

## Modulation of the leniency bias in the discursive dilemma

Gustavo Bodanza<sup>1,2</sup>, Esteban Freidin<sup>2</sup>, Sebastián Linares<sup>2</sup>,  
and Fernando Delbianco<sup>3</sup>

<sup>1</sup>Departamento de Humanidades, Universidad Nacional del Sur, Buenos Aires, Argentina

<sup>2</sup>Instituto de Investigaciones Económicas y Sociales del Sur (IIESS, UNS-CONICET), Buenos Aires, Argentina

<sup>3</sup>Departamento de Economía, Universidad Nacional del Sur, and Instituto de Matemática de Bahía Blanca (INMABB, UNS-CONICET), Buenos Aires, Argentina

We experimentally approach the discursive dilemma to gain insight into people's procedural appropriateness judgments. We relied on a vignette in which three people had formed opinions about two skills (premises) of a candidate to decide whether to hire her/him (conclusion). The dilemma arises when different outcomes (hire vs. not hire) are achieved depending on whether the majority opinion is independently considered for each premise or for the global conclusion of each judge. Participants were asked to choose the procedure they thought to be more appropriate to reach a decision. In Experiment 1, we found a leniency effect (a bias to prefer the aggregation procedure that led to hiring the candidate), which was reduced by introducing the participant as a juror with an exogenously provided negative opinion about the candidate's skills. In Experiment 2, we replicated the opinion effect, even when subjects did not participate as jury members. In Experiment 3, we found that the leniency bias was only reduced when participants' negative opinion was aligned with a majority of negative premises, but not with a majority of negative conclusions. We discuss present findings in terms of the identification of empirical regularities that may affect people's procedural legitimacy judgments.

**Keywords:** Discursive dilemma; Collective decision making; Leniency bias; Confirmation bias.

Kornhauser and Sager (1993) have referred to the problem arising in collegial courts when different decisions can be obtained depending on the aggregation method used. This inconsistency can happen when a conclusion is reached by aggregating the votes of individual judges either on the overall subject of discussion or on each relevant issue separately. Table 1 resumes a decision context in which a jury formed by three judges has to make a decision for "yes" or "no" with regards to a proposition  $r$ , which is implied by the conjunction of two propositions  $p$  and  $q$ . The opinion of each judge about  $p$ ,  $q$  and  $r$  is showed in the rows. A dilemma is here manifest: when applying majority voting on the conclusion  $r$  directly (henceforth, *compound* procedure), we obtain a different outcome than when applying majority voting on the elements  $p$  and  $q$  and then "deriving" the collective opinion on  $r$  (henceforth, *elemental* procedure).

This is an old paradox known as *discursive dilemma* (DD) or *doctrinal paradox* (Kornhauser & Sager, 1993)

and gave rise to contemporary Judgement Aggregation Theory (List & Puppe, 2009). Several philosophical, normatively oriented proposals have been advanced to solve the paradox (Bovens & Rabinowicz, 2003, 2006; Brennan, 2001; Cariani, Pauly, & Snyder, 2008; Chapman, 2002; Dietrich & List, 2007; Hartmann & Sprenger, 2012; List, 2006; Pettit, 2001). The vast majority of developments around collective judgement aggregation are oriented towards responding what are the normatively appropriate responses to paradoxes or other unsolved situations (Regenwetter et al., 2009). Little attention, however, has been directed at introducing a behavioural perspective to its analysis, in particular, for the DD (Bonneton, 2007, 2010). This study aims at contributing to this last perspective.

We address the empirical question of what procedure people would preferentially choose to escape the DD paradox. People's judgement on procedural adequacy for group decisions and the factors that modulate it have been

Correspondence should be addressed to Gustavo Bodanza, Viamonte 688, (8000) Bahía Blanca, Argentina. (E-mail: bodanza@gmail.com)

This work was partially supported by ANPCyT (PICT 2013-1489, PICT 2017-1702), Universidad Nacional del Sur (PGI 24/I223) and CONICET (PIP 2014 112-201301-00600-CO), Argentina.

G. B. is contributed to conception, design, collection, analysis and interpretation of data, drafting the article. E. F. is contributed to conception, design; analysis and interpretation of data; drafting and revising the article. S. L. is contributed in conception; interpretation of data; critical revision of the article for important intellectual content. F. D. is contributed to analysis and interpretation of data, statistical analysis, Classification and Regression Tree (CART) analysis and software management.

**TABLE 1**  
The discursive dilemma

	<i>p</i>	<i>q</i>	<i>r</i>
Judge 1	Yes	No	No
Judge 2	No	Yes	No
Judge 3	Yes	Yes	Yes
Majority	Yes	Yes	No

widely studied in the field of social-psychology (Lind & Tyler, 1988). The pioneering work of Thibaut and Walker (1975) showed that perceptions of procedural justice result in increased satisfaction, and are among the most important determinants of procedural preferences. This may apply to a variety of social contexts in which the aggregation of votes, decisions or opinions matter, such as in private organisations, political parties, or legal institutions (Tyler, 2011). In turn, this knowledge may offer useful insights for policymakers or authorities, more generally, given that understanding the factors that modulate perceptions of legitimacy provides elements to enhance people's satisfaction with authorities and institutions, and also normative attitudes and behaviours (Tyler, 2006).

