Some annoying grading scale effects under different versions of evaluative voting

Antoinette Baujard
Université de Lyon at Saint-Etienne, UJM, GATE Lyon Saint-Etienne CNRS
with Isabelle Lebon, Herrade Igersheim
Université Caen Normandie, CREM CNRS / CNRS Paris School of Economics / CNRS University of Strasbourg
Thanks to the team VOTER AUTREMENT
including Thierry de Cordoue Hecquart, Jean-François Laslier, Sylvain Bouveret, Renaud Blanch, Vincent Merlin, Annick Laruelle, the elected officials, city staffs and the volunteer citizens of Hérouville Saint-Clair, Strasbourg, Grenoble, Crolles, Allevard...

3rd ILLC Workshop on Collective Decision Making
Institute for Logic, Language and Computation
University of Amsterdam, 6-7 June 2019

In a number of voting rules allowing voters to have a say on every candidates, there may be different balloting devices :

- Yes or no, approval or non-approval
- Numbers on a given grade scale : 0,1 or 0,1,2 or -1,0,1 or -2,-1,0,1,2 or -0.5,0,+1 or 0 to 20 or 0 to $+\infty$...
- Appreciation such as "excellent, very good, good, passable, mediocre, inadequate", or "excellent, good, fair, poor"...

Which balloting devices are more desirable than others ?

Most voters are more satisfied with more expressive rules, everything being equal.
Is the choice of one balloting device just a matter of more or less fine-grained information ? Alternatively, has it an impact on voters behaviors and election outcome ?
Some think that more options is likely to favor inclusive candidates and disfavor populism and corruption (e.g. Janacek, D2.1). Is this bias confirmed and the only one ? Can we formulate desirable properties concerning balloting ?

Evaluative Voting (also called grade voting, range voting, utilitarian voting)

A balloting device Voters grade each candidate independently from a given grade scale.

An aggregation rule The candidate who gets the higher sum of grades is the winner.

	0	1	2
Nicolas Dupont-Aignan			
Marine Le Pen			
Emmanuel Macron			
Benoît Hamon			
Nathalie Arthaud			
Philippe Poutou			
Jacques Cheminade			
Jean Lassalle			
Jean-Luc Mélenchon			
François Asselineau			
François Fillon			

Different possible scales for evaluative voting, e.g., EV(-1,0,1,2), EV(0,1,2,3,4,5)...

	-1	0	1	2
Nicolas Dupont-Aignan				
Marine Le Pen				
Emmanuel Macron				
Benoît Hamon				
Nathalie Arthaud				
Philippe Poutou				
Jacques Cheminade				
Jean Lassalle				

	-1	0	+1
Nicolas Dupont-Aignan			
Marine Le Pen			
Emmanuel Macron			
Benoît Hamon			
Nathalie Arthaud			
Philippe Poutou			
Jacques Cheminade			
Jean Lassalle			

	0	1	2
Nicolas Dupont-Aignan			
Marine Le Pen			
Emmanuel Macron			
Benoît Hamon			
Nathalie Arthaud			
Philippe Poutou			
Jacques Cheminade			
Jean Lassalle			

	0	1	2	3
Nicolas Dupont-Aignan				
Marine Le Pen				
Emmanuel Macron				
Benoît Hamon				
Nathalie Arthaud				
Philippe Poutou				
Jacques Cheminade				
Jean Lassalle				

	0	1	2	3	4	5
Nicolas DUPONT-AIGNAN						
Marine LE PEN						
Emmanuel MACRON						
Benoît HAMON						
Nathalie ARTHAUD						
Philippe POUTOU						
Jacques CHEMINADE						
Jean LASSALLE						

This paper scrutinizes wether and how variations of length scales matter and whether and how the introduction of a negative grade matter.

Evaluative voting One round voting rule where voters can assess every candidate, independently, on a given grade scale (EV).

Use in multicriteria decisions: Sport competition, School grading system

Use in political elections: AV in open list system : France, Swiss ; Cumulative voting in Germany, Luxembourg ; Negative voting : Latvia

The length and the negative grade in particular, might or might not matter.

Voters satisfaction: Experimental data confirm that voters like better rules and scales with more opportunity for expression than less, e.g. the possibility to give negative grades and (not too) larger scales

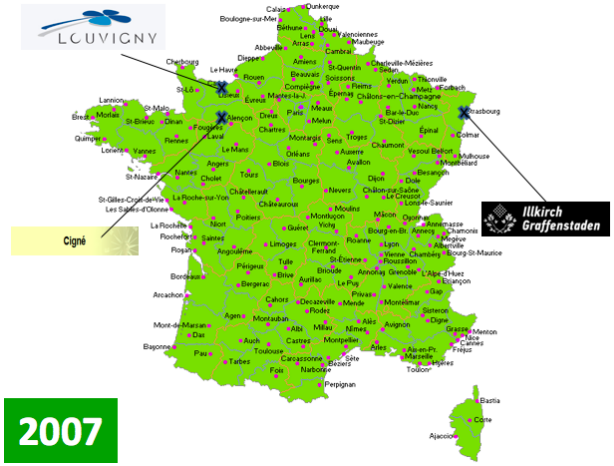
Strategic issue: Larger scale favor and give more power to more strategic voters, which is debatable

