Self-Signaling in Voting

Lydia Mechtenberg, Grischa Perino, Nicolas Treich, Jean-Robert Tyran, and Stephanie Wang[1]

February 28, 2022

Abstract

This paper presents a two-wave survey experiment on self-image concerns in voting. We elicit votes on a ballot initiative in Switzerland that spurred campaigns involving widely shared normative values. We investigate how messages that change the self-signaling value of a Yes vote affect selection and processing of information, and reported voting behavior. We find that a message enhancing the self-signaling value of a Yes vote is effective: voters agree more with arguments in favor of the initiative, anticipate more frequently voting in favor, and report more frequently having voted in favor of the initiative.

JEL: C93, D72, D91

Keywords: voting, multi-wave field experiment, information selection, information processing

[1] Mechtenberg (corresponding author): University of Hamburg, Lydia.Mechtenberg@uni-hamburg.de. Perino: University of Hamburg, Grischa.Perino@uni-hamburg.de. Treich: University Toulouse Capitole, INRAE, Toulouse School of Economics, nicolas.treich@inrae.fr, Tyran: University of Vienna, University of Copenhagen and CEPR (London), Jean-Robert.Tyran@univie.ac.at, Wang: University of Pittsburgh, swwang@pitt.edu. We thank Claudia Schwirplies for helpful comments. Nicolas Treich acknowledges support from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program), and IDEX-AMEP and FDIR chairs at the Toulouse School of Economics (TSE-P). Lydia Mechtenberg acknowledges support from the EU-Consortium DEMOS under grant agreement ID 822590 (Horizon 2020). This study was registered with the AER RCT Registry as #AEARCTR-0003551.

1 Introduction

How do people decide on their votes? One prominent motivation for voters is to act upon their normative beliefs (Feddersen et al., 2009; Morgan and Várdy, 2012).[2] Not surprisingly, value-based claims about policies, candidates, and/or parties are prevalent in political campaigns and discourse (Sandel, 2005; Haidt 2012; Enke 2020). After all, it is well-documented that people, at least to some extent, like perceiving themselves as being "good" and form self-serving judgments about the consequences of their actions (Bodner and Prelec, 2003; Bénabou and Tirole, 2011; Di Tella et al., 2015; Grossman and van der Weele, 2017). Playing the ethical card during political campaigns can hence mobilize voters to support a good cause. However, it may also mobilize them to support a policy that serves special interest since an ethical frame can trigger self-image concerns, which may bias voters' information selection or processing (Niehaus 2020). Self-image concerns may lead voters to neglect or downplay information that indicates that the suggested policy is not as ethical as proclaimed.[3] The reason is that believing in the ethical value of the proposed policy is necessary for self-signaling being a good person by voting in favor of the policy. Hence, campaigners may find it profitable to play the ethical card without going in-depth about why the suggested policy does indeed serve its proclaimed ethical purpose. In light of these dynamics, this paper addresses the question of how campaigning that targets self-image can affect voters' selection and processing of information about electoral issues and consequently their votes.

We conducted a baseline survey before our participants made their choices, implemented informational interventions in this survey, and elicited final decisions in a second survey afterwards.[4] In our case, the final decisions of interest are votes in an animal-welfare ballot in Switzerland (see Cantoni et al., 2019, who studied the effect of changing beliefs on self-reported participation in a political protest). With our treatments, we test if and how moralizing campaigns affect voters' information selection, information processing, and voting

[2] This study is agnostic about whether the utility derived from acting upon moral beliefs is expressive, or instrumental, or both. See the detailed discussion on how expressive and instrumental motives interrelate (and can well be aligned) in the introduction to Morgan and Várdy (2012). See also Borah (2019).

[3] Note that this differs from the trade-off between material selfishness and moral actions that is prominently studied in the literature. For experimental studies on this latter trade-off, see Golman et al. (2017) and the literature cited therein.

[4] Ethical approval has been granted by the dean of the social-science faculty at Hamburg University. The form can be obtained from the authors by request. The experiment was pre-registered at the AEA RCT Registry (https://doi.org/10.1257/rct.3551-1.0) under a different title.

behavior. In testing both whether moralizing campaigning affects information selection and/or information processing, we investigate the precise channels through which individuals, in particular voters, may bias their beliefs to sustain their self-image, as modeled in Bodner and Prelec (2003), Bénabou and Tirole (2011), and the literature that builds on them.

By choosing their information sources, voters can to some extent select information to sustain their self-image.[5] In addition, there is another possible informational strategy that they can use to manipulate their own beliefs: They can downplay information that contradicts the supposed ethical value of their choices, and they can overweigh information supporting this supposed ethical value (Golman et al. 2017). Such information-processing strategies of self-manipulation are as consistent with the prevalent models of motivated bias as pure information avoidance is. Similar to confirmation bias (Rabin and Schrag 1999), biased ethical beliefs resulting from biased information-processing strategies can survive all contrary campaigns that expose voters to fact-based information (Fryer et al., 2019). If people do indeed use biased information-processing to improve their self-image, campaigners will be heavily tempted to play the ethical card to trigger these strategies in voters and win them over. Their opponents will be at a loss of what to do against this – other than engaging in the same strategy and thereby escalating political polarization (Fryer et al., 2019; Garrett and Bankert, 2020).

