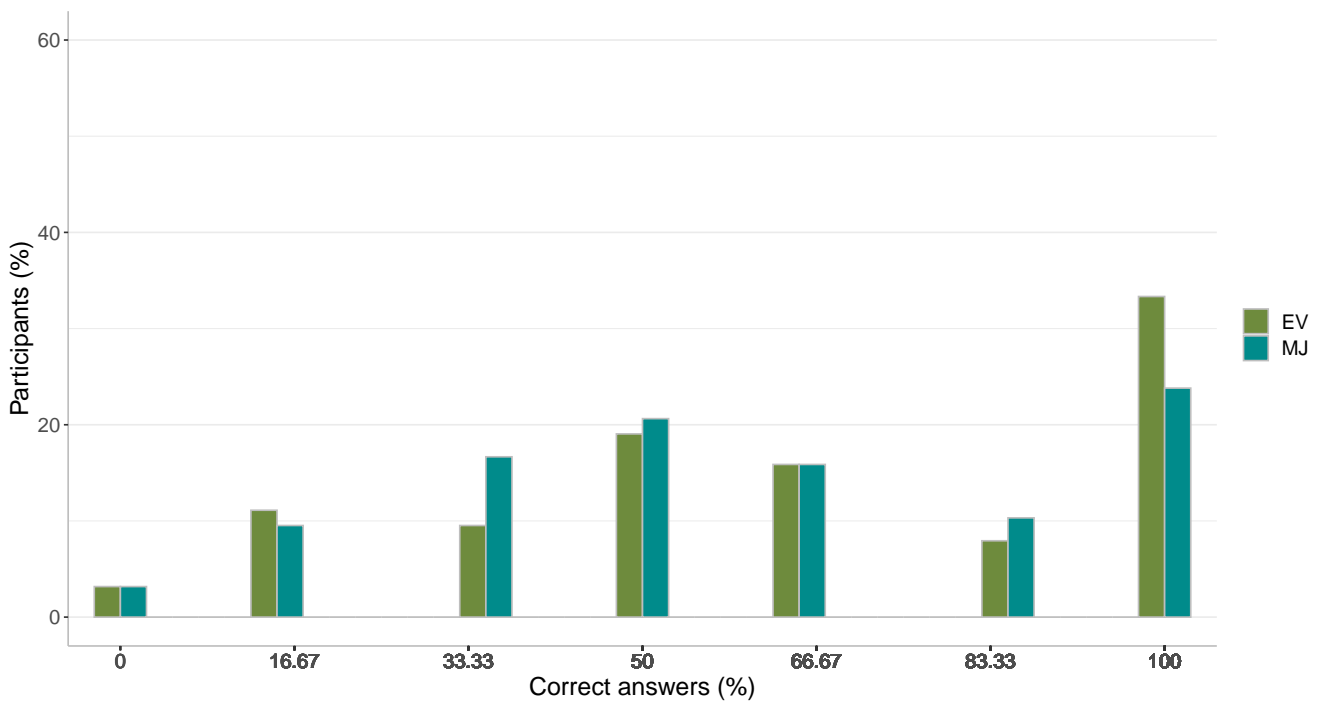
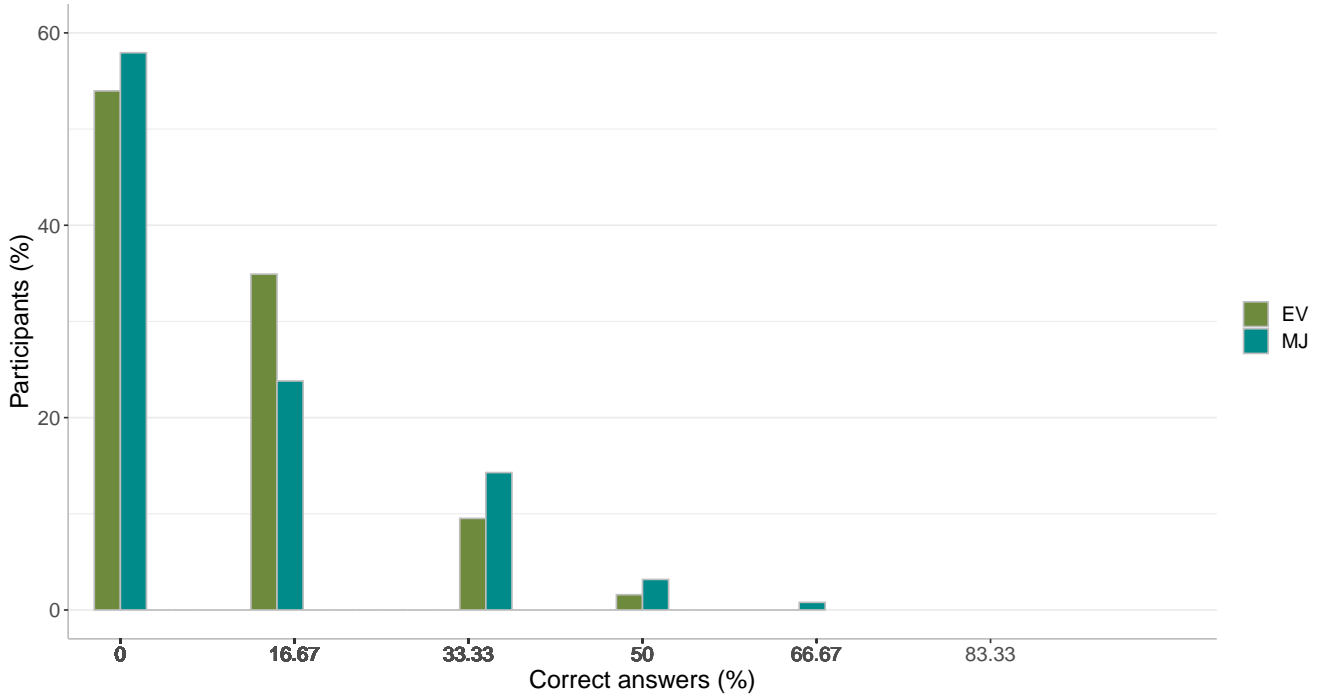
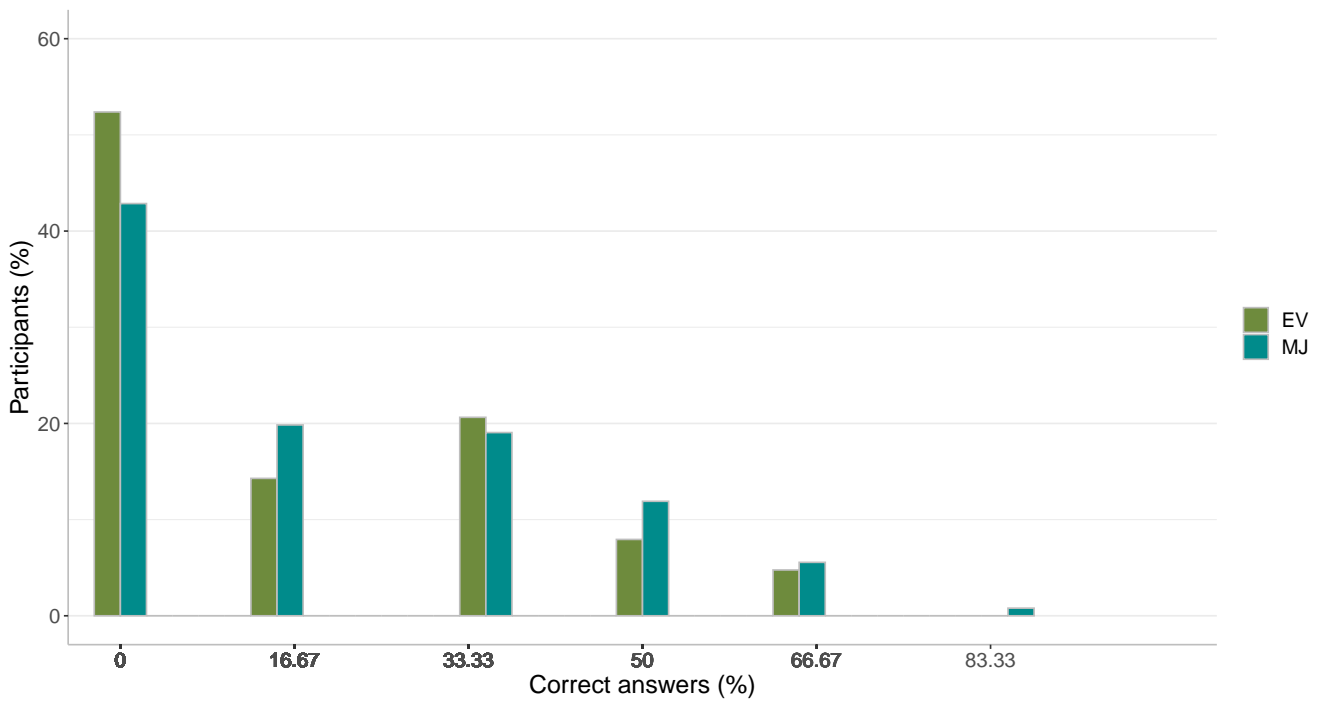