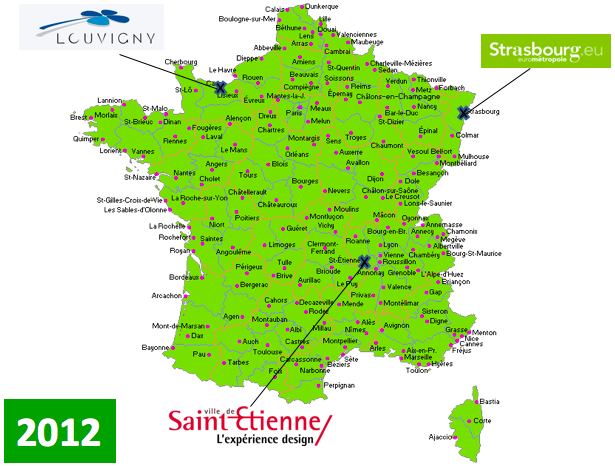
Behavioral calibration: The presence of negative grades significantly affect behaviors hence outcome (Baujard & al, 2018). But only a conjecture could be formulated about length : it might be stable under different lengths.

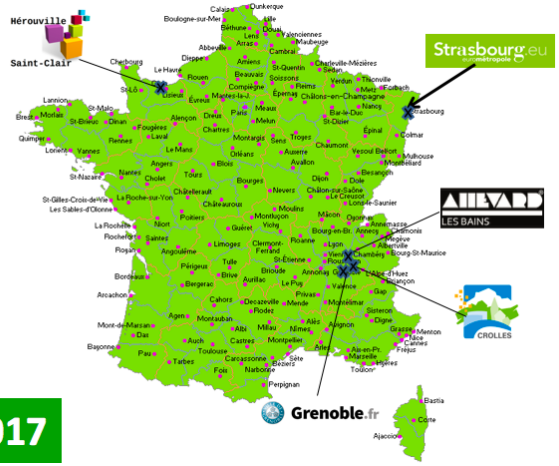
The aims of the paper

1. Expose a new protocole to provide reliable data likely to compare grading behaviors for different scales (with no correction, selection bias...)
2. Confirm and specify results concerning label effects and the role of the negative grade. The introduction of a negative grade clearly distorts the scores in disfavor of some candidates, but this paper enables to identify which one in particular.
3. Test the assumption that there is no significant length effect. Clearly, this conjecture does not stand up to scrutiny. With a longer scale, more voters refrain from using extreme grades. Again, we highlight that certain minor unkown candidates are likely to be favored by longer scales.

Design


LCUVIGNY
Cigné
Illkirch Graffenstaden
2007


LCUVIGNY
Strasbourg.eu
Saint-Étienne
L'expérience design
2012


Hérouville
Saint-Clair
Strasbourg.eu
ALLEVARD
LES BAINS
CROLLES
Grenoble.fr
2017

Before the experiment, information step

LE DOSSIER DU JOUR | EN ISÈRE

CITOYENNETÉ | Des chercheurs, dont certains sont isérois, proposent,

Et si on pouvait voter

LE DOSSIER DU JOUR | EN ISÈRE

le 23 avril, de tester d'autres moyens de voter à Grenoble, Crolles et Allevard

autrement ?

3 149 Grenoblois invités à voter autrement au premier tour

À Crolles, ce sera un vote par approbation à deux tours

ÉLECTRICES, ÉLECTEURS DE CROLLES
PARTICIPEZ A UNE EXPERIMENTATION DE VOTE

LE 23 AVRIL 2017, LORS DU PREMIER TOUR DES ELECTIONS PRESIDENTIELLES
APRES LE VOTE OFFICIEL, DANS VOTRE BUREAU DE VOTE

En tant qu'électrice(eur)(s) de Crolles, vous êtes invité(e)(s) à expérimenter de nouveaux modes de scrutin à l'occasion du premier tour de la présidentielle de 2017.
Suite à une initiative des samedi-citoyens, vous serez invité à venir voter expérimentalement, après avoir voté officiellement. Un bureau de vote expérimental sera installé près de votre bureau de vote officiel avec l'accord de la municipalité et de la préfecture. Votre vote-test est totalement anonyme et les résultats expérimentaux n'ont aucun impact sur le résultat officiel. L'expérimentation, de nature scientifique, est conçue par des chercheurs du CNRS, des Universités de Grenoble, Lyon Saint-Étienne, Caen Normandie, Paris, Strasbourg et Bilbao. Cette opération sera organisée simultanément dans d'autres communes françaises : Allevard-les-Bains et Grenoble (Isère), Strasbourg (Bas-Rhin), Hérouville-Saint-Clair (Calvados).

Notre but est de mieux comprendre le fonctionnement et l'impact des modes de scrutin. Toutes les règles de vote ont pour objectif de désigner démocratiquement le candidat élu. Cependant tous les pays n'utilisent pas la même règle, et chacune a des caractéristiques différentes qui peuvent modifier le résultat.
En France, le scrutin en vigueur pour les élections présidentielles est le vote uninominal à deux tours. Notre objectif est d'expérimenter et d'étudier les effets de modes de scrutin alternatifs qui offrent plus de possibilités aux électeurs de s'exprimer. Chaque électeur crollois pourra tester, aléatoirement, soit le vote par approbation à 1 tour, soit le vote par approbation à 2 tours (voir bulletins au dos).