The ballot in our study followed an initiative in Switzerland that claimed to improve animal welfare. This initiative demanded writing the dignity of horned animals into the Swiss constitution and to cross-subsidize farmers with horned cattle who refrain from dehorning. Both the campaigners for the initiative and their opponents used ethical arguments, the major pro-argument claiming that dehorning amounts to violence, and the major contra-argument claiming that the likely alternative, namely tethering that comes close to immobilizing the horned cattle, amounts to violence as well. This initiative provides the ideal setting for our study because of its strong ethical dimension, the consensus on the ethical goal (improving animal welfare), and the lack of consensus on whether the proposed policy (cross-subsidizing

[5] A sizeable literature starting with Dana et al. (2007), and Ehrich and Irwin (2005) provides lab evidence of participants avoiding costless information on whether a preferred, materially selfish action has pro- or anti-social consequences. Our approach differs from theirs in three respects. First, we are not interested in the trade-off between a materially selfish and a moral action, but in the trade-off between a seemingly moral and a truly moral action. Second, we are interested in how moral campaigning affects voting behavior. Third, we are focused on behavior in the field.

farmers who refrain from de-horning even if instead they tether their cattle) does indeed serve this goal.[6]

Our main treatments manipulate the self-signaling value of voting for the suggested policy.[7] We provide subjects with the truthful, if simplified, message that good-hearted people tend to be good to animals, too. This manipulation using a simple message follows from the psychology literature that has shown that self-image (and related concepts such as self-esteem, self-view or self-awareness) can be momentarily altered (Heatherton and Polivy 1991, Gao et al. 2009).[8] Note that our message is not informative about how animal-friendly or ethical the suggested policy is. Instead, this message contains information about a correlation between being good-hearted in general and an animal-friendly attitude. Hence, this message does not provide ethical orientation as to whether or not one should vote in favor of the initiative if one wants to make the ethical decision. Instead, our intervention increases the self-signaling value of voting for the policy. Treated subjects can use a Yes vote as a means to self-signal both being good to animals and being good in other respects, thus entailing a better self-image than subjects who we did not treat with our message. Importantly, our intervention does not provide additional information on the policy's effectiveness or conformity with ethical duties or rights. In another intervention that is designed to reduce the self-signaling value of a Yes vote, we tell our subjects that being good to animals does not necessarily imply being generally good-hearted. Both interventions are compared with the control in which no extra message is provided. We then study how our interventions affect the willingness to read arguments pro and contra the policy, agreement with these arguments, subsequent voting intentions, and reported voting behavior.

We find that increasing the self-signaling value of voting *Yes* through our informational intervention has significant effects: First, it enhances the intensity of the subjects' agreement with arguments stating that the policy at stake would indeed benefit the animals concerned, while the intensity of their agreement with arguments disputing this remains unchanged. (Surprisingly, however, our intervention does not affect the choice of which type of arguments

[6] See Osborne and Turner (2010) on the appropriateness of referenda for collective decisions with a dominant common-interest component.

[7] In this respect, our study is related to Schneider (2020) who, in a moral consumption context, varies the self-signaling value of particular product choices.

[8] Reporting a laboratory experiment, Falk (2021) shows that a manipulation of self-image can significantly change the incidence of immoral behavior. In contrast, our research interest is not on de-facto morality but in self-signaling.

to read, pro or contra the initiative.) Hence, we find that self-signaling is sustained by biased information processing in the field, at least in our sample. This evidence highlights one important channel through which "motivated bias" is generated and sustained in natural environments. Second, our intervention increases the number of subjects intending to vote *Yes* and, indeed, the number of subjects who report having actually voted *Yes* after the ballot. Decreasing the self-signaling value of voting *Yes* has no effect. This asymmetry is in line with other studies that find updating after self-serving news but no reactions to news that, if processed correctly, would undermine self-serving biases (Eil and Rao, 2011; Sunstein et al., 2016). To summarize, our findings are consistent with our subjects signaling to themselves a high value of their character through biased information processing. Gaining a positive self-image indeed seems to be an important motivation when choosing how to vote. This experimental result is consistent with a basic model of voting with motivated beliefs, as shown in the appendix B (see also Le Yaouanq 2021).

Voting is particularly vulnerable to self-signaling motives as the likelihood of an individual voter being pivotal is typically rather small (Shayo and Harel, 2012). Irrespective of the potential material outcomes of a ballot, the expected costs of voting one way or another tend to be negligible for an individual voter. Hence, even if the self-image concerns are relatively small, they may still exert a sizable influence on voting decisions[9]. Consistent with our finding, there exists ample evidence of biased information processing on politically contentious, value-laden issues such as death penalty (Lord et al. 1979), abortion (Pomerantz et al. 1995), attitude about homosexuality (Munro and Ditto 1997), justification for war (Nyhan and Reifler 2010) or climate change (McCright and Dunlap 2011). Given the importance of voting in determining the fate of democratic societies, understanding the role of self-signaling motives in this context is of particular interest.