Figure A.1.2: Distribution of share of wrong answers

(a) Questions on evaluative voting



(b) Questions on majority judgment



## A.2 Summary statistics

Table A.2.1: Summary statistics by treatment group

	Treatment EV (N=63)	Treatment JM-G (N=63)	Treatment JM-I (N=63)
<i>Gender</i>			
Female	0.54	0.51	0.54
Male	0.46	0.49	0.46
<i>Age</i>			
[18-24]	0.94	0.83	0.87
[24+]	0.063	0.17	0.13
<i>Occupation</i>			
Student	0.95	0.86	0.92
Other	0.048	0.14	0.079
<i>Field of study</i>			
Economics and business	0.44	0.51	0.33
Engineering	0.37	0.22	0.33
Other	0.19	0.27	0.33
<i>Higher education institution</i>			
Ecole management Lyon	0.51	0.71	0.45
Ecole centrale	0.39	0.16	0.28
Other	0.11	0.13	0.27
<i>Agglomeration size</i>			
Rural municipality	0.17	0.063	0.17
2,000-19,999 inhabitants	0.14	0.19	0.21
20,000-99,999 inhabitants	0.21	0.13	0.16
More than 100,000 inhabitants	0.30	0.37	0.32
Paris agglomeration	0.17	0.25	0.14

### A.3 Answers to True/False questions

Table A.3.1: True/false questions on preferences aggregation

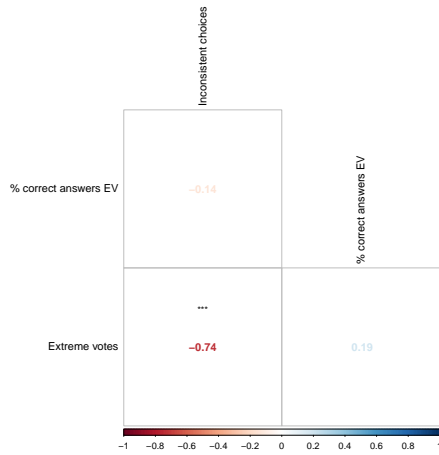
<b>Questions on evaluative voting</b>	Correct answer	Correct	Wrong	Did not know
The grade corresponds to the average score given to this candidate by all voters	True	47.62%	32.28%	20.11%
The grade corresponds to the average rating given to this candidate by half plus one of the voters who are most in favor of him/her.	False	70.90%	5.29%	23.81%
The grade corresponds to the score given by the voter who divides in two the set of voters ranked from the most favorable to the most unfavorable.	False	67.72%	10.05%	22.22%
<b>Questions on majority judgment</b>	Correct answer	Correct	Wrong	Did not know
The majority grade corresponds to the average rating given to this candidate by all voters	False	51.85%	18.52%	29.63%
The majority grade corresponds to the average rating given to this candidate by half plus one of the voters most in favor of him/her	False	53.97%	14.29%	31.75%
The majority grade corresponds to the judgment given by the voter who divides in two the set of voters ranked from the most favorable to the most unfavorable	True	43.39%	25.40%	31.22%