Testez le vote par approbation. Il se déroule en un seul tour de scrutin. L'électeur indique les candidats qu'il approuve. Il peut ainsi voter pour un seul candidat ou pour plusieurs candidats. Le candidat qui réunit le plus grand nombre d'approbations est élu. Les électeurs qui testeront le vote par approbation se verront aussi proposer de donner leur opinion sur tous les candidats, en les évaluant par une note de 0 à 20.

Testez le vote par approbation à deux tours. Il se déroule en deux tours de scrutin. Les électeurs approuvent ou non chaque candidat au premier tour. Sont sélectionnés pour le duel du second tour les deux candidats ayant réuni le plus grand nombre d'approbations. Nous ne pouvons pas prévoir quel serait le vrai second tour dès le 23 avril ; pour déduire qui serait élu au second tour, nous vous demanderons de classer vos candidats préférés par ordre.

Un compte-rendu des résultats du vote expérimental sera disponible après les élections législatives sur notre site internet, et les analyses seront publiées dans des revues scientifiques internationales.

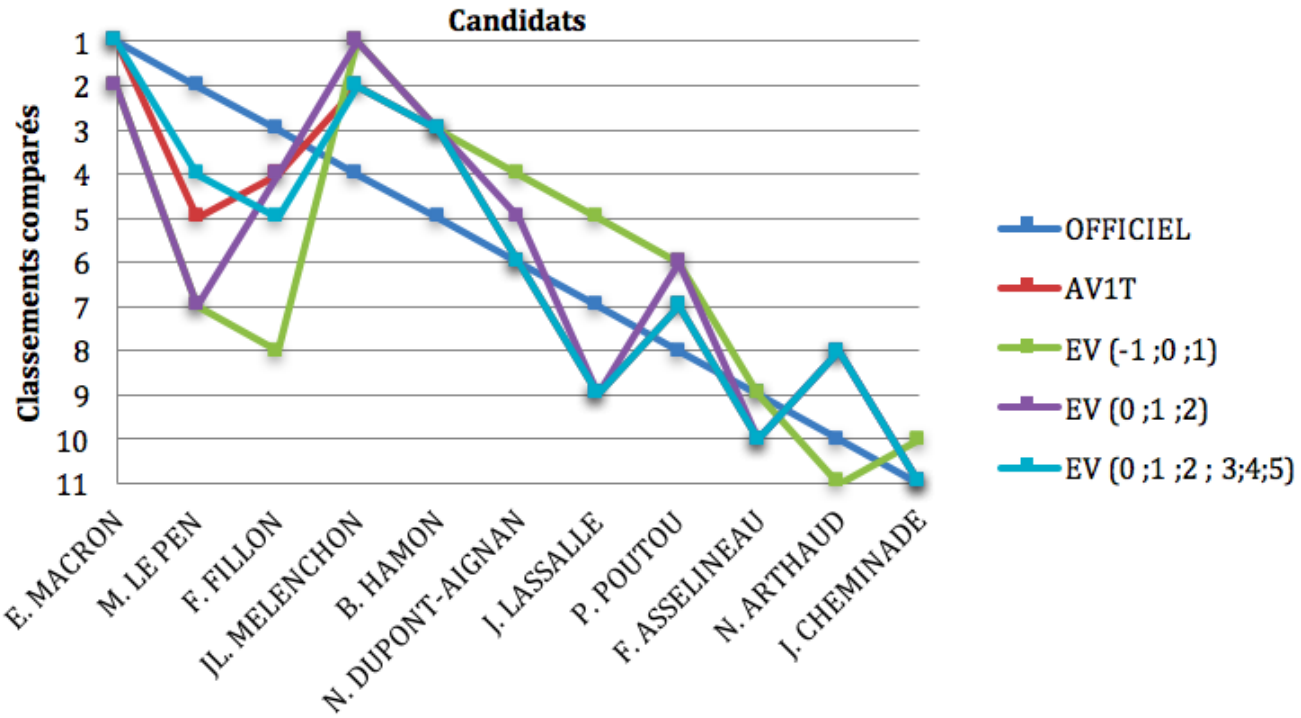
QUESTIONS ? COMMENTAIRES ? ENVIE D'EN SAVOIR PLUS ?
REUNION D'INFORMATION LE 6 AVRIL 2017 A 19H, SALLE BORIS VIAN A CROLLES

https://www.facebook.com/votaction2017/
Contact : Antoinette Baujard, directrice-adjointe du laboratoire CNRS GATE Lyon Saint-Etienne, Antoinette.baujard@univ-st-etienne.fr
Et notre site web : https://www.gate.cnrs.fr/spip.php?article580