2 Procedures and predictions

2.1 Data and descriptive statistics

On November 25, 2018, the Swiss voted on the proposal of a grass-root initiative colloquially called "Horncow Initiative". This initiative demanded to pin down the dignity of horned

[9] See Gerber et al. (2008) and DellaVigna et al. (2016) for evidence of the effect of social motives on voter turnout and behavior.

animals in the Swiss constitution. In addition, it asked for subsidizing farmers who do not cauterize their animals' horns. These subsidies, it requested, should be financed by cutting agricultural subsidies elsewhere and should hence be without effect on taxes or prices. We chose this ballot for its near-absence of substantial economic impact: The cost and consumption effects that the initiative's proposal would have on most voters in case of success would be negligible, and their self-interest would not be touched.[10] Hence, voters' instrumental concerns would be mainly altruistic, i.e., directed toward the proposal's true consequences on animal welfare. This provides an incentive to gain as objective, unbiased information about these consequences as possible. Moral self-signaling concerns, by contrast, make it attractive to remain ignorant about the proposal's potential negative consequences, or the potential absence of positive consequences. We conducted a pre-registered and IRB-approved two-waves survey experiment timed before and after the ballot. In the second wave, we re-contacted only subjects that completed the first wave. The first wave was implemented in the two weeks prior to the final day of the ballot, and the second wave a few days after. We restricted our experiment to the German-speaking part of Switzerland. The Swiss standing LINK Institute panel was employed for recruitment, and written consent was obtained from all subjects as part of wave 1.[11] Only truthful information was given to them. Subjects were informed as part of their consent that the survey in wave 1 might vary across participants. We screened out early voters who had voted already before the start date of wave 1 and participants not eligible to vote.

In the first wave, we conducted nine randomized versions of one survey. All versions elicited, among relevant demographics, (1) variables measuring information selection, (2) a variable measuring information processing, and (3) the intended vote. In addition, we elicited control variables such as the PriorAttitude toward the initiative's proposal and prior informedness (Informed). A full list of variables and their explanations is relegated to Table A.1 in the appendix.

Information selection. We measure information selection as follows. A booklet that the Swiss government sent to all Swiss voters several weeks before our experiment started contained three arguments in favor and three arguments against the initiative's proposal. We used these

[10] There are of course non-negligible cost effects on farmers. We elicit if our subjects are farmer or related to farmers and control for this.

[11] LINK institute cooperates with the Swiss government on a regular basis, implementing post-ballot polls, which are considered reliable.

and one other argument widely circulating in the media to create a balanced information menu: Three arguments in favor of the proposal claimed that dignity and physical well-being of animals and justice among farmers would improve, should the initiative be approved in the ballot. Three arguments against the proposal addressed these same three goals and argued that none of them would be reached in case of the proposal's success (See Table 1 for the precise formulation of the six arguments).[12] Our subjects had to choose which arguments to read: all six, only the three in favor, only the three against, or none at all. At the point of choice, they did not know that we were offering them only arguments they already were highly likely to know from the official booklet or the media. Thereby, without biasing their information set, we could measure the willingness of those predisposed in favor of the initiative to avoid negative information on the proposal's potential consequences, either by avoiding information in general or by reading only the supportive arguments. Similarly, we could measure the willingness of all voters to avoid reading arguments that contradict their prior attitude toward the initiative.

Table 1: Arguments of the endogenous information-acquisition mechanism.

Arguments for the Horncow Initiative	**Arguments Against the Horncow Initiative**
Dehorning violates the dignity of animals and is tantamount to a mutilation. It must mean something if nature gave horns to cows. For instance, these horns help the cows sorting out their hierarchy within their herd.	*It is well possible that the Horncow Initiative does not improve the dignity of animals. The reason is that in order to get subsidized, farmers could resolve to fixate their animals (e.g., by tethering). Their motive: Wounds caused by horns lower profits but may be prevented not only by dehorning but also by resolute fixation of the cattle, i.e., by limiting their range of motion to the greatest extent. Hence, farmers who nowadays dehorn their animals could, in case of the initiative's success, switch to permanent tethering of their cattle.*
Horns are organs well supplied with blood. Dehorning cows requires cauterizing the sockets of the horns to prevent them growing. This is a substantial medical intervention. Even though this intervention is legally required to be conducted under anaesthetization, many calves suffer from pain after cauterization, partly for long time.	**It is well possible that the Horncow Initiative does not prevent cruelty to animals. Resolute limiting of their range of motion in the stable or wounds caused by horns of other cows could result from subsidizing farmers with horned cattle. Possibly cows suffer more from tethering (or, alternatively, wounds caused by skirmishes with other horned cows in the stable) than from the dehorning.*
Since horned animals need more space and care from their farmers, a compensation for farmers holding horned animals is justified. Hence, farmers holding horned animals should be subsidized. Since the initiative does not demand a legally banning dehorning animals, the farmers' freedom of choice is preserved.	*Subsidizing farmers with horned animals may put those farmers at a disadvantage who breed hornless cattle. Even nowadays there are such farmers in Switzerland. There is no scientific evidence that cattle that is born hornless is "less natural" or suffers more than horned*

[12] Table 1 contains the English translation of the original German arguments.

	cattle. Hence, one should not put farmers who breed hornless cattle at a disadvantage.

Note: *This argument has been taken from the media. The Swiss booklet sent to all Swiss voters mentioned an argument almost identical to the second in the right column here.