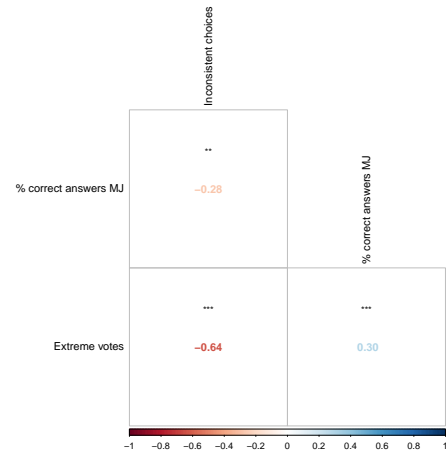
## A.4 Correlation among levels of understanding

Figure A.4.1: Correlation matrix - Levels of understanding

(a) EV Treatment



(b) MJ-G and MJ-I treatments



**Note:** The variable “Inconsistent choices” represents the number of inconsistent choices in the three last elections of every series. “Extreme votes” represents the sum of the extreme votes given by each participant in the three last elections of every series. % correct answers EV (MJ) is the share of correct answers to the questions on how votes are aggregate in evaluative voting (majority judgment). \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

## B Experiment details

This section contains a French-to-English translation of the experimental instructions. We highlight changes from one treatment to another by bold characters at the beginning of a chunk of text. For instance, the mention **[EV]** indicates that the next chunk of text exposes the instructions for the *Evaluative voting* treatment. Note that evaluative voting has been named grade voting (*vote par note*) in the experiment. Answer options are in *italic*, separated by a semicolon.

### B.1 Session presentation

#### B.1.1 Purpose of the study

A team of researchers (Antoinette Baujard, Isabelle Lebon, Simone Marsilio) is testing different voting systems.

This experiment takes just over an hour and its aim is purely scientific. The data collected is completely anonymous.

The voting methods to be tested will be explained as the experiment progresses.

#### B.1.2 Study procedure

This experiment will consist of three parts:

- The discovery of alternative voting methods by testing them in a real-life situation.
- Mock elections for any candidates, to test these alternative voting methods.
- Questions on the different voting methods tested.

Participation in this experiment is remunerated. For every 2 points obtained, you will earn 1 euro. Your remuneration is made up of:

- A fixed payoff of 10 points.
- A variable payoff of 0 to 20 points depending on your decisions in the mock election series
- A variable payoff of 0 to 20 points depending on your answers to 10 questions on voting methods

You have the right to withdraw freely before the end of the experiment without having to justify yourself, but you will not be entitled to your gratification.

If you need any help during the experiment, ask the experimentalist in the room.

Are you ready to begin? *Yes*

## B.2 Contextualized experience

### B.2.1 The context

The aim of this first series of questions is to enable you to experience new voting methods in a given context of collective decision-making. You can vote according to your personal tastes.

To do this, we ask you to imagine a concrete situation: you are a member of a condominium assembly, which has to decide on a single color for all the walls in the common areas (knowing that the floors and doors are made of untreated wood). The condominium manager proposes a list of candidate colors, which the condominium meeting then votes on.

Here are the possible colors:



From now on, you will be able to express your opinion and vote for a color according to different voting methods.

### B.2.2 The expression of opinion

The purpose of this question is to gather your opinion on these different colors. Your answer will not be used to decide on the color of the walls.



Evaluate each of the 8 candidate colors by giving them a score between 0 and 20 by moving the slider. 0 is the worst score, 20 is the best. If you can't or don't want to comment on a color, tick the "I have no opinion" box. *Slider between 0 and 20 for each color*

### B.2.3 The "two-round" voting method. Presentation

The official voting method is called the "two-round" system. This is the method used, for example, in the French presidential election.

Each elector votes for a single candidate color. For this color to be chosen, it must obtain an absolute majority in the first round. If no color exceeds the 50% threshold in the first round, the two candidate colors with the highest scores in the first round go head-to-head in a second round. The color that obtains the absolute majority of votes cast wins.

## B.3 Other voting methods and contextualized application

### B.3.1 Approval voting - Explanation

Approval voting is a voting method that allows voters to decide on each color. For each color, each voter can either approve it (vote for it), or not approve it (not vote for it). They can approve as many colors as they wish (between 0 and 8 colors). With approval voting, the winning color is the one with the highest number of approvals.

### B.3.2 Approval voting - Further explanation

With approval voting, a candidate's score corresponds to the number of approvals he or she has received.

Let's take the following example: 2 voters E1 and E2 vote on 3 candidates C1, C2 and C3. E1 votes as follows

	Not approved	Approved
C1		X
C2	X	
C3		X

E2 votes as follows

	Not approved	Approved
C1	X	
C2		X
C3		X

The results of this election are as follows:

Candidates	Number of approvals
C1	1
C2	1
C3	2

C3 wins the election because C3 has more approvals than C1, and than C2.

### B.3.3 Approval voting - First experience

A voter wants to support candidate X. How do you think he should fill his ballot? *I do not approve (I do not vote for); I approve (I vote in favor).*

### B.3.4 Approval voting - Vote in context

A single color will be chosen by the co-owners' assembly for the walls of the co-ownership. Which of these candidate colors do you approve of? *I approve; I do not approve for every color*

### B.3.5 Approval voting - Satisfaction

On a scale from 0 (if you think it's not a good voting system) to 20 (if you think it's a good voting system), how would you rate approval voting as a voting system? *Slider between 0 and 20.*

### B.3.6 Evaluative voting, Explication

Grade voting is a voting method that allows voters to express their opinion on each candidate by assigning a grade on the proposed scale (-1; 0; +1; +2; +3; +4).

The evaluations awarded to each candidate are added together. The winner is the candidate with the highest sum of evaluations, *i.e.* the one with the highest average evaluation.

### B.3.7 Evaluative voting, further explication

Grade voting can also be called evaluative voting, or value voting. The score obtained by a candidate corresponds to the sum of the grades he or she has received.

Let's take an example. 2 voters E1 and E2 decide on 3 candidates C1, C2 and C3.

E1 votes as follows:

	-1	0	+1	+2	+3	+4
C1	X					
C2				X		
C3						X

E2 votes as follows:

	-1	0	+1	+2	+3	+4
C1						X
C2		X				
C3			X			

The results of this election are as follows:

Candidates	Total score (sum of scores)
C1	0+4=4
C2	1-1=0
C3	4+2=6

C3 wins the election because C3's total score is higher than that of C1 and C2.