D-day– April 23, 2017 : Experimental voting progress

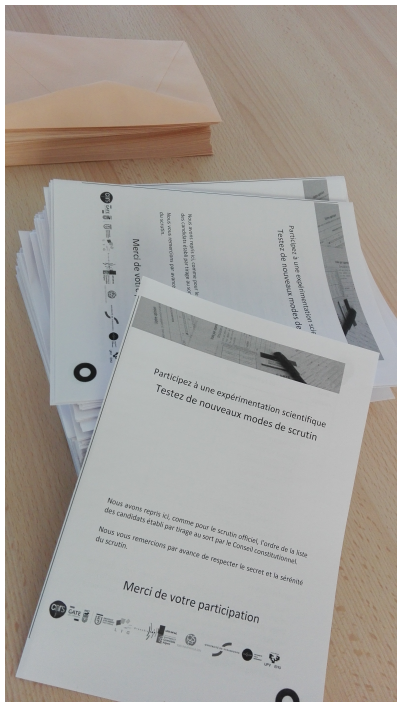
Crolles, France, 23 avril 2017
Premier tour de l'élection présidentielle française

Compared ranking of candidates for different voting rules tested in situ, on the basis of corrected data



Source : Compte-rendu de l'expérimentation VOTER AUTREMENT lors du premier tour de l'élection présidentielle française le 23 avril 2017 à Allevard-les-Bains, Crolles, Grenoble, Hérouville-Saint-Clair, Strasbourg et sur internet, 27 juin 2017. En ligne sur gate.cnrs.fr/vote.

Randomized allocation of ballot papers per voting station



Hérouville-Saint-Clair

- **AV** and **EV[0,1,2,3]**
- **AV** and **EV[0,1,2,3,4,5]**

Strasbourg

- **AV** and **EV[0,1,2]**
- **AV** and **EV[-1,0,1]**
- **AV** and **EV[0,1,2,3]**
- **AV** and **EV[-1,0,1,2]**

Grenoble

- **AV** and **EV[0 ;1]**

Crolles

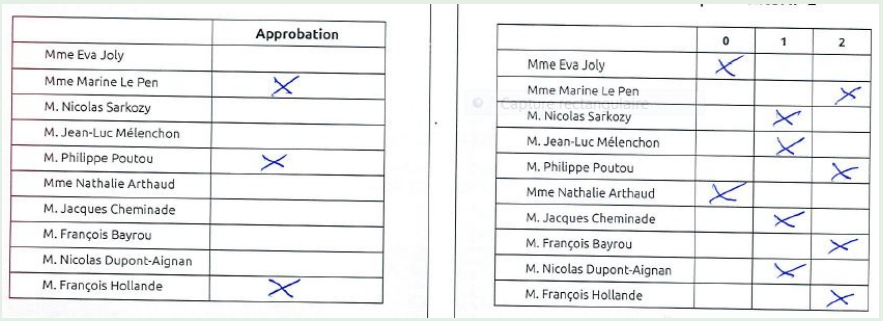
- **AV** with a survey on pol. opinion on a {0-20} scale
- **AV2T**

Allevard-les-Bains

- **D&A**
- **Double D&A**
- **Semi D&A**

Voting with grades consists in dividing the set of candidates into 2, 3, 4 or 6 (x) equivalence classes, and rank them according to their preferences.

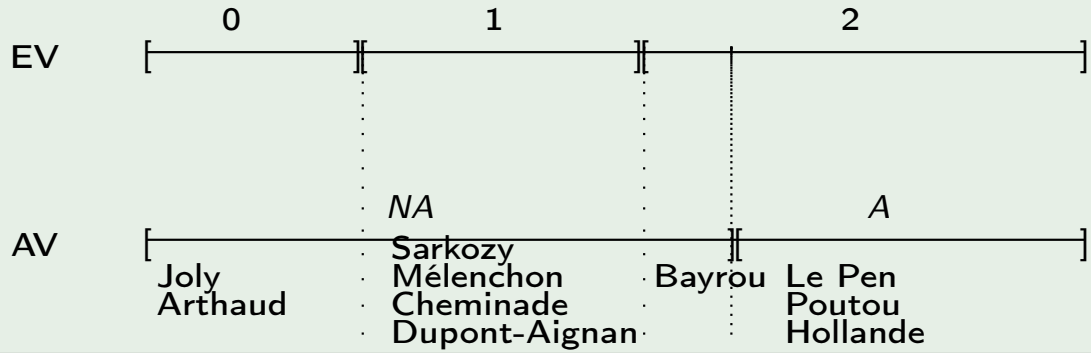
An illustration. Ballot 10, Saint-Etienne 2012



	Approbation
Mme Eva Joly	
Mme Marine Le Pen	X
M. Nicolas Sarkozy	
M. Jean-Luc Mélenchon	
M. Philippe Poutou	X
Mme Nathalie Arthaud	
M. Jacques Cheminade	
M. François Bayrou	
M. Nicolas Dupont-Aignan	
M. François Hollande	X

	0	1	2
Mme Eva Joly	X		
Mme Marine Le Pen			X
M. Nicolas Sarkozy		X	
M. Jean-Luc Mélenchon		X	
M. Philippe Poutou			X
Mme Nathalie Arthaud	X		
M. Jacques Cheminade		X	
M. François Bayrou			X
M. Nicolas Dupont-Aignan		X	
M. François Hollande			X

$\{$ Joly ; Arthaud $\} \prec \{$Sarkozy ; Mélenchon ; Cheminade ; Dupont-Aignan $\} \prec \{$ Bayrou $\} \prec \{$ Le Pen ; Poutou ; Hollande $\}$



Given that various evaluative voting choices reflect consistently the ordinal preferences for candidates when scales differ, whichever in labels or lengths, research questions concerns the variations in the expression of preferences and their impact on the result when scale lengths vary.