Information processing. Even if acquired information was unbiased, it might be processed in a biased way by our subjects, as modeled in most of the work of Bénabou and Tirole (e.g., 2002, 2006, 2011) and confirmed in lab experiments both by economists (Eil & Rao, 2011) and neuroscientists (Sharot et al., 2011; Sharot & Garrett, 2016; Kuzmanovic et al., 2018). As memories need to be reconsolidated after access (Lee et al., 2017) treatments can introduce long-run biases even to information that was acquired before the start of the experiment if participants access these memories during the experiment (Yao et al., 2021). In our survey, we therefore asked for each type of argument, supportive or unsupportive of the proposal, how much the subjects who read it agreed with it ranging from 'not at all' to 'fully'. This allows us to compute a measure of change relating the PriorAttitude, i.e. the degree to which participants reported to be leaning in favor or against the initiative before being exposed to treatment and information, to the degree to which they agree with pro or contra arguments post treatment. To this end we normalize all variables and take the difference between ex-post and ex-ante variables normalizing the result to the interval [-1,1]. The resulting change measures are ΔAgreementPRO and ΔAgreementCON. Moral self-signalers have to believe in the moral value of a Yes vote to use it as a profitable signal to themselves. Hence, the more they want to self-signal, the more they have an incentive to process arguments in a biased way, assigning more weight to the supportive type.

Intended and actual votes. Voting plans tend to function as commitment devices (Nickerson and Rogers 2010). We hence elicit the immediate effect of our treatment variations explained below on planned voting behavior by asking our subjects whether they intend to turn out and, if they do, how they intend to vote. In combination with the variable PriorAttitude, this allows us to measure changes in how subjects evaluate the proposal after being treated in the experiment. After the ballot, we re-contacted all subjects who completed the first wave and elicited their actual vote by asking whether, and how, they voted.

Treatments. Our treatments are depicted in Figure 1. First, we vary the self-signaling value of a Yes vote: HIGH and LOW treatments differ from NEUTRAL treatments in that in both former types, we give subjects true information that we expect will enhance (in HIGH) or lower (in

LOW) the salient moral value of a Yes vote. In HIGH, we cite evidence for the positive correlation between cruelty towards animals and cruelty towards humans and conclude that good-hearted people tend to be good to animals. In LOW, we cite evidence indicating that the correlation between empathy with animals in need and empathy with humans in need is less than perfect. Note that neither information touches the question whether the success of the initiative would improve the animals' situation. (See Appendix A.I for the precise wording and the scientific foundation of our interventions.) These informational interventions are implemented after we elicit prior attitudes and prior informedness but before subjects have to make their choice of arguments to read. In the NEUTRAL treatments, we refrain from any such informational intervention.

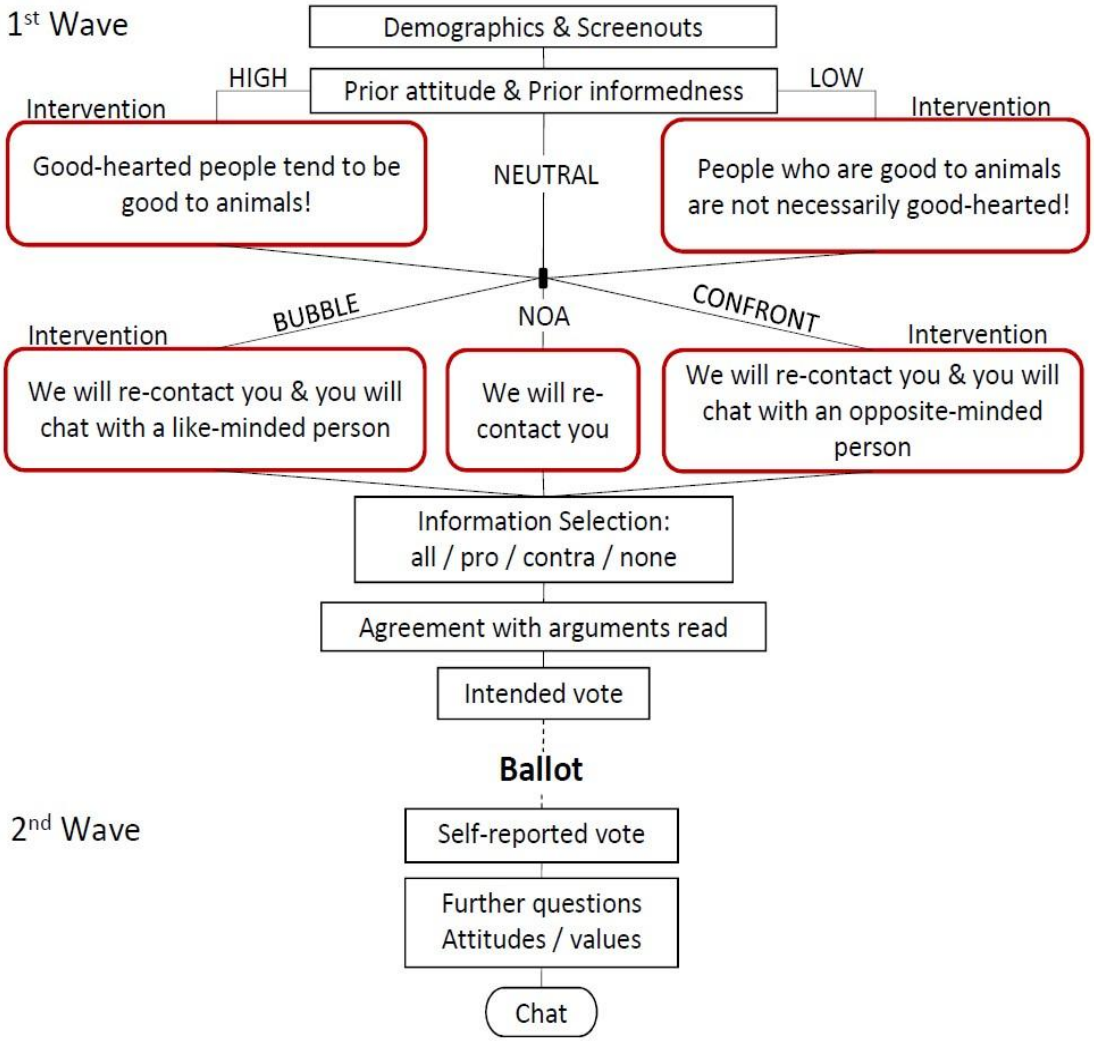
Second, and orthogonal to this variation, the treatments BUBBLE, CONFRONT and NOA vary what and how much subjects were told about communication in the second wave: in BUBBLE (CONFRONT), we told them that they would chat with someone of similar (different) prior attitude toward the Horncow Initiative after the second survey. In NOA (for "no anticipation"), we only told them that they would be re-invited for a second survey. Hence, BUBBLE and CONFRONT but not NOA induce anticipation of a social situation in which subjects may discuss their votes. While BUBBLE lets subjects expect social insulation of their prior opinion, CONFRONT promises confrontation with a different opinion.