### **B.3.8 Evaluative voting, comprehension**

By voting, a voter wants to give maximum support to X, his preferred candidate. In your opinion, with grade voting, how should he fill his ballot?  $-1, 0, +1, +2, +3, +4$

### **B.3.9 Evaluative voting, vote in context**

The voting method for choosing the color of the condominium wall is now by grade.

How would you rate each of the colors proposed by the union council?  $-1; 0; 1; 2; 3; 4$  for every color

### **B.3.10 Evaluative voting, satisfaction**

On a scale from 0 (if you think it's not a good way to vote) to 20 (if you think it's a good way to vote), how would you rate note voting? *Slider between 0 and 20.*

### **B.3.11 Majority judgment, explanation**

Majority judgment is a voting system that allows voters to express their opinion on each candidate by assigning one of the following judgments: To be rejected; Insufficient; Acceptable; Fairly good; Good; Very good.

At the end of the vote, each candidate receives the judgment that at least 50% of voters gave him or her, or a better rating. The inventors of this rule call this "majority judgment".

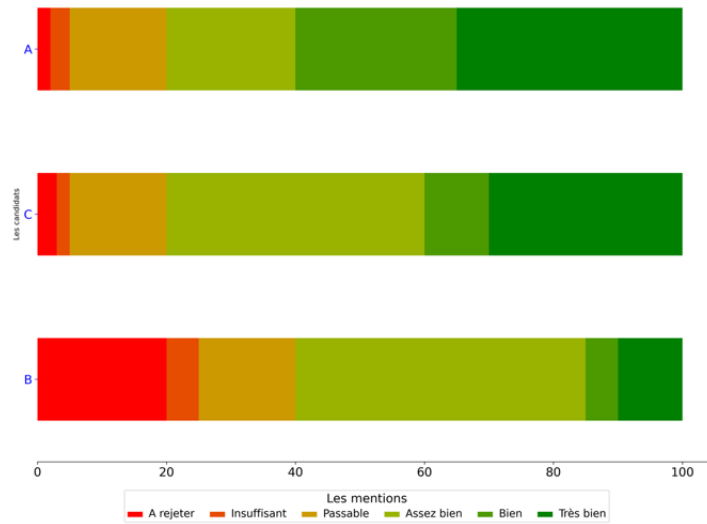
The elected candidate is the one who obtains the best majority score.

When several candidates have the same majority score, a tie-breaking rule must be applied to rank them (see explanations below).

### **B.3.12 Majority judgment, explanation. Merit profiles**

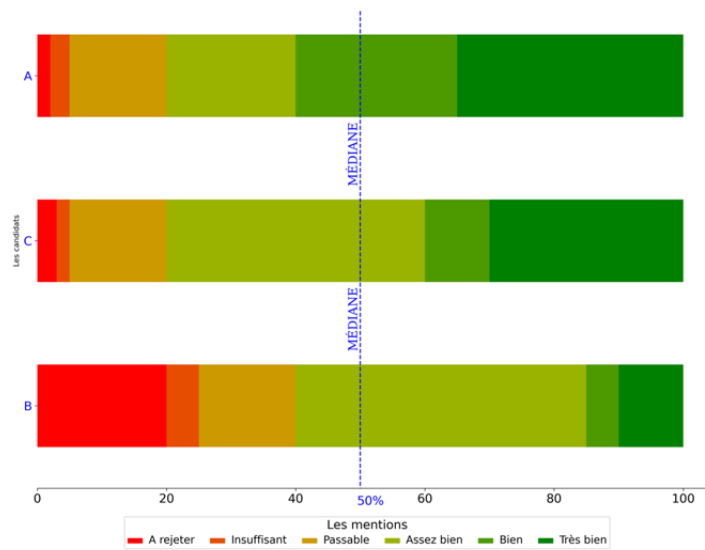
Each candidate obtains a "merit profile" corresponding to the % of votes in each category. These can be represented, for example, in order from worst to best.

For example, we obtain the following merit profiles for candidates A, B and C.



### B.3.13 Majority judgment, explication. Majority grade

The result obtained by a candidate corresponds to the median of the judgments it received. In other words, it is the score with which a majority of voters agree, since at least 50% of voters gave the candidate that judgment or a better one.



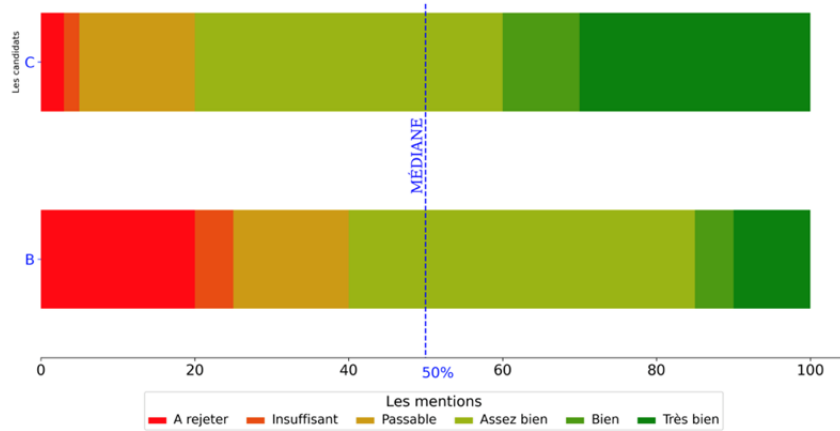
The candidate with the best majority score wins. In this example, candidate A wins the election because he obtains a better majority mark (Good) than candidates C and B (Fairly good).

### B.3.14 Majority judgment, explication. Tie-breaker judgment

[MJ TREATMENT, majority gauge explanation] What happens if two candidates receive the same majority score? For large populations of voters, we can apply the majority gauge method. In this case, the tie is broken by considering that the largest block of voters who disagree with the endorsement

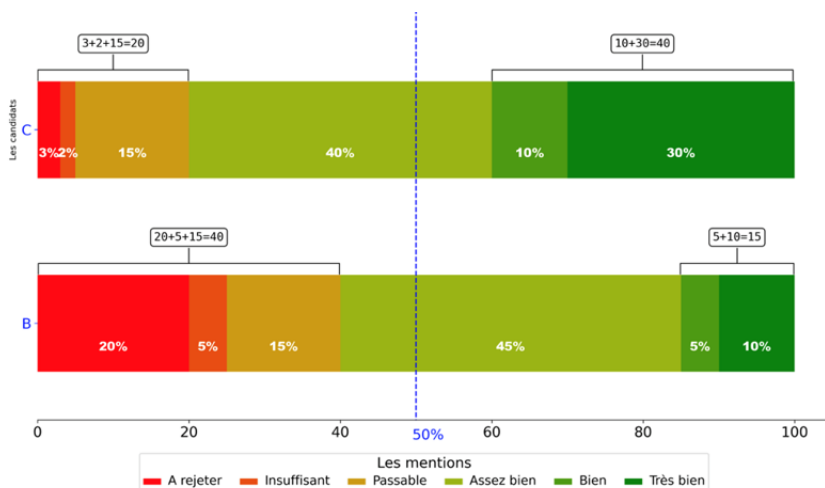
decides: the largest group of voters who think that the candidate was worth more (or less) than the endorsement gets the nod.