A specific design

Each participant vote for the 11 official candidates. Default grade is always the lowest one (0 for EV, or -1 for EVneg). Each voter pick a ballot randomly and vote under :

- EV2=AV, i.e. a 2-step scale : EV(0,1)

And one of either of the following rules

- EV3 or EV3neg, i.e. a 3-step scale : EV(0,1,2) or EV(-1,0,1)
- EV4 or EV3neg, i.e. a 4-step scale : EV(-1,0,1,2) or EV(0,1,2,3)
- EV6, i.e. a 6 step-scale : EV(0,1,2,3,4,5)

- Randomization in ballot distribution allows the results to be used directly to compare the impact of the grading scales. Because the participants can be considered homogeneous in each place.
- The low or high participation of voters according to their political orientation is not a problem. But the scores obtained can have no political interpretation.

Theoretical literature on EV.

- Axiomatic charaterizations of evaluative voting : Smaoui and Lepelley (2013), Pivato (2013), Macé (2018) ;
- Axiomatics on similar rules : Aleskerov (2007), Gaertner and Xu (2012), Alcantud and Laruelle (2014), Gonzalez, Laruelle and Solal (2019) ;
- How preferences express in individual EV ballots : Ceron and Gonzalez (2019)

How strategic or sincere voting occurs in evaluative voting allows to introduce the issue of translation of preferences into their expression in balloting.

- Theoretical prediction : A strategic voter should use only extreme grades : Nunez and Laslier (2014) :
- Empirical observation : voters use intermediate grades, even more in real contexts (than in controlled contexts), eg. Igersheim et al. (2016).
- In sincere voting, there exists some underlying true evaluation of the candidates, as may be captured in the ballots, hence a problem of *representation* in the sense of the theory of measurement (Narens 1985), coined the *calibration* problem (Baujard et al. 2018).

The calibration problem

How preferences are translated into grades for different contexts, e.g. different evaluative scales.

Properties of Voting rule Different candidates may be affected differently by the changing scales– hence different results for different voting rules/scales.

A behavioral issue We know little about why and how variations of scales matter – hence an experimental inquiry is necessary

Following the principle of relative utilitarianism (Dhillon and Mertens 1999), a voter uses the extreme grades for her best and worse candidates, and will accommodate the rating of the others upon her preferences and the grading scale.
Or alternatively, all candidates judged poorly may be given bad grades, which suppose to imagine two hypothetical extreme candidates –one absolutely good and one absolutely bad– to calibrate linearly.
Linear calibration implies that, when comparing EV3 or EV6 on two samples of the same population, the observed fractions of first halves of grades should be equivalent in both.

Invariance with negative grade

Assumption AN. Numerical scales of same length but different labels are linearly equivalent

- AN1 For each candidate, the score is translated by one unit when comparing EV3 with EV3neg and EV4 with EV4neg.
- AN2 Up to a 1-point translation, the different scales with and without negative grade generate the same proportion of lowest grade for each candidate when he/she not approved.
- AN3 Up to a 1-point translation, the different scales with and without negative grade generate the same proportion of highest grade for each candidate when he/she approved.

Invariance with length

Assumption. Numerical scales of different length are linearly equivalent :

- AL1 The distribution of grades associated with each candidate remains stable for the scales AV, EV4 and EV6 when reduced in two classes.
- AL2 The normalized scores of each candidates are unchanged under various scales.
- AL3 The length of the scale does not change the use of the extreme grades for any candidates. For each candidate, the score is translated by one unit when comparing EV3 with EV3neg and EV4 with EV4neg.

Assumption. Length does not change the propensity to use the entire scope of the grading scale.

- AG1 For all scale lengths, all voters use the entire scale of grades
- AG2 The lengthening of the scale does not modify the proportion of voters who use the full extent of the grading scale.

Major vs minor

- Viable (Cox, 1997) / Serious (Myerson, 2002) / Major : candidates who have a reasonable chance to win the election
- Non viable / Minor candidates

Yet, the notion of "viable" candidate must be adapted to the voting rule : A multinominal voting rule (for which the outcome is not sensible to close/clone candidates) may allow a large number of viable candidates.

- 5 major candidates : the 4 viable candidates of the official voting : E. Macron (EM), M. Le Pen (MLP), F. Fillon (FF) and JL Mélenchon (JLM) + B. Hamon (BH) as a viable candidates under EV
- 6 minor candidates : N. Dupont-Aignan (NDA), Jean Lassalle (JL), P. Poutou (PP), F. Asselineau (FA), N. Arthaud (NA) and J. Cheminade (JC).

Exclusive vs. inclusive (major candidates)

Baujard et al. 2014, for the French Pres. election, distinguished :

- The exclusive candidates : who arouse strong feelings, whether positive from their voters, or negative from the other voters ;
- The inclusive candidates : who are supported by a large number of voters (but not necessarily strongly valued).