The second wave of our experiment included a short survey identical for all subjects and a partner-chat in which, depending on whether the subject was in BUBBLE or CONFRONT, they chatted with an (ex ante) like-minded or opposite-minded partner. (Participants in NOA were randomly assigned to either a like-minded or an opposite-minded partner.) We re-invited all 2,112 subjects who completed the first survey and had 1,057 completing the second. Apart from eliciting self-reported actual votes, the second survey contained questions on whether consequences (GoodEffect) or intentions (GoodIntent) are more important when morally evaluating – and rewarding – a particular action.

Two pathways. Before we state our predictions, we clarify one important distinction that distinguishes two pathways toward voting in order to self-signal a good character in our experiment: the distinction between information selection and biased information processing. While we can measure information avoidance by whether subjects neglected arguments, we can measure biased information processing by the agreement with arguments in favor or against the initiative only for those who read both types of arguments. Hence, if we find

sizeable information avoidance, we will not be able by design to find biased information processing, because biased information selection implies that the sample of those reading both types of arguments is biased. Treatment effects on agreement with arguments in favor or against the initiative could not be reliably attributed to biased information processing in this case.

Figure 1: Treatments



Apart from this technicality, there is another design feature that marks the two types of informational bias as openings into two different pathways of behavior.[13] To see this, consider first a subject supportive of the initiative who ponders which arguments to read: all, or only those in favor, or only those against, or none. Our treatment interventions are designed to influence her choice. However, since we composed the set of arguments based on the booklet

[13] In principle, subjects could use both pathways simultaneously.

sent to all Swiss voters by the government, with the exception of one argument that was prominent in the media, the subject's information set should not depend on which reading choice she would finally make. Therefore, by design subjects who exhibit biased information selection in our experiment should not be driven by that to biased voting behavior, be it planned or actual. Hence, biased information selection in our experiment should not lead in itself to biased voting, compared to how our subjects would have voted without our treatment interventions.

However, it is easy to see that the situation is different when considering biased information processing. Here, we were unable to preclude by design that the bias on the informational stage translates into biased voting. A subject reading both types of arguments but influenced by our treatment intervention HIGH (LOW) to put more (less) weight on those in favor of the initiative may well become more (less) likely to vote *Yes*. Hence, we get the two potential pathways *A (*for *Avoid)* and *B (*for *Bias)* below.

Pathway A: In treatment HIGH (LOW), subjects become weakly more (less) likely to skip arguments against the initiative, compared to NEUTRAL, respectively. But they remain unaffected in their planned and actual voting behavior.

Pathway B: In treatment HIGH (LOW), subjects become weakly more (less) likely to overweigh arguments in favor of the initiative, relative to those against, than in NEUTRAL, respectively.

Both pathways, A and B, are rooted in the theoretical literature discussed above, in particular in the work of Bénabou and Tirole. While the existing literature on information avoidance has already documented subjects using pathway A (Matthey and Regner, 2011; Nyborg, 2011; Feiler, 2014; Grossman, 2014; Serra-Garcia and Szech, 2019; Freddi, 2019), pathway B still lacks empirical evidence.[14]

2.2 Predictions

Below, we state the predictions for both pathways. If pathway A is clearly refuted, we will get a large enough subsample of subjects who read all arguments, which allows us to test pathway B if that sample turns out to be unbiased. Note that biased processing of arguments, e.g., overweighting pro arguments compared to con arguments, implies a higher probability of biased voting intentions and behavior, while biased argument selection does not have such a

[14] In the pre-registration of this study, we only mentioned pathway A.

clear-cut implication in our setting. The reason is that we composed the set of arguments from the booklet that the Swiss government sent to all voters and one argument circulating in the media. Hence, a participant selecting arguments may not encounter unknown information. However, only biased, but not Bayesian, processing of information that is already known could lead to a change in intended and actual behavior.

Pathway A. We now state all hypotheses relating to pathway A.

Hypothesis H1.A (Self-Image and Information Selection).