Some empirical studies on deliberation styles of juries set relevant precedents for the present work. Hastie, Penrod, and Pennington (1983) showed that individuals were more satisfied with the group decision when they could deliberate about the elements supporting decisions than when they deliberated only on individual verdict preferences. Moreover, Kameda (1991) showed that individuals who deliberated about the elements were more prone to accept a decision contrary to their preference than those who just voted on the verdict. Furthermore, MacCoun and Kerr (1988) reported a series of experiments in which participants showed a tendency to choose procedures leading to benevolent outcomes in mock juries, and referred to this phenomenon as a leniency bias. The authors observed an *asymmetry effect* in dichotomous decisions: in juries with equal initial splits about the verdict and with no clear predominant individual preference, the election was more inclined to acquittal than to condemn. Furthermore, some authors argue that the tendency to prefer decisions leading to favourable outcomes is stronger than the preference for fair procedures (Esaiaasson, Persson, Gilljam, & Lindholm, 2016).

Regarding the DD, Bonnefon (2007, 2010) asked what procedure, elemental or compound, individuals would choose in that context. He found a pervasive leniency bias with French university students as subjects. The dilemma was presented in the context of the evaluation of an employee according to her/his competence and motivation for the job. Two versions, one conjunctive (*competent and motivated*) and one disjunctive (*competent or motivated*) were combined with two descriptive framings, one positive (*competent, motivated*)

and one negative (*incompetent, unmotivated*). Participants had to evaluate each deliberation style in a five points scale and then answer the question: *Is it true that the employee is [in]competent [and/or] [un]motivated?* In each case, subjects tended to choose the procedure promoting the acceptance (rejection) of the conclusions *competent and/or motivated (incompetent and/or unmotivated)*. This confirmed a leniency effect, understood in the general sense of a compassionate attitude.

## PROPOSAL

In view of Bonnefon's findings, we wondered whether such leniency bias can be modulated by other factors or it is a robust phenomenon in the context of the DD. Given the previously mentioned MacCoun and Kerr's (1988) asymmetry effect, we were also concerned with a possible incidence of participants' opinions when they were not neutral. In this sense, we tested for the effect of the presence/absence of an exogenously provided opinion to participants. This is relevant in the present context because a confirmation bias would incline people to choose the procedure which outcome coincided with their own opinion, which, if negative, would not coincide with predictions derived from a leniency bias. We contrasted the strengths of these two biases. Moreover, since individuals in Bonnefon's experiments were consulted as non-jury members, we studied whether individuals engaged in the decision as members of the jury would show a different inclination in the choice of a deliberative style, relative to non-juries with the same opinion. Indeed, Thibaut and Walker (1975) argued that procedures were viewed as fair when disputants could voice their concerns in an effort to influence the decision outcome. In this sense, we expected provided opinions to exert a stronger effect in conditions in which participants had a jury role, than when they did not. Last but not least, present experiments test the leniency bias with an Argentine sample, which jointly with previous studies in other countries, contribute to forging a cross-cultural perspective in the study of legitimacy perception of decision-making processes.

We report here a series of three experiments, relying on a vignette in which a jury of three members had to evaluate a candidate for an academic position with regards to two dimensions, teaching and research skills (Bovens & Rabinowicz, 2003; List, 2006).

## EXPERIMENT 1

### Participants

Participants were 539 voluntary undergraduates from Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina. We had no hypothesis regarding either gender or age, so we did not collect data on those variables

in these experiments. We relied on students from a wide variety of disciplines (mainly accountancy, philosophy, geography, architecture and law) which usually deliver a balanced representation of gender. In terms of age, the vast majority of students at UNS are on their twenties.

## Material and procedures

The protocol used in our experiments was approved by the Ethics Committee of the Hospital Municipal “Dr. Leónidas Lucero” of Bahía Blanca, Argentina. The paradox was framed within the context of the acceptance (lenient result) or rejection (negative result) of a candidate for an academic position, based on the evaluation of her/his academic skills on teaching and research. We expected to observe a leniency bias tilting the choice towards the deliberation style promoting the positive outcome. Moreover, we wanted to test the incidence of an exogenously provided positive or negative opinion. Our hypothesis was that a positive opinion would reinforce the leniency bias, while a negative opinion would counterbalance it.

The study used two main conditions that were randomly distributed among the participants. In the opinion condition ( $n = 401$ ), participants were assigned the role of jury members and an opinion about the candidate’s relevant skills. Jointly with their opinion, participants read the opinions of two other juries. In the no-opinion condition ( $n = 138$ ), participants simply read the same three opinions, but no opinion was attributed to themselves. The methodological decision of exogenously assigning an opinion to participants was done to warrant the display of the DD, which is the context in which we wanted to evaluate participants’ choice of a deliberation style.