Polarizing/(un-)popular/medium (4 types of candidates)

Darmann et al. 2017, for the Austrian parties of the Styrian Parliament :

- polarizing : Strong support from a large part of society and strong negative support from another large part.
- popular : Strong support for a specific segment of society and seen positively by a large part of society.
- medium : Acceptable by a large part of society and strong (positive or negative view) in a small group.
- unpopular : Strong support from only a small group and seen negatively by a large portion of society.

Yet... impossible qualification or comparisons upon uncorrected results.

Inclusive and *Exclusive* candidates among voters who do not approve them

Major candidates may be either :

Inclusive candidates attract positive feelings, including for voters who do not approve them. EM, JLM, BH

Exclusive candidates who are strongly rejected by voters who do not approve them. MLP, FF

Unpopular and uncovered

Minor candidates may be either :

Unpopular with a quite known political line. Voters can form an opinion. NDA, PP, NA.

Uncovered who cannot be seen neither positively nor negatively because almost unknown, because of a low media coverage outside electoral moments. JL, FA, JC

Assumption AN – Invariance with negative label

For each candidate, the distribution of the grades is translated by one unit when comparing EV3 with EV3neg, and EV4 with EV4neg.

- AN1. Effects on scores
- AN2 & AN3. Effects on the frequency of use of lower or larger grades

We can work this out by comparing EV2 (benchmark), EV3 and EV3neg, i.e. by using the 2017 Strasbourg data.

EV-Test Group	Sample size
{0 ;1 ;2}	251
{-1 ;0 ;1}	247
{0 ;1 ;2 ;3}	282
{-1 ;0,1 ;2}	236

Effect on scores – Assumption AN.1

For each candidate, the score is 1-unit shifted when comparing EV3 with EV3neg, and EV4 with EV4neg.

Comparison of scores in scales with and without negative grade

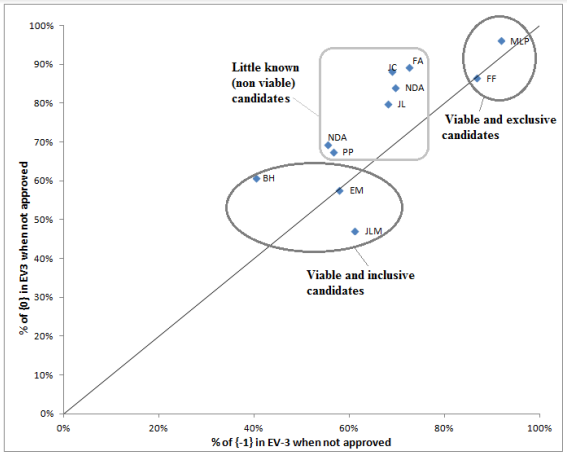
Candidates	EV3	EV3neg normalized	Test EV3-EV3neg	EV4	EV4neg normalized	Test EV4-EV4neg
EM	1.01	1.11	Non Sign.	1.37	1.40	Non Sign.
MLP	0.15	0.19	Non Sign.	0.30	0.23	Non Sign.
FF	0.31	0.42	Non Sign.	0.55	0.55	Non Sign.
JLM	1.22	1.21	Non Sign.	1.57	1.79	Significant
BH	1.09	1.28	Significant	1.48	1.74	Significant
NDA	0.27	0.51	Significant	0.44	0.54	Non Sign.
JL	0.30	0.44	Significant	0.45	0.62	Significant
PP	0.57	0.74	Significant	0.86	1.07	Significant
FA	0.20	0.36	Significant	0.26	0.48	Significant
NA	0.47	0.63	Significant	0.62	0.91	Significant
JC	0.14	0.37	Significant	0.25	0.49	Significant

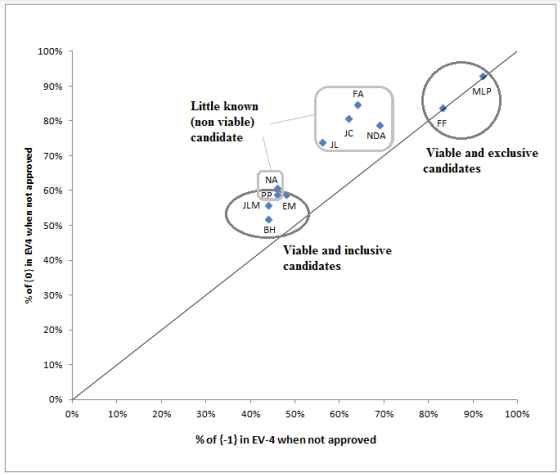
The introduction of the negative grade significantly increases the scores of all minor candidates and BH. – hence a (only) relative impact on major candidates, and especially disfavor exclusive candidates.

(Hence the necessity to disentangle behaviors concerning approved or non approved candidates.)

Effect on the frequency of lower grades – AN2

Up to a 1-pt translation, the different scales with and without negative grade generate the same proportion of lower grade for each candidate when not approved.