(a) HIGH increases direct avoidance of arguments against the Horncow Initiative, compared to NEUTRAL for those not initially opposing the Initiative.

(b) LOW decreases direct avoidance of arguments against the Horncow Initiative, compared to NEUTRAL for those not initially opposing the Initiative.

Hypothesis H2.A (Social Dimension and Information Avoidance).

(a) BUBBLE increases direct avoidance of arguments opposing the participant's own prior attitude, compared to NOA.

(b) CONFRONT decreases direct avoidance of arguments opposing the participant's own prior attitude, compared to NOA.

Pathway B. We now state all hypotheses relating to pathway B.

Hypothesis H1.B (Self-Image and Information Processing).

(a) HIGH increases the agreement with arguments supportive of the Horncow Initiative, compared to NEUTRAL for those who have read both types of arguments.

(b) LOW decreases the agreement with arguments supportive of the Horncow Initiative, compared to NEUTRAL for those who have read both types of arguments.

Conditional Hypothesis H2.B (Self-Image and Votes).

(a) If H1.B (a) is true, then HIGH increases the likelihood of (i) intended and (ii) reported *Yes* votes.

(b) If H1.B (b) is true, then LOW decreases the likelihood of (i) intended and (ii) reported *Yes* votes.

We now proceed to testing these hypotheses. We correct for multiple-hypotheses testing using the Romano-Wolf correction.

3 Results

3.1 Data and descriptive statistics

The first wave of the survey was completed by 2,112 participants in German-speaking parts of Switzerland recruited from the standing LINK Institute panel that is representative for the Swiss adult population. The second wave was completed by 1,057 participants. Summary statistics of both waves can be found in Table 2. Attrition was not random but independent of treatment assignment. Most importantly, assignment to the HIGH treatment did not affect the likelihood of dropping out (Mann-Whitney, p = 0.988). Among those completing wave 1 but not wave 2 there were significantly more (Mann-Whitney, p = 0.0000) women and subjects were less well informed, more emotional about and more inclined toward supporting the initiative, compared to subjects completing both waves. Despite the overrepresentation of women among those dropping out, the share of women in the final sample is still above the national average (54.0 percent in the sample vs. 50.4 percent in the population) and treatment assignment did not affect the share of women among the dropouts. The share of participants that supported (opposed) the initiative in the final sample are comparable to those that participated in the ballot (see Table A.2). With respect to farmers in general, farmers with horned animals, and age, there was no significant difference between the samples. Unless stated otherwise, we report results for the subsample that completed both waves of the survey. For all outcome variables elicited in wave 1, we also report results for all participants completing wave 1 to check whether sample attrition is a relevant driver.

Table 2: Summary statistics of survey waves 1 and 2

	Wave 1			Wave 2		
Variable	**Obs.**	**Mean**	**std. dev.**	**Obs.**	**Mean**	**std. dev.**
HIGH	2,112	.340	.474	1,057	.341	.474
LOW	2,112	.315	.465	1,057	.305	.460
BUBBLE	2,112	.244	.430	1,057	.242	.429
CONFRONT	2,112	.232	.423	1,057	.237	.426
Female	2,108	.592	.492	1,054	.540	.499
age (categories)	2,112	3.836	1.623	1,057	3.855	1.666
Farmer	2,109	.0123	.110	1,056	.0123	.110
FarmHorn	2,112	.00473	.0687	1,057	.00473	.0686
Informed[#]	2,098	.0686	1.830	1,056	.3570	1.627
PriorAttitude	1,825	4.023	2.062	1,057	3.741	1.967

Emotions#	1,964	.1996	1.884	1,031	-.1077	1.698

Note: Variables tagged with a # are measured on an integer scale in the range {-3,3}.

Table 3: Information selection across treatments

	HIGH	LOW	NEUTRAL	BUBBLE		CONFRONT	NOA	ALL
Arguments read	**Percentages**							**Frequ.**
Both	77.50	79.50	79.73	82.81	79.28	76.91	78.90	834
None	16.67	14.29	13.60	12.50	13.94	16.36	14.85	157
Only PRO	4.17	4.04	4.27	2.73	5.58	4.18	4.16	44
Only CONTRA	1.67	2.17	2.40	1.95	1.20	2.55	2.08	22
Opposing	78.3	81.7	81.9	84.0	80.5	79.1	80.6	852
Frequ.	360	322	375	256	251	550		1,057

Note: 'Opposing' refers to the set of arguments that support the opposite position towards the initiative compared to that expressed by the participant prior to exposure to treatments. The first four rows are mutually exclusive and exhaustive, i.e. add up to the full sample completing both waves of the survey. The fifth row overlaps with rows 1, 3 and 4.

3.2 Pathway A: information selection

Strategic avoidance of information would be most obvious if treatments induced one-sided information selection. Hypothesis 1.A implies that the share of those only reading the PRO arguments should be higher (lower) in HIGH (LOW) than in NEUTRAL. Hypothesis 2.A implies that the share of those reading both arguments should be lower (higher) in BUBBLE (CONFRONT). Table 3 indicates that neither is the case. Mann-Whitney tests show that the distributions are not significantly different across treatments (Table 4). Logit regressions in Tables A.3 and A.4 confirm this. Hence, there is no clear treatment effect on direct information avoidance, i.e., Hypotheses 1.A and 2.A are not confirmed.