The opinion condition was further divided according to which deliberative style (elemental or compound) was associated with the lenient outcome, and whether the assigned opinion was positive or negative. This design resulted in four between the subject conditions nested in the opinion condition: (a) elemental-lenient negative opinion ( $n = 104$ ), (b) elemental-lenient positive opinion ( $n = 40$ ), (c) compound-lenient negative opinion ( $n = 128$ ) and (d) compound-lenient positive opinion ( $n = 129$ ). In turn, the no-opinion condition was divided in two between-subject conditions according to which deliberation style led to the lenient outcome: (e) elemental-lenient no-opinion ( $n = 54$ ); and (f) compound-lenient no-opinion ( $n = 84$ ). Sample sizes in the different experimental conditions could not be determined a priori because some of the tested factors had not been explored before. In this sense, differences in sample sizes among conditions were driven by differences in effect sizes among comparisons in the same experiment. Importantly, the robustness of the main findings was checked with subsequent independent conditions (see subsequent experiments).

Participants were provided with a written sheet of paper (in Spanish) with a preamble saying:

*Next you will read a situation in which a decision has to be made and we ask you to evaluate, according to your own opinion, what is the best criterion for the decision to be based on.*

Next, the situation was described. In the opinion conditions, it read as follows:

*You are a member of an academic jury which has to decide by vote the hiring of a new professor. There are two relevant conditions for accepting a candidate, namely her/his teaching and research skills.*

In order to obtain the elemental-lenient conditions ([a], [b] and [e]), a conjunctive requirement was presented:

*If, in the jury’s opinion, both conditions are fulfilled then the candidate will be accepted.*

In order to obtain the compound-lenient conditions ([c], [d], and [f]), a disjunctive requirement was presented:

*If, in the jury’s opinion, at least one of those conditions is fulfilled then the candidate will be accepted.*

The text continued informing that there was an applicant to the position. Next, the opinion of the participant was provided in the phrase:

*In your opinion, the candidate ...*

followed by:

- *has enough background in teaching but not in research, hence your vote is for rejecting the candidate. (condition [a])*
- *has enough background in both teaching and research, hence your vote is for accepting the candidate. (condition [b])*
- *has not enough background neither in research nor in teaching, hence your vote is for rejecting the candidate. (condition [c])*
- *has enough background in research but not in teaching, hence your vote is for accepting the candidate. (condition [4])*

The opinions of the other two members of the jury were introduced in such a way that the paradox ensued. A remark was made about the fact that the majority of vote on the acceptance/rejection of the candidate was inconsistent with the majority verdict on each of the relevant elements (i.e., teaching and research backgrounds) considered separately, thus making the DD explicit.

The no-opinion conditions ([e] and [f]) were the same as the opinion versions but without any reference of the participant as a jury member.

A table summarising the information on opinions and conclusions was introduced afterwards, similarly to Table 1, but leaving blank the decision cell in the lower right-most corner. Moreover, the cells corresponding to the majority's opinion about each element, teaching and research, included the legend "2 out of 3 accept/reject the candidate's teaching/research skills."

Finally, the same question was asked in all versions, just varying which deliberative style led to the lenient outcome as expressed below:

*In your opinion, what should be the right procedure for the jury to arrive to a decision? Choose (1) or (2):*

(1) *To consider the jury members' opinions on each element (teaching and research) separately, and then accept/reject the candidate, or.*

(2) *To consider the general conclusion about the candidate of each jury member and, then, reject/accept the candidate.*

## Results and discussion

The results are succinctly presented in Table 2. Overall, there was a predominant tendency of participants to prefer the procedure leading to the benevolent result for the candidate, except for the compound-lenient negative opinion condition (c), in which participants showed indifference between deliberation styles.

We began by comparing the decisions in conditions with no-opinion ([e] and [f]): the majority preferred aggregating premises when that led to accepting the candidate, whereas only a minority chose the process of aggregating premises when that led to rejecting the candidate. This same trend was found in the conditions in which participants were provided with a positive opinion of the candidate ([b] and [d]): participants were inclined to aggregate premises when that led to accepting the candidate, whereas only a minority preferred to aggregate premises when that led to rejecting the candidate. As mentioned before, the only instance in which the lenient bias was not observed was when participants' opinion was to reject the candidate in the compound-lenient condition (c). Indeed, the lenient bias appeared again in the elemental-lenient condition even when participants' provided opinion was against accepting the candidate (a). These last two results suggest that participants' opinion for rejecting the candidate interacted with the aggregation method leading to the lenient outcome. In fact, in the compound lenient conditions, having a negative opinion of the candidate (c) led to a significantly higher proportion of choices for the procedure leading to rejection than in the condition without opinion (f). In contrast, the prevailing tendency to choose the procedure leading to the lenient outcome was not significantly reinforced when subjects' provided opinion favoured that result. Actually, the elemental-lenient condition with a positive

opinion (b) showed a lower (though not significant) proportion of choices for aggregating the premises than the elemental-lenient condition with no opinion (e).