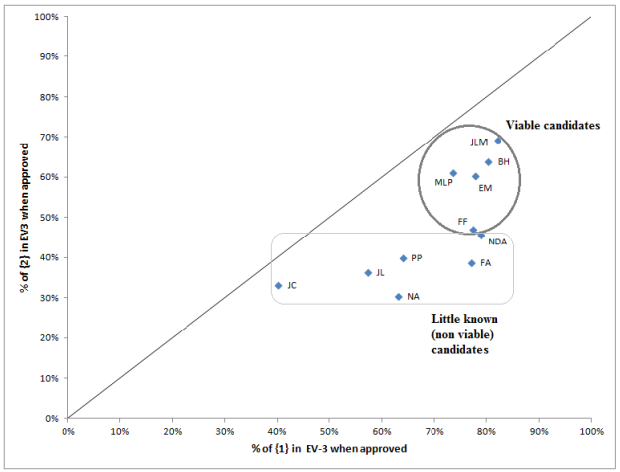


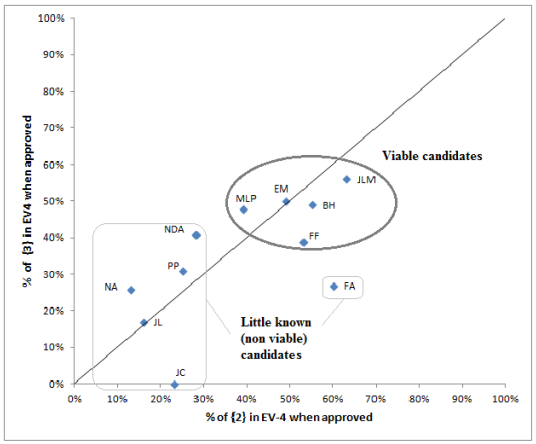
We observe a negative grade effect in both scales :

- AN2 is rejected for minor candidates : as they are usually not rejected as such, many voters do not want to give them the lowest grade if negative.
  Minor candidates receive less -1 than 0.
- AN2 is not rejected for exclusive candidates : voters do give them the lowest grade, whichever -1 or 0.
  Exclusive candidates receive as much lowest grade in both grading scales.
- AN2 is not verified for inclusive candidates : Not a sharp effect as inclusive candidates have few lowest grades in any case.

Effect on the frequency of highest grades – AN3

Up to a 1-point translation, the different scales with and without negative grade generate the same proportion of highest grade for each candidate when he/she approved.





- From EV3 to Ev3neg, AN3 is clearly rejected : all approved candidates receive more highest grade :
- This is no longer true from EV4 to EV4neg. No obvious trend is emerging either for major candidates or for minor candidates.
- The differentiated evolution of the scores depends mainly on the behavior of the voters when they do not approve candidates. The changes due to introduction of negative grade are mostly favorable to minor candidates.

Assumption AL –Invariance with length

Numerical scales of different lengths are linearly equivalent.

We can work this out by comparing EV2 (benchmark), EV4 (x2), and EV6 (x3), i.e. by using the 2017 Hérouville-Saint-Clair data.

EV-Test Group	Sample size
{0 ;1}	667
{0 ;1 ;2 ;3}	350
{1 ;2 ;3 ;4 ;5}	311

Invariance with length – AL.1

Comparisons of distribution. The distribution of grades is left unchanged for each candidate for various x-step scale (AV, EV4 and EV6) when reduced linearly to two classes.

Reduction to two classes : grade distribution (Hérouville data)

Candidates	AV (N=667)		EV4 (N=354)		EV6 (N=313)	
	$\{0\}$	$\{1\}$	$\{0,1\}$	$\{2,3\}$	$\{0,1,2\}$	$\{3,4,5\}$
EM	51 %	49 %	52 %	48 %	49 %	51 %
MLP	88 %	12 %	87 %	13 %	88 %	12 %
FF	81 %	19 %	80 %	20 %	81 %	19 %
JLM	46 %	54 %	44 %	56 %	48 %	52 %
BH	48 %	52 %	50 %	50 %	50 %	50 %
NDA	87 %	13 %	87 %	13 %	87 %	13 %
JL	95 %	5 %	97 %	3 %	95 %	5 %
PP	75 %	25 %	80 %	20 %	76 %	24 %
FA	95 %	5 %	97 %	3 %	94 %	6 %
NA	85 %	15 %	87 %	13 %	88 %	12 %
JC	95 %	5 %	97 %	3 %	95 %	5 %
Average	77 %	23 %	78 %	22 %	77 %	23 %

- At the 1% threshold, one cannot reject the hypothesis that the two probabilities of choosing in the scope of lower grades class are the same under EV4 and EV6. Hence AL1 is verified.
- But, looking closely, significant differences arise in the details, but they happen to be compensated in our specific data set.

Length effect – AL2

The normalized scores of each candidates are left unchanged for various scales.

Reduction to two classes : comparison of scores (Hérouville data)

Candidates	AV	EV4 normalized	EV6 normalized	Test AV-EV4	Test AV-EV6	Test EV4-EV6
EM	0.49	0.48	0.45	Non Sign.	Non Sign.	Non Sign.
MLP	0.12	0.13	0.11	Non Sign.	Non Sign.	Non Sign.
FF	0.19	0.19	0.19	Non Sign.	Non Sign.	Non Sign.
JLM	0.54	0.53	0.49	Non Sign.	Non Sign.	Non Sign.
BH	0.52	0.48	0.48	Non Sign.	Non Sign.	Non Sign.
NDA	0.13	0.14	0.15	Non Sign.	Non Sign.	Non Sign.
JL	0.05	0.1	0.12	Significant	Significant	Non Sign.
PP	0.25	0.24	0.25	Non Sign.	Non Sign.	Non Sign.
FA	0.05	0.06	0.09	Non Sign.	Significant	Significant
NA	0.15	0.18	0.17	Non Sign.	Non Sign.	Non Sign.
JC	0.05	0.06	0.08	Non Sign.	Significant	Significant

- The differences of scores of candidates are not significant for many candidates.
- A noticeable exception : longer scales favor uncovered candidates.