Table 4: Information Avoidance Non-Parametric Tests

Outcome Variable	**Treatments**			**Mann-Whitney** (*p*-values)	**N**
AvoidanceCONTRA	HIGH	vs.	NEUTRAL	0.309	735
	HIGH	vs.	NEUTRAL & LOW	0.280	1,057
	LOW	vs.	NEUTRAL	0.876	697
	LOW	vs.	NEUTRAL & HIGH	0.704	1,057
ReadOpposingAttitute	BUBBLE	vs.	NOA	0.102	806
	BUBBLE	vs.	NOA & CONFRONT	0.116	1,057
	CONFRONT	vs.	NOA	0.652	801
	CONFRONT	vs.	NOA & BUBBLE	0.953	1,057

Note: While the impact of BUBBLE is close to being significant at the 10%-level, the direction of the impact is the opposite of that conjectured in Hypothesis H2.A a).

In sum, testing Hypotheses 1.A and 2.A does not provide any evidence for pathway A: we do not find that voters use more information avoidance to sustain a moral self-image when salience of morality increases (HIGH) or less information avoidance when salience of morality decreases (LOW). Neither do we find that they avoid more information in a harmonious social setting and less in a confrontational setting (BUBBLE / CONFRONT) than under social insulation (NOA). Hence, either moral self-signaling and the degree of social insulation are irrelevant motivations in our voting context, or some or all of these motivations work through pathway B rather than pathway A, i.e., via biased information processing rather than biased information selection. We hence turn to testing pathway B.

3.3 Pathway B: information processing

Biased information processing is observable only for those arguments participants chose to read. The subsample of participants that read all arguments is therefore our starting point in this part of the analysis. Since we find treatment effects on the perception of PRO arguments, we check robustness of results for the sample of participants that selected the PRO but not necessarily the CONTRA arguments as well. Note that only about two percent of participants exclusively read the CONTRA arguments (Table 3).

The decision to read PRO or both PRO and CONTRA arguments is not affected by the four interventions HIGH/LOW and BUBBLE/ANT (Table 5). Only BUBBLE comes close to having a significant impact on sample composition. Hence, should BUBBLE turn out to be a significant driver of biased information processing or voting, then we would need to treat that result with caution.

Table 5: Test for sample selection (Logit)

	(1)		(2)	
	Read PRO & CON		Read PRO	
HIGH	-0.019	(0.523)	-0.012	(0.649)
LOW	-0.007	(0.823)	-0.003	(0.905)
BUBBLE	0.060	(0.059)	0.048	(0.098)
CONFRONT	0.021	(0.492)	0.041	(0.160)
Informed	-0.015	(0.083)	-0.017	(0.039)
PriorAttitude	0.007	(0.251)	0.016	(0.005)
Farmer	0.114	(0.516)	0.080	(0.586)
FarmerHorn	-0.219	(0.342)	-0.037	(0.861)
Female	-0.004	(0.875)	-0.013	(0.576)
Age categ.	*Yes*		*Yes*	

N	1049	1049

Note: Dependent variables: Dummies if PRO & CON or PRO arguments have been read, respectively. Marginal effects are presented. p-values in parentheses unadjusted for multiple hypothesis testing.

Testing the hypotheses for pathway B (H1.B – H2.B) requires some care in choosing the identification strategy. *PriorAttitude*, the variable capturing a participant's attitude towards the initiative before any of the treatment interventions took place, is highly correlated with both post-treatment agreement with PRO arguments (Pearson's r = -0.6104, p = 0.0000) and with anticipated (Pearson's r = -0.7374, p = 0.0000) and reported voting (Pearson's r = 0.7235, p = 0.0000). This is not surprising, as exposure to the survey and treatment interventions are unlikely to fundamentally uncouple a participant's preferences and voting behavior from her position at the beginning of the survey. While *PriorAttitude* has immense explanatory power for the outcome variables of interest in pathway B, it also is arguably correlated with the error term as it is highly likely that it is causally affected by unobserved variables that also causally affect the outcome variables of interest. In the same direction points that both coefficients and significance levels of treatment dummies are highly sensitive to the inclusion of *PriorAttitude* as a control variable and that t-values for *PriorAttitude* are an order of magnitude higher than those of other explanatory variables (see Table A.5 in the appendix). Hence, estimated coefficients in regressions that include *PriorAttitude* as a control variable are likely to suffer from an omitted variable bias.

To address this problem while still using the highly relevant information contained in *PriorAttitude*, we apply a diff-in-diff approach where *PriorAttitude* serves as the reference point. Because both *PriorAttitude* and the outcome variables *AgreementPRO*, *IntendedVote* and *ReportedVote* are measured in different but intuitively compatible categorical scales we normalize each of them to the interval (-1,1) and (0,1), respectively, before taking differences. The latter are again normalized to the interval (-1,1). This has the added advantage that the distributions of the resulting variables *ΔAgreementPRO*, *ΔIntendedVote* and *ΔReportedVote* are now close to continuous and we therefore use OLS instead of (ordered) logit regressions which eases both interpretation of coefficients and multiple hypothesis testing. In total we test ten hypothesis in this paper taking into account that we use two outcome variables, *ΔIntendedVote* and *ΔReportedVote*, for hypotheses H2.B(a) and H2.B(b). We use the Romano-Wolf (Romano and Wolf 2005a,b, 2016 and Clarke et al. 2020) procedure based on 10,000 replications to calculate p-values adjusted for twenty hypotheses (five outcome variables

times four treatment variables) and eight hypotheses (outcome variables *AvoidanceCONTRA, ΔIntendedVote, ΔReportedVote, ΔAgreementPRO* for treatments HIGH and LOW). Due to the restrictions imposed by the implementation in STATA, we could not test the exact set of hypotheses and we therefore went for a very conservative (twenty hypotheses) and a somewhat laxer (eight hypotheses) version. The latter was necessary for those specifications that did not include the BUBBLE and CONFRONT treatments as explanatory variables. As the implementation in STATA does not allow us to specify outcome-variable specific estimation methods we use OLS for all. For brevity, we only report adjusted *p*-values in those instances where the unadjusted *p*-values point towards a significant effect.