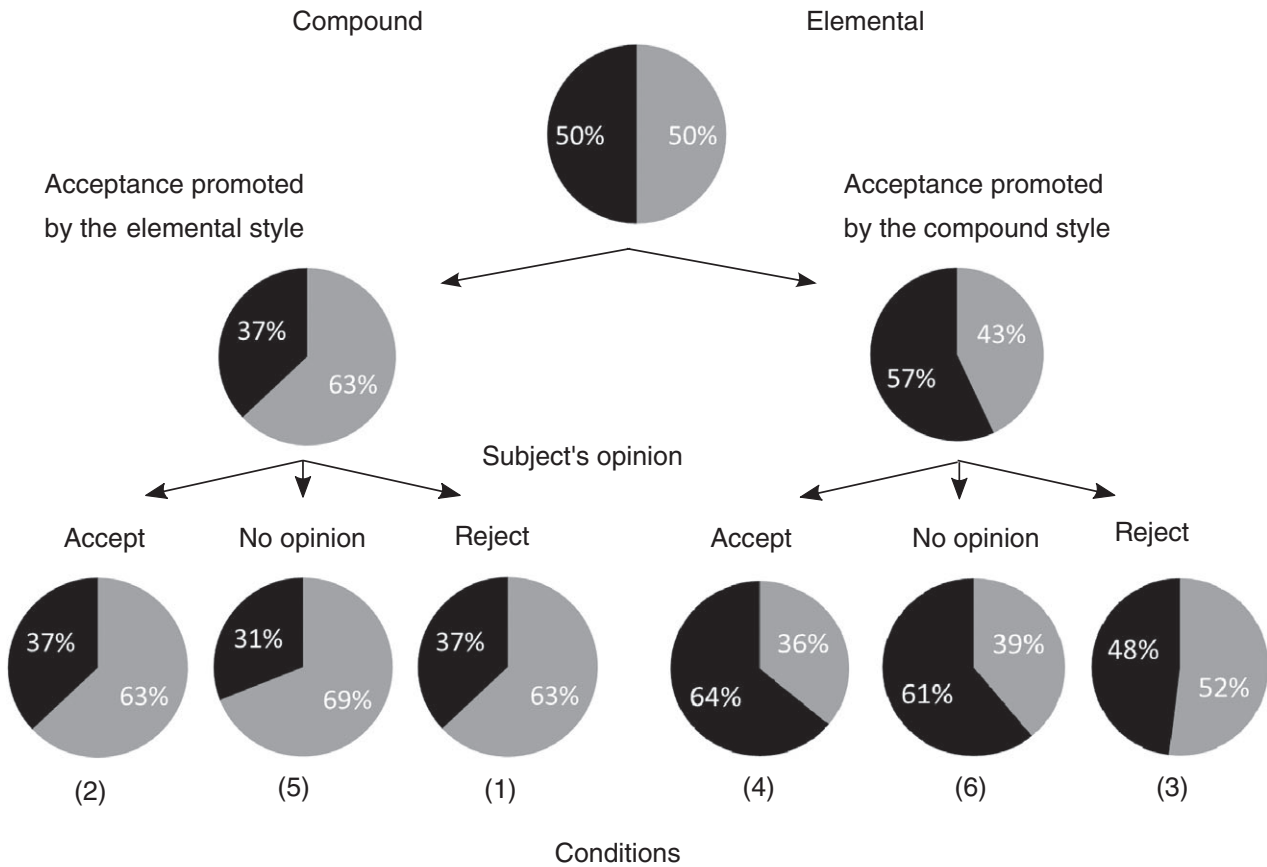
All in all, we conclude that the leniency bias was pervasive across conditions. Only an opinion of rejecting the candidate jointly with a majority of premises pointing in that direction led to a decrease in the leniency bias, but only to take choices to indifference between deliberation styles.