A shift in the use of extreme vs. intermediate grades

AL1 is totally verified (grade distribution stability for our specific data set) but AL2 is only partially verified (score stability), i.e. not for uncovered candidates. Why ?

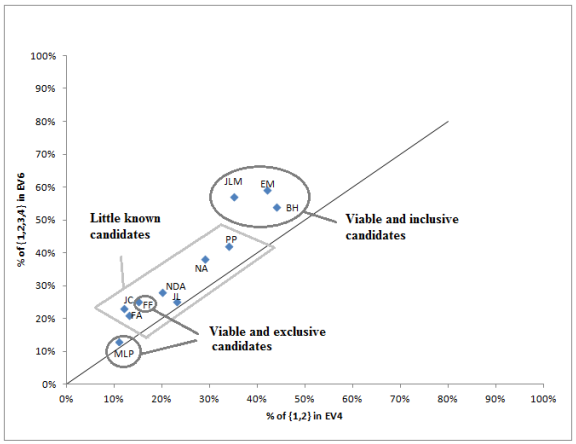
- Most voters have not strategic concern but sincere/expressive concern. The fact that they use many intermediate grades may potentially affect the candidates scores.
- The outcome results from two opposite effects :
  - When a non-approval becomes an intermediate grade (rather than the lowest grade), the score tends to increase.
  - When an approval becomes an intermediate grade (rather than the highest grade), the score tends to decrease

Hence non approved candidates are more likely to be favored,
and approved candidates are more likely to be unfavored by longer lengths.
Among non-approved candidates, exclusive candidates or well identified unapproved candidates keep with the lowest grades. But there is tendency to shift to a low intermediate grade when candidates are unknown. This is favorable for uncovered candidates.

A shift in the use of extreme vs. intermediate grades

Strategic voting. Hypothesis (AL3)

The length of the scale does not change the use of extreme grades for any candidates.

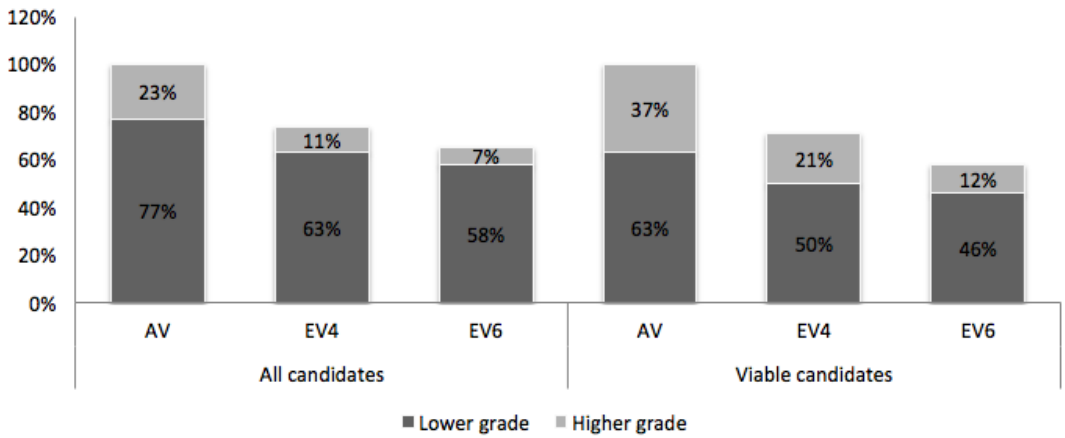


AL3 is rejected :

- The longer the scale, the more voters use intermediate grades.
- More for inclusive candidates than for all others.

The use of the entire grading scale.

- AG1 : For all scale lengths, all voters use the entire scale of grades
- AG2 : The lengthening of the scale does not modify the proportion of voters who use the full extent of the grading scale.



AG1 and AG2 are rejected.

- an inclination toward expressive rather than strategic behavior (as if a confusion between grading and expressive behavior)
- a limit to the equal representativeness of all voters

Concluding remarks

Grading scales matter : Voters rank candidates consistently but in distinct manners for different grading scales, hence different electoral outcomes (or significant yet compensated differences that may not always turn into different outcomes).

1. Scales with negative grades favor minor candidates. This result is much sharper than the standard impression that negative grades disfavor exclusive candidates, which is just an indirect result.
2. Length matters, as far as longer scales favor uncovered candidates
3. These observation derive from the shift in using intermediate grades, induced by an expressive vote in EV rather than strategic behaviors.

As far as longer scales are likely to favor unapproved uncovered unpopular candidates, and induce unequal weight among voters, we question the desirability of longer lengths.

Thank you !