For example, B and C obtained the same majority score of “Fairly good”.



We will identify their tie-breaking judgments:

- For B, we ignore the proportion of votes that gave her the highest judgment, in this case “Fairly good”. We compare the number of votes that gave him a better grade (5%+10%, *i.e.*, 15%) with the total number of votes that gave him a worse grade (20%+5%+15%, *i.e.*, 40%). The largest number of people who disagreed with the majority score “Fairly good” were those who would like to give a worse judgment. This means that the tie-breaker for candidate B is “Acceptable”.
- For C, on the other hand, those who give a better judgment (10%+30%=40%) than the majority score are more numerous than those who give a worse judgment (3%+2%+15%=20%). The tie-breaker for C is therefore the judgment “Good”.



As “Good”, C’s tie-breaking score, is better than “Acceptable”, B’s tie-breaking score, C will therefore be ranked higher than B.



### B.3.15 MAJORITY DECISION, EXPLANATION. Tie-breaking decision

[MJ TREATMENT, Iterative deletion explanation] What happens if two candidates receive the same majority score? In this case, we use the iterative deletion method.

Each of these candidates is given a majority score. If one of the candidates then obtains a better majority score than the other, it will be ranked higher than the other. If not, a new majority score is withdrawn from each of the tied candidates until the tie is broken.

In our example, 100 voters judge that the two candidates B and C have the same majority score “Fairly good”. We therefore remove one “Fairly good” judgment for candidate B and one “Fairly good” judgment for candidate C.

- If, once these judgments have been removed, B and C still have the same median judgment, the process is repeated.
- If, once these judgments have been removed, one of these two candidates obtains a better median than the other, we can then rank them.

Click to watch the iterative process through to the tie-breaker

[Interactive video showing the iterative deletion method]

### B.3.16 Majority judgment - First vote experience

How do you think a voter who wants to give maximum support to candidate X is going to vote? *To be rejected; Insufficient; Acceptable; Fairly good; Good; Very good.*

### B.3.17 Majority judgment - Vote in context

The voting method for choosing the color of the condominium wall is now majority judgment.

How do you rate each of the colors proposed by the syndicate council? *To be rejected; Insufficient; Acceptable; Fairly good; Good; Very good.* for every color

### B.3.18 Majority judgment - Satisfaction

On a scale of 0 to 20 (0 if you think it's not a good voting system; 20 if you think it's a good voting system), how would you rate majority voting? *Slider between 0 and 20.*

### B.3.19 The four levels of satisfaction - 1

You have just used four different voting methods to allow co-owners to choose a color for the condominium walls. Which of the four rules did you prefer? *The two-round system (official rule); Approval voting; grade voting; Majority judgment.*

### **B.3.20 The four levels of satisfaction - 2**

You've just experienced grade voting and majority judgment to choose a color for the walls of your condominium in a single round. In both cases, you were able to accurately assess the candidate colors on a 6-level scale.

In your participation in a collective choice, how do you prefer to express yourself? Place your cursor:

- Slider between "I prefer 100% to qualify my vote by giving my opinion on each candidate" to "I 100% prefer to choose a single candidate"
- Slider between "I 100% prefer to use notes to decide on each candidate" to "I 100% prefer to use judgments to decide on each candidate."

### **B.3.21 The four levels of satisfaction - 3**

Which statement(s) do you feel best represent your opinion?

To vote, I appreciate: (Place your cursor between yes and no)

- To know exactly how my vote will be taken into account in the final calculation by expressing myself with (numerical) scores
- To express myself in a common language by evaluating candidates with judgments (words)
- To give my opinion with a ranking
- To give my opinion with smileys
- To vote for a single candidate

### **B.3.22 The four levels of satisfaction - 4**

Which statement do you think best represents your opinion? *Voting means expressing yourself; Voting means saying what you choose; Voting means choosing together.*

### **B.3.23 The four levels of satisfaction - 5**

In these alternative voting methods, you have appreciated that the winner is chosen by: *Sum of grades or approvals; The best judgment grade, with a tie-breaking rule; I liked neither; I liked both*

You've just experienced three different new voting methods for choosing the color of the walls in your condominium. Which of the three rules did you like least? *Approval voting; Grade voting; Majority Judgment.*

Why? *Open ended question*

## B.4 Introduction to decontextualized experimentat

You are now going to use evaluative voting/majority judgment in a series of virtual elections. Your remuneration will depend on your gains in these virtual elections, which we will now introduce to you.

Would you like to find out more about voting by mark/majority judgment before moving on? *I would like to return to the “find out more” explanations; No thanks, I will go on to the next step.*

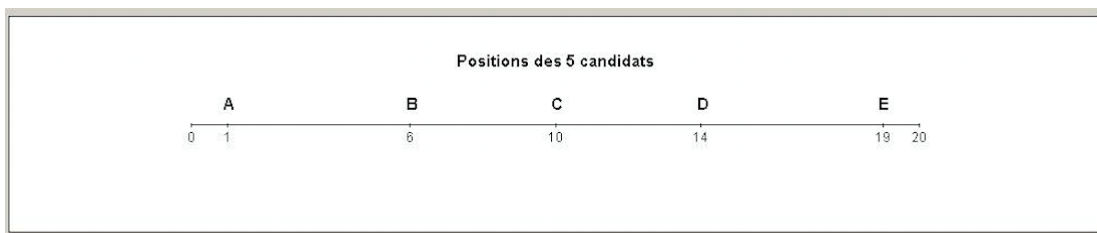
### B.4.1 Presentation of the decontextualized experiment - procedure

During each election, the 21 people in this session will elect a winner from five candidates, whom we will call candidates A, B, C, D, E. These candidates are located along an axis numbered from 0 to 20.