As a corroboration of the previous results, we present a classification and regression tree analysis CART (Breiman, Friedman, Olshen, & Stone, 1984) made with R software. This method delivers a hierarchy of explaining factors obtained from the raw data, that is, without any theoretical pre-specification. The analysis predicts a response or class  $Y$  from inputs or explicative variables  $X_1, X_2, \dots, X_p$  by growing a tree with binary branches. At each internal node in the tree, a test to one of the inputs, e.g.  $X_i$ , is performed. Depending on the outcome of the test, the next step is to go either to the left or to the right sub-branch of the tree. Eventually, a leaf node is reached, where a new prediction is made. This prediction aggregates or averages all the training data points which reach that leaf. A tree for classifying subjects according to the choice Elemental/Compound was programmed, yielding the data plotted in Figure 1 (all percentages rounded). As seen at the tree root, 539 observations gave a mean of 50% of the individuals choosing each style. The first partition is made with respect to the deliberation style promoting the accept result. Both nodes show the tendency to choose the procedure yielding the accept outcome, elemental in the left node and compound in right node. So far, the choice of deliberation style seems to be inclined towards that yielding the more benevolent result, though the inclination is greater when a positive majoritarian opinion is observed about the elements. Now, a second partition is made according to whether the subjects are provided with an acceptable opinion, a reject opinion, or no opinion. On the left hand, where the elemental style yields the accept outcome, the pies corresponding to conditions (b), (e) and (a) are very similar. Hence, the choice of an elemental style when it yields the accept outcome is not significantly influenced by the opinion assigned to the subjects. On the right hand, where the accept outcome is promoted by the compound style, the choice was inclined towards the procedure leading to the lenient outcome when subjects were either provided with the accept opinion (d) or no opinion (f). On the contrary, a slight majority of the subjects who were provided with the reject opinion were inclined to the elemental style, which led to rejecting the candidate (c). This shows that the benevolent outcome is partially reverted when participants' provided negative opinion on the candidate coincided with a rejection based on an elemental style (i.e., based on negative premises).



**TABLE 2**  
Dependency of the choice of deliberation style on the result promoted by each style and the subject's assigned opinion.

Condition	Independent variables			Dependent variable		(*)	Binomial prob. test
	Outcome by style:		Subject's opinion	Subject's choice:			
	Elemental	Compound		Elemental (%)	Compound (%)		
(a)	Accept	Reject	Reject	62.5	37.5	a	$p < .007$
(b)	Accept	Reject	Accept	62.5	37.5	a	$p < .08$
(c)	Reject	Accept	Reject	52.3	47.7	a	$p < .33$
(d)	Reject	Accept	Accept	36.4	63.6	b	$p < .001$
(e)	Accept	Reject	—	68.5	31.5	a	$p < .004$
(f)	Reject	Accept	—	39.3	60.7	b	$p < .03$



**Figure 1.** CART analysis of the data shown in Table 2.

The observed leniency bias could be the consequence of an implicit rapport or identification of the participants, undergraduate students, with a candidate to occupy an academic position. We believe this to be unlikely given Bonnefon’s (2010) results showing a similar positive bias towards the candidate in a different, non-academic, framing about the evaluation of an employee. Still, one could imagine scenarios where the leniency or benevolence could diminish. Indeed, to complement this point, we tested a framing revolved around the hiring of a surgeon in a hospital, considering her/his degree of knowledge to solve complex cases and experience to easily solve the

most common cases. Naturally, we expected a significantly lesser benevolence towards the candidate (relative to the professor) given the potentially harmful consequences of hiring an incompetent surgeon. We used the medical scenario with accountancy undergrads ( $N = 210$ ) to test whether differences in leniency were observed when a provided negative opinion coincided with a rejection based on the elemental style. We tested two conditions structurally similar to (c) and (d). As a result, we replicated the significant difference in the leniency bias (around 16 percentage points). The compound style, leading to the lenient outcome, was chosen by 47% of the

participants when their provided opinion was positive, while only 31% of the participants with negative opinion chose that style (Fisher's exact test;  $p < .05$ ). Overall, we found that beyond the specific level of leniency shown by participants in a given frame, the modulation of such leniency by the interaction between negative premises and a participant's own negative opinion is robust to framing changes.

We derived two follow-up questions from the results of Experiment 1. First, we asked whether the particular effect was seen in condition (c) depended just on the provided negative opinion or on that opinion together with the participants' role as jury members. A confirmation bias effect may not distinguish between these two possibilities. However, it can be thought that having an active voice (vote) may strengthen the bias towards the provided opinion. Indeed, some findings suggest that people perceive procedures as fairer when they warrant a "voice" to those involved, or in other words, they allow the expression of opinions (Hulst, van den Bos, Akkermans, & Lind, 2017; Lind & Tayler, 1988; van den Bos, 2005). We tackled this question in Experiment 2 by providing participants with a negative opinion on the candidate, but, this time, without them being part of the jury. The second question about the result observed in condition (c) had to do with how strong the provided opinion is and whether that affects the modulation of the leniency bias observed. This was tackled and further explained in Experiment 3.

## EXPERIMENT 2

### Materials and procedure

We recruited 144 voluntary undergraduate students of various disciplines (mainly literature, geography, biology, computation, architecture and medicine) at UNS. We tested whether positioning participants as jury members had an effect on the choice of a deliberative style. That is, in Experiment 1, we observed that a negative opinion assigned to subjects positioned as members of the jury, together with a majority of negative premises (condition [c]), led to a decrease in the leniency bias relative to that shown by participants in the other conditions. We tested the compound-lenient negative opinion condition but, this time, subjects were positioned out of the jury (we call this condition (c'): compound-lenient no-jury negative-opinion). The sample size was intended to be similar to that of condition (c). Table 3 shows the summary information seen by the participants in this condition.