### B.4.2 Presentation of the decontextualized experiment - the candidates

Candidate A is in position 1, candidate B in position 6, candidate C in position 10, candidate D in position 14, and candidate E in position 19. These positions are fixed and will remain the same throughout the session.

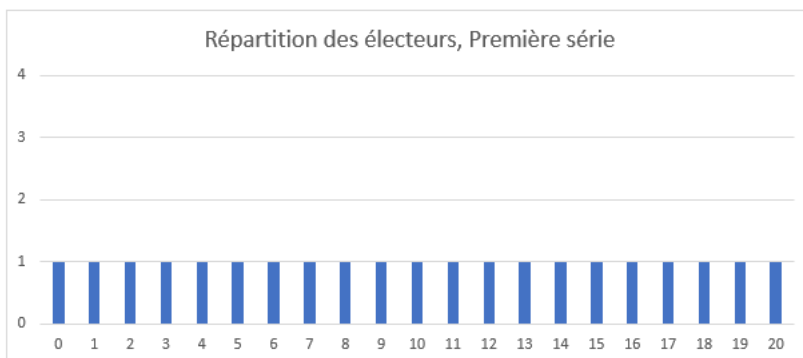
Figure B.4.1: Positions of the 5 candidates



### B.4.3 The voters

In the first round of elections, the 21 voters are distributed from 0 to 20, from 1 to 1. Your position will be drawn by lot.

Figure B.4.2: Positions of the 5 candidates



### B.4.4 Procedure

You will take part in 3 series of 5 repeated elections. The candidates' positions remain fixed throughout the experiment. A voter's position changes from one round to the next, but remains the same during each round. The rank of each winning candidate is publicly announced after each intermediate election.

### B.4.5 Remuneration method

For this series of fictitious elections, your gain will be greater the closer the position of the candidate elected in the last election is to your position.

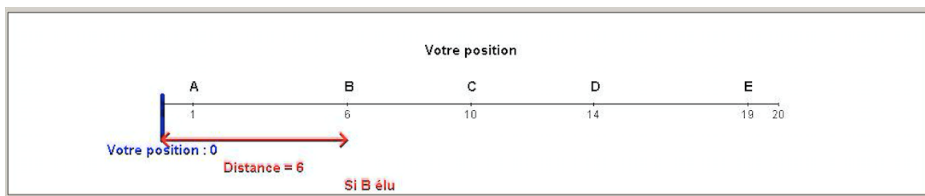
Your gain (in points) = 20 - Distance between the elected candidate's position and yours.

After each election, you will be informed of your corresponding payoff.

### B.4.6 Remuneration method - Example 1

For example, if you're in position 0, here's your payoff if candidate B is elected:

Figure B.4.3: Example of remuneration

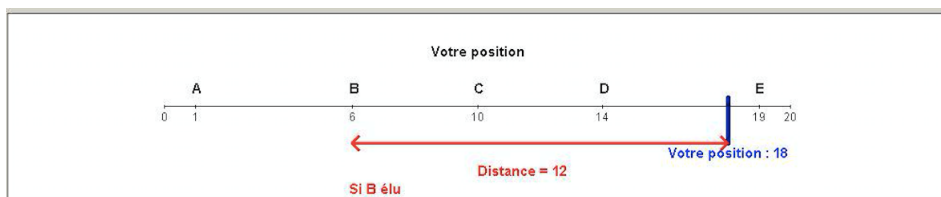


Gain = 20 - 6 = 14 points. *I understood; I would like more explanations*

### B.4.7 [If they answer “I would like more explanations” in the previous question] Remuneration method - Example 2

For example, if you're in position 18, here's your payoff if candidate B is elected:

Figure B.4.4: Example of remuneration



Gain = 20 - 12 = 8 points. *I understood; I would like more explanations*

In each series of five elections, your actual payoff will ultimately depend only on your win in the fifth election. However, after each election you will be informed of the corresponding potential earnings.

You can earn a minimum of 1 point (if the average distance between you and the elected candidate is 19) and a maximum of 20 points (if you are in the exact position of the elected candidate). Your payoff for winning this election can therefore vary from 0.5 to 10 euros.

### B.4.8 Remuneration

At the end of this experiment, you will have taken part in 15 elections (3 series X 5 repetitions).

To calculate your share of earnings for these fictitious series of elections, the computer will take the points earned in the last election of each of the three series in which you will take part.

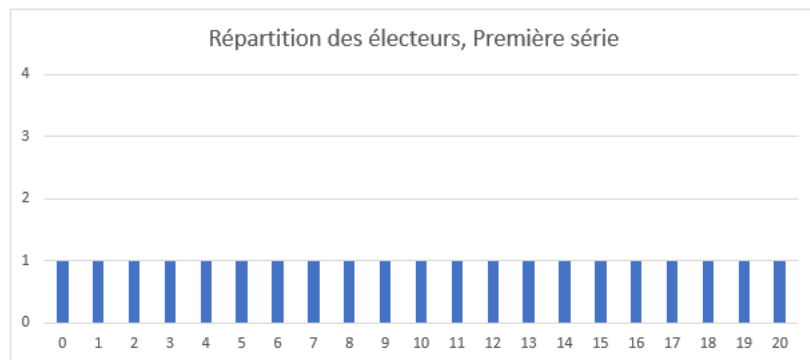
Your position in the first series of elections is drawn at random. The positions you will be assigned in the second and third series are designed to ensure that all participants have equal earnings expectations. In other words, the protocol is designed so that your payoff does not depend on the random draw.

### B.4.9 Series 1 - Your position

[EV] You are going to vote with grade voting. [MJ] You are going to vote with majority judgment.