The negative opinion was assigned as follows:

*In your opinion, the candidate has not enough background, neither in research nor in teaching.*

**TABLE 3**  
Information summary read by participants (Experiment 2)

	<i>Teaching</i>	<i>Research</i>	<i>Accept or not?</i>
Judge 1	No	No	No
Judge 2	Yes	No	Yes
Judge 3	No	Yes	Yes
Majority	2 out of 3 reject the teaching skills	2 out of 3 reject the research skills	—

### Results and discussion

We observed a very similar result to that observed in the analogous condition (c) from Experiment 1 in which participants were jurors: whereas in Experiment 1, 52.3% of participants chose the elemental style leading to rejecting the candidate, here 52.1% (75 out of 144) of the participants chose that same response (Binomial Probability test,  $p = .34$ . Comparison between these conditions: Fisher's exact test,  $p > .50$ ). In addition, 47.9% (69 out of 144) of participants in condition (c') chose the procedure leading to the lenient outcome, which is significantly less than the 60.71% of participants choosing the lenient outcome in condition (f) (compound-lenient no-opinion) in Experiment 1 (Fisher's exact test,  $p < .05$ ). These two results suggest that providing participants with a negative opinion of the candidate decreased the probability of choosing the procedure leading to the lenient outcome, in particular, when that procedure was based on aggregating conclusions. Moreover, results from Experiment 2 suggest that the provided opinion exerted a similar effect regardless of whether participants were in the role of jury members.

Results from Experiments 1 and 2 took us to attempt to answer another question: Why did the modulation of the leniency effect only occur when participants' provided negative opinion coincided with a majority of negative premises, but not when it coincided with a majority of negative conclusions? Note that in conditions in which the modulation of the leniency effect was found in Experiments 1 and 2, participants were positioned in a strongly negative opinion of the candidate, that is, they rejected both candidate's teaching and research skills, in the context of a majority of negative premises of the overall jury. In contrast, when we tested the effect of providing a negative opinion in the context of a majority of negative conclusions of the overall jury, participants' provided opinion was only weakly negative (rejecting one skill while accepting the other). This asymmetry between negative opinion conditions (strongly negative vs. weakly negative) had to do with necessary conditions for obtaining the paradox present in the DD. If participants would have a strong negative opinion in the context of an elemental-lenient condition (i.e., in which the aggregation of conclusions leads to

**TABLE 4**  
Information summary read by participants in Experiment 3

	<i>Teaching</i>	<i>Research</i>	<i>Accept or not?</i>
Judge 1	No	Yes	No
Judge 2	Yes	No	No
Judge 3	Yes	Yes	Yes
Majority	2 out of 3 accept the teaching skills	2 out of 3 accept the research skills	—

rejecting the candidate), there would be no paradox neither with a disjunctive nor with a conjunctive condition (see Appendix).

### EXPERIMENT 3

#### Materials and procedure

The participants were 60 voluntary undergraduate students of different study disciplines (mainly from literature, geography, biology, computation, architecture and medicine) at UNS. We used the conjunctive version in which the compound style led to rejecting the candidate (condition [e]), but this time the participants were provided with a strong negative opinion (rejection of both teaching and research skills), though they did not participate as jury member. The sample size was purposely similar to that used in condition (e). The information for the subjects was summarised in Table 4.

To position the subjects in a strongly negative opinion we include the following sentence:

*Assume that you have expertise to evaluate the candidate to the same extent as the juries and, in your opinion, the candidate neither has enough skills in teaching nor in research.*

The same questionnaire as that used in Experiment 1 was introduced next.

#### Results and discussion

The elemental style, yielding the lenient outcome, was chosen by 68.3% of the individuals (41 out of 60; Binomial Probability test,  $p = .003$ ). This percentage is almost the same as that observed in the similar condition (e) with no opinion (68.5%) and, curiously, slightly greater—though not significantly so—to condition (1) with a positive opinion (62.5%) in Experiment 1. Therefore, we conclude that providing participants with a strongly negative opinion was not enough to moderate the leniency bias.

### GENERAL DISCUSSION

The choice of a deliberative style in the context of the DD has been shown to be sensitive to framing effects

(Bonnefon, 2007, 2010). In particular, a bias towards choosing procedures leading to outcomes described in more lenient terms has been reported in the DD as well as in other decision situations (e.g., Esaiasson et al., 2016). The present study aimed at contributing with behavioural experiments to the incidence of specific factors that could tilt the choice of a deliberative style. We used a vignette scenario to test the incidence of three factors: the leniency bias, the confirmation bias and the positioning of participants as jury members. Indeed, we confirmed a pervasive leniency bias. Participants showed an overall preference for the procedure (elemental or compound) leading to the favourable outcome for the professorship candidate in 6 of the 8 conditions, whereas they showed indifference between deliberative styles in the remaining 2 conditions. This bias was only reduced when participants were provided with a strong negative opinion of the candidate while the jury's majority opinion on the candidate's skills was negative as well (this effect was indeed robust to framing changes). Therefore, we conclude that the pervasive leniency effect was modulated by a confirmation bias, but only under very specific conditions. Last, we found that this reduction of the leniency bias occurred independently of whether the participant was involved as jury or, instead, had an outsider opinion.