Here is the first distribution of voters along the axis from 0 to 20

Figure B.4.5: Positions of the 5 candidates



**Your position is:** [0-20]

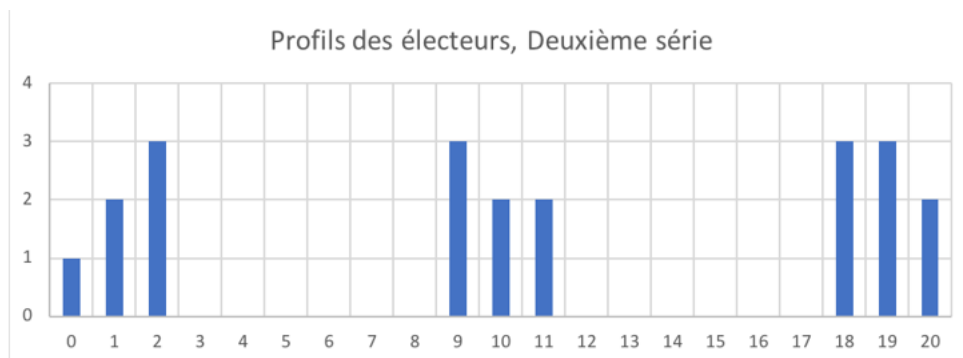
The position of all voters is maintained for the next 5 elections.

Your win = 20 - distance between the elected candidate and [0-20]

### B.4.10 Series 2 - Your position

You're going to vote at MJ/EV. Here is the first distribution of voters along the axis from 0 to 20

Figure B.4.6: Positions of the 5 candidates



**Your position is:** [0-20]

The position of all voters is maintained for the next 5 elections.

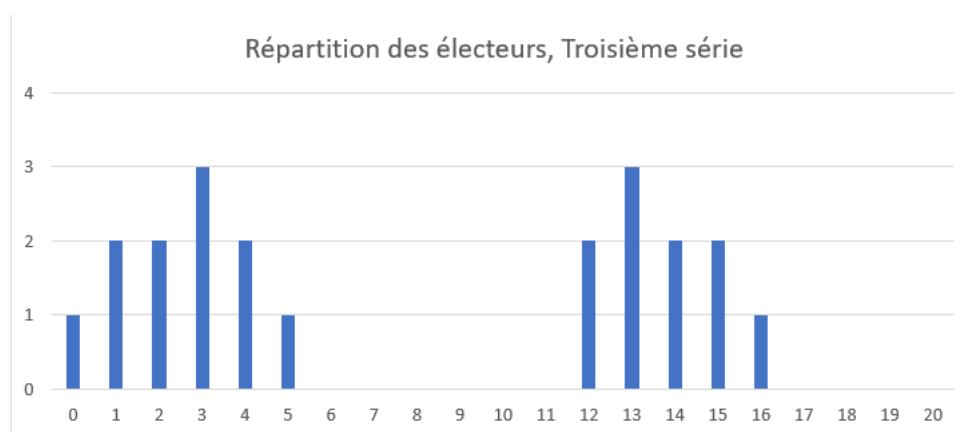
Your win = 20 - distance between the elected candidate and [0-20]

### B.4.11 Series 3 - Your position

[EV] You are going to vote with grade voting. [MJ] You are going to vote with majority judgment.

Here is the first distribution of voters along the axis from 0 to 20

Figure B.4.7: Positions of the 5 candidates



**Your position is:** [0-20]

The position of all voters is maintained for the next 5 elections.

Your win = 20 - distance between the elected candidate and [0-20]

## B.5 The vote with majority judgment

The voting system is: majority judgment

Your position is XX.

How would you rate each candidate?

Candidates	Judgments					
	To reject	Insufficient	Acceptable	Fairly good	Good	Very good
A						
B						
C						
D						
E						

Confirm your vote.

### B.5.1 Majority judgment - Results

Here is the merit profile obtained by the candidates for the election in which you have just participated:

Candidates	Profil de mérite						
A	Dark Green	Light Green	Light Green	Light Green	Grey	Grey	Yellow
B	Dark Green	Dark Green	Dark Green	Light Green	Grey	Yellow	Yellow
C	Light Green	Light Green	Light Green	Light Green	Grey	Grey	Yellow
D	Light Green	Light Green	Light Green	Light Green	Light Green	Grey	Grey
E	Dark Green	Light Green	Light Green	Light Green	Light Green	Grey	Grey

Elected candidate: XX

Your calculated payoff for this election: XX

### B.6 The vote with evaluative voting

The voting system is: grade voting

Your position is XX.

How would you rate each candidate?

Candidates	Judgments					
	-1	0	+1	+2	+3	+4
A						
B						
C						
D						
E						

Confirm your vote.

#### B.6.1 Evaluative voting - Results

Here are the scores obtained by the candidates for the election in which you have just participated:

Candidates	Sum of grades
A	56
B	43
C	12
D	78
E	32

Elected candidate: D

Your calculated payoff for this election: XX

## B.7 Comprehension - Presentation

We are now going to ask you ten questions about the voting systems presented in this experiment. Each question can earn you up to 2 points.

Your remuneration will depend on your answers:

- 2 points for each correct answer
- 0 points for wrong answers
- 0.5 points if you ticked “I do not know”.

## B.8 Comprehension - Evaluative voting

### B.8.1 Comprehension - Vote 1

Consider an election involving 2 candidates: candidate A and candidate B. Each of the 11 voters has given a grade to each of the 2 candidates. This results in the following distribution of the grades awarded to each of the two candidates (Each X represents a vote)

Grades	-1	0	1	2	3	4
Candidate A	XX	XX	X	XX	XXX	X
Candidate B	X	X	XXXX	X	X	XXX

With grade voting, which candidate do you think was elected by these 11 voters? Please select an answer below. *Candidate A; Candidate B; I do not know*

### B.8.2 Comprehension - Vote 2

Consider an election involving 2 candidates: candidate A and candidate B. Each of the 11 voters has given a grade to each of the 2 candidates. This results in the following distribution of the grades awarded to each of the two candidates (Each X represents a vote)

Grades	-1	0	1	2	3	4
Candidate A	XXXX	X		X	XX	X
Candidate B	XX	XX	XX	X	X	X

With grade voting, which candidate do you think was elected by these 11 voters? Please select an answer below. *Candidate A; Candidate B; I do not know*



**B.8.3 Comprehension - Vote 3**

Consider an election involving 2 candidates: candidate A and candidate B. Each of the 11 voters has given a grade to each of the 2 candidates. This results in the following distribution of the grades awarded to each of the two candidates (Each X represents a vote)

Grades	-1	0	1	2	3	4
Candidate A	XXXX	X		XXX	XX	X
Candidate B	XX	XX	XX	X	XXX	X

With grade voting, which candidate do you think was elected by these 11 voters? Please select an answer below. *Candidate A; Candidate B; I do not know*

**B.8.4 Comprehension - Sentences**

For this question, if at least one answer is incorrect, your payoff is 0. If all answers are “I do not know”, your payoff will be +0.5. If all answers are correct, your payoff is +2.