The relevance of the present research relies on its connection with procedural fairness research. The study of people's preferences for group or institutional decision procedures may serve to gain knowledge on the predictors of legitimacy judgments, decision acceptance and people's inclination to comply with authorities' decisions. At its best, present research may help authorities to implement procedures for which there is evidence of strong acceptance. Nonetheless, as present and previous findings indicate (Bonnefon, 2010), people's procedural preferences may be affected by factors beyond those concerned with decision procedures. In this sense, there is no apparent procedural preference, but participants' choices of decision style switched according to the outcome that it provided. In line with this, Esaiasson et al. (2016) posed that the importance of procedural fairness preferences is eclipsed by this type of outcome favourability.

Present results confront us with the challenge of finding a suitable explanation for the interaction between the leniency effect and the confirmation bias. More specifically, the outcome favourability was mitigated when participants' strong negative opinion on the candidate coincided with a majority of negative opinions on the candidate's skills by the jury. Interestingly, participants with an assigned strong negative opinion of the candidate still preferred the lenient outcome when the jury's majority conclusion was to reject him/her and the majority opinion on the candidate's skills judged separately was positive. According to List (2006), in a disjunctive decision problem, the elemental deliberation style could be interpreted

as avoiding more false positives (e.g., accepting the candidate when he/she did not deserve it) than the compound style, whereas the compound style would avoid more false negatives (e.g., rejecting the candidate when he/she deserved to be hired). Bonnefon (2010) suggested that avoidance of false negatives is more of a welfare concern than avoidance of false positives, when the framing of the situation is positive as it was the case in the present protocol. The pervasive leniency effect in most of the present conditions could be interpreted under this light, that is, as avoiding a false negative (not hiring a proper candidate). In turn, participants were only ready to make their negative opinion count when a majority of negative opinions on the premises reassured that a false negative was unlikely to be reached. Following an argument by MacCoun and Kerr (1988: 31), if the evidence for a particular verdict is clear (as suggested by the converging negative positions about the facts of both the majority of the jury and the participant), then the favoured verdict should be demonstrably “correct” and the “bias” for that position should be evident. This suggests that the stronger and clearer the evidence against the candidate is, the less pronounced the general leniency bias should be, as we found in present experiments.

### Limitations of the present study

We can mention some limitations of the experiments described here which may merit further research to confirm present findings more generally.

First, we here relied on a convenient sample (university students), whereas the implications could be meant to apply, on the one hand, to the general public, and on the other hand, to group decisions, from courts, through academic panels, to political committees, etc. Therefore, it remains to be corroborated whether present experimental effects apply to more representative samples and/or experienced decision-makers.

Second, present experiments relied on hypothetical scenarios and decisions. Of course, this may limit the generalizability to real-world contexts in which decisions have real consequences and may put decision-makers’ reputations at stake. We believe this could be partially tackled with experimental economics methods in future studies (i.e., in experiments with monetary incentives).

Finally, a more specific methodological concern may arise when considering the present test of a confirmation bias. The procedure used here may seem conservative in this respect, since participants’ opinions were not self-generated, but were exogenously provided to them. Nonetheless, different arguments justify and back the present methodological choice. First, the DD arising spontaneously out of participants’ own opinions on a given topic would have been a very unlikely event; hence, it would have been impractical to rely on participants’

true opinions to study procedural preferences in the DD. Second, even if the present procedure looks artificial, the confirmation bias has been shown to be very pervasive, even appearing under conditions in which subjects had no material stake or obvious personal interest (Nickerson, 1998). And third, results indeed corroborated the effect of the exogenously provided opinions on participants’ decisions under some circumstances in the present experiments, meaning that the present manipulation was actually effective. Having defended this methodological choice, we could still see its limits. In particular, it could be argued that we might have just tested a lower bound of a confirmation bias effect, and that relying on participants’ true opinions could have revealed much stronger effects than the ones observed here.