Knowing that the candidate with the highest score is elected, the grade voting assigns each candidate a score that

1. Corresponds to the average score given to this candidate by all voters. *True; False; I do not know.*
2. Corresponds to the average rating given to this candidate by half plus one of the voters who are most favorable to him/her. *True; False; I do not know.*
3. Corresponds to the score given by the voter who divides in two the set of voters ranked from the most favorable to the most unfavorable. *True; False; I do not know.*

**B.9 Comprehension - Majority judgment**

**B.9.1 Comprehension - Vote 1**

Consider an election involving 2 candidates: candidate A and candidate B. Each of the 9 electors has voted on the 2 candidates, giving them one vote each. This results in the following distribution of judgments for each of the two candidates:

Grades	To reject	Insufficient	Acceptable	Fairly good	Good	Very good
Candidate A	XXX		X		XX	XXX
Candidate B	XX	X	X	XX		XXX

With the majority judgment, which candidate do you think was elected by these 11 voters? *Candidate A; Candidate B; I do not know*

### B.9.2 Comprehension - Vote 2

Consider an election involving 2 candidates: candidate A and candidate B. Each of the 9 electors has voted on the 2 candidates, giving them one vote each. This results in the following distribution of judgments for each of the two candidates:

Grades	To reject	Insufficient	Acceptable	Fairly good	Good	Very good
Candidate A	X	XXX	X		X	XXX
Candidate B	XX		XX	XX		XXX

With the majority judgment, which candidate do you think was elected by these 11 voters? *Candidate A; Candidate B; I do not know*

### B.9.3 Comprehension - Vote 3

Consider an election involving 2 candidates: candidate A and candidate B. Each of the 9 electors has voted on the 2 candidates, giving them one vote each. This results in the following distribution of judgments for each of the two candidates:

Grades	To reject	Insufficient	Acceptable	Fairly good	Good	Very good
Candidate A	X	XXX	X		XXX	XXX
Candidate B	XX		XX	XX		XXXXX

With the majority judgment, which candidate do you think was elected by these 11 voters? *Candidate A; Candidate B; I do not know*

### B.9.4 Comprehension - Sentences

For this question, if at least one answer is incorrect, your payoff is 0. If all answers are “I do not know”, your payoff will be +0.5. If all answers are correct, your payoff is +2.

Bearing in mind that the candidate with the best score is elected, the majority judgment awards each candidate a score that:

1. Corresponds to the average rating given to this candidate by all voters. *True; False; I do not know.*
2. Corresponds to the average rating given to this candidate by half plus one of the voters most in favor of him/her. *True; False; I do not know.*
3. Corresponds to the mention given by the voter who divides in two the set of voters ranked from the most favorable to the most unfavorable. *True; False; I do not know.*

## B.10 Comprehension - Approval voting

You have also tested approval voting in the choice of colors. Do you remember how this voting system works? Take a look at this example.

11 voters approve candidates A and B.

	Not approved	Approved
Candidate A	XXXXXXXXX	XXX
Candidate B	XXXXX	XXXXXX

With approval voting, which candidate do you think was elected by these 11 voters? *Candidate A; Candidate B; Candidate C; Candidate D; I do not know*

## B.11 Satisfaction and assessment of voting methods

### B.11.1 Satisfaction - Results

What is your position on the following statement: Citizens must be able to easily deduce the election result after the count, once the contents of the ballots are known. *Useless; Not necessary; Desirable; Fundamental; Other: Would you like to specify?*

### B.11.2 Satisfaction

You have discovered four different voting systems. Which of the four rules did you prefer? *Two-round uninominal voting (the official rule); Approval voting; Evaluative voting; Majority judgment.*

Why did you prefer [what they chose before]?

### B.11.3 Knowledge of the rules (Only in the last 5 sessions)

Did you know about two-round uninominal system before starting this experiment? *Yes; No*

Did you know about "approval voting" before starting this experiment? *Yes; No*

Did you know about "grade voting" before starting this experiment? *Yes; No*

Did you know about "majority judgment" before starting this experiment? *Yes; No*

## B.12 Identity

How old are you? *18-24 years old; 25-34 years old; 35-49 years old; 50-64 years old; 65 years old or more.*

Are you? *Man; Woman; I prefer not to answer.*

What is your occupation? *Farmer, craftsman, self-employed, company director; Executive; Intermediate profession; Employed; worker; retired; inactive; student; I do not wish to answer; I want to precise*

What is your highest level of qualification? *CAP/BEP; High school; 1-year college or equivalent; 2-year college or equivalent; 3-year college or equivalent; 1-year Master degree or equivalent; 2-year master or equivalent; Ph.D or equivalent; No degree? Other*

In which field did you mainly study? *Arts, languages, literature; Biology, chemistry, geology, mathematics, physics; Law; Geography; Management / Trade; History; Informatics; Engineering; Medicine, Physiotherapy, health care; economics; sociology; sport; Technique, Technology; Handicrafts, Crafts; other*

What is the size of the agglomeration where you vote? *Rural municipality; 2,000 to 19,999 inhabitants; 20,000 to 99,999 inhabitants; More than 100,000 inhabitants; Parisian agglomeration.*

### **B.13 Remuneration**

The experiment is now over. All data is recorded strictly anonymously. Only your remuneration is associated with your name.

Your remuneration will be XX euros.

For your information, your remuneration consists of:

- A fixed gain for your participation in the experiment: 5 euros fixed
- A variable gain based on your performance in the repeated election series: XX euros
- A variable gain depending on your performance in answering the questions on voting methods: XX euros.

### **B.14 Last screen**

Please wait until all participants have completed the experiment.

We will let you know when you can proceed to the office to collect your payment. Take only your label and completed payment receipt with you. Please do not communicate the content of this experience with anyone else.

Thank you for your participation! The experiment is now over. All data is recorded strictly anonymously. Only your payment is associated with your name. Your remuneration will be XXX euros. You can take it in turns to go to the office to receive your payment.