Manuscript received January 2018

Revised manuscript accepted October 2018

### REFERENCES

- Bovens, L., & Rabinowicz, W. (2003). Democracy and argument—tracking truth in complex social decisions. In A. van Aaken, C. List, & C. Luetge (Eds.), *Deliberation and decision*. Aldershot, U.K.: Ashgate Publishing.
- Bovens, L., & Rabinowicz, W. (2006). Democratic answers to complex questions – An epistemic perspective. *Synthese*, 150(1), 131–153. <https://doi.org/10.1007/s11229-006-0005-1>
- Bonnefon, J.-F. (2007). How do individuals solve the doctrinal paradox in collective decisions? *Psychological Science*, 18(9), 753–755. <https://doi.org/10.1111/j.1467-9280.2007.01972.x>
- Bonnefon, J.-F. (2010). Behavioral evidence for framing effects in the resolution of the doctrinal paradox. *Social Choice and Welfare*, 34(4), 631–641. <https://doi.org/10.1007/s00355-009-0421-5>
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software.
- Brennan, G. (2001). Collective coherence? *International Review of Law and Economics*, 21(2), 197–211. [https://doi.org/10.1016/S0144-8188\(01\)00056-4](https://doi.org/10.1016/S0144-8188(01)00056-4)
- Cariani, F., Pauly, M., & Snyder, J. (2008). Decision framing in judgment aggregation. *Synthese*, 163(1), 1–24. <https://doi.org/10.1007/s11229-008-9306-x>
- Chapman, B. (2002). Rational aggregation. *Politics, Philosophy & Economics*, 1(3), 337–354. <https://doi.org/10.1177/1470594X02001003004>
- Dietrich, F., & List, C. (2007). Strategy-proof judgment aggregation. *Economics and Philosophy*, 23(03), 269–300. <https://doi.org/10.1017/S0266267107001496>
- Esaiasson, P., Persson, M., Gilljam, M., & Lindholm, T. (2016). Reconsidering the role of procedures for decision acceptance. *British Journal of Political Science*, 1–24. <https://doi.org/10.1017/S0007123416000508>
- Hartmann, S., & Sprenger, J. (2012). Judgment aggregation and the problem of tracking the truth. *Synthese*, 187(1), 209–221. <https://doi.org/10.1007/s11229-011-0031-5>



Hastie, R., Penrod, S., & Pennington, N. (1983). *Inside the jury*. Cambridge, MA: Harvard University Press.

Hulst, L., van den Bos, K., Akkermans, A., & Lind, E. (2017). On the psychology of perceived procedural justice: Experimental evidence that behavioral inhibition strengthens reactions to voice and no-voice procedures. *Frontiers in Psychological and Behavioral Science, 6*(1), 1–12.

Kameda, T. (1991). Procedural influence in small-group decision making: Deliberation style and assigned decision rule. *Journal of Personality and Social Psychology, 61*(2), 245–256. <https://doi.org/10.1037/0022-3514.61.2.245>

Kornhauser, L. A., & Sager, L. G. (1993). The one and the many: Adjudication in collegial courts. *California Law Review, 91*, 1–51.

Lind, E. A., & Tyler, T. R. (1988). *The social Psychology of procedural Justice*. New York, NY: Springer Science+Business Media, LLC.

List, C. (2006). The discursive dilemma and public reason. *Ethics, 116*(2), 362–402. <https://doi.org/10.1086/498466>

List, C., & Puppe, C. (2009). Judgment aggregation: A survey. In P. Anand, P. Pattanaik, & C. Puppe (Eds.), *The handbook of rational and social choice*. Oxford, U.K.: Oxford University Press. Chapter 19

MacCoun, R. J., & Kerr, N. L. (1988). Asymmetric influence in mock jury deliberation: Jurors' bias for leniency. *Journal of Personality and Social Psychology, 54*(1), 21–33. <https://doi.org/10.1037/0022-3514.54.1.21>

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175–220. <https://doi.org/10.1037/1089-2680.2.2.175>

Pettit, P. (2001). Deliberative democracy and the discursive dilemma. *Philosophical Issues, 11*(1), 268–299. <https://doi.org/10.1111/j.1758-2237.2001.tb00047.x>

Regenwetter, M., Grofman, B., Popova, A., Messner, W., Davis-Stober, C. P., & Cavagnaro, D. R. (2009). Behavioural social choice: A status report. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1518), 833–843. <https://doi.org/10.1098/rstb.2008.0259>

Thibaut, J., & Walker, L. (1975). *Procedural justice: A psychological analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Tyler, T. R. (2006). *Why people obey the law*. Princeton, NJ: Princeton University Press.

Tyler, T. R. (2011). *Why people Cooperate. The role of social motivations*. Princeton, NJ: Princeton University Press.

van den Bos, K. (2005). What is responsible for the fair process effect? In J. Greenberg & J. Colquitt (Eds.), *Handbook of organizational justice* (pp. 273–300). Mahwah, NJ: Lawrence Erlbaum Associations.

APPENDIX

TABLE A1  
Non-paradoxical conditions

Teaching	Research	Accept or not?
(a) Disjunctive version with a strong negative opinion provided to the participant and aggregation of conclusions leading to reject the candidate		
Judge 1 No	No	No
Judge 2 Yes	No	Yes
Participant No	No	No
Majority 2 out of 3 reject the teaching skills	2 out of 3 reject the research skills	Reject (no paradox)
(b) Conjunctive version with a strong negative opinion provided to the participant and aggregation of conclusions leading to reject the candidate		
Judge 1 No	Yes	No
Judge 2 Yes	No/yes	No/yes
Participant No	No	No
Majority 2 out of 3 reject the teaching skills	2 out of 3 reject/accept the research skills	Reject (no paradox)