Table 6: Information Processing Non-Parametric Tests

Outcome Variable	**Treatments**			**Mann-Whitney**	***N***
ΔAgreementPRO				(*p*-values)	
Sample					
Read PRO & CON / Wave 1 & 2	HIGH	vs.	NEUTRAL	0.010	573
Read PRO & CON / Wave 1	HIGH	vs.	NEUTRAL	0.005	961
Read PRO / Wave 1 & 2	HIGH	vs.	NEUTRAL	0.026	604
Read PRO & CON / Wave 1 & 2	LOW	vs.	NEUTRAL	0.676	548
Read PRO & CON / Wave 1	LOW	vs.	NEUTRAL	0.808	924
Read PRO / Wave 1 & 2	LOW	vs.	NEUTRAL	0.813	577
ΔAgreementCON					
Read PRO & CON / Wave 1 & 2	HIGH	vs.	NEUTRAL	0.242	572

Notes: Mann-Whitney tests on impact of HIGH treatment on agreement with PRO and CON arguments for different subsamples. 'Read PRO & CON / Wave 1 & 2': read all arguments offered and completed waves 1 & 2; 'Read PRO & CON / Wave 1': read all arguments offered and completed wave 1; 'Read PRO / Wave 1 & 2': read only PRO arguments and completed waves 1 & 2.

Non-parametric tests of the impact of treatments on the agreement with the arguments in favor of the initiative are reported in Table 6. Irrespective of the sample, HIGH significantly increases the agreement with arguments in favor of the initiative relative to participants' attitude before the intervention. This is in line with hypothesis H1.B(a). There are no significant effects on agreement with arguments opposing the initiative nor of the LOW treatment on any of the outcome variables. Hypotheses H1.B(b) is not confirmed.

Table 7 uses regression analysis to confirm the above findings controlling for different sets of exogenous variables that were all elicited before treatment interventions and multiple hypothesis testing. Regressions (1) and (2) show for the sample of participants that read both

PRO and CONTRA arguments that HIGH significantly increases agreement with the PRO arguments. To test for robustness and reduce the risk of issues with sample selection, we repeat the analysis with two further samples. Regression (3) extends the sample to those participants only completed wave 1 of the survey adding another 565 observations. Regressions (4) and (5) cover all participants completing both waves that read the PRO arguments, i.e. compared to the sample of regressions (1) and (2) they include the 44 participants that did not read the CONTRA arguments (see Table 3).

Table 7: Biased information processing

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			ΔAgreementPRO				*ΔAgreementCON*
	Read PRO & CON Wave 1 & 2		PRO&CON Wave 1	Read PRO Wave 1 & 2		Read PRO & CON Wave 1 & 2	
HIGH	0.061	0.056	0.055	0.054	0.051	-0.016	-0.017
	(0.005)	(0.010)	(0.002)	(0.011)	(0.016)	(0.513)	(0.492)
Romano-*Wolf*	(0.012)	(0.035)		(0.014)	(0.035)		
LOW	-0.002	-0.005	0.004	0.001	-0.001	0.016	0.015
	(0.914)	(0.816)	(0.813)	(0.979)	(0.979)	(0.524)	(0.553)
BUBBLE		-0.006	-0.003		-0.004		-0.012
		(0.773)	(0.855)		(0.866)		(0.633)
CONFRONT		-0.026	-0.025		-0.023		0.008
		(0.246)	(0.161)		(0.302)		(0.740)
Informed		0.013	0.008		0.009		0.008
		(0.049)	(0.117)		(0.173)		(0.259)
Farmer		0.070	0.025		0.072		-0.097
		(0.476)	(0.734)		(0.464)		(0.382)
FarmHorn		0.230	0.041		0.206		-0.006
		(0.199)	(0.730)		(0.206)		(0.979)
Female		0.026	0.000		0.022		0.035
		(0.161)	(0.991)		(0.213)		(0.089)
Age categ.	*No*	*Yes*	*Yes*	*No*	*Yes*	*No*	*Yes*
_cons	0.119	0.121	0.119	0.120	0.119	0.007	0.001
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.688)	(0.980)
N	825	824	1389	869	867	823	822
R^2	0.013	0.049	0.024	0.009	0.040	0.002	0.019
F	5.277	2.762	2.242	4.090	2.352	0.788	1.051
aic	120.748	116.900	260.183	126.182	126.648	295.024	307.259
bic	134. 9	192.3	344.0	140.5	202. 9	309.2	382.6

Notes: OLS regressions
Dependent variables: *ΔAgreementPRO* and *ΔAgreementCON* capture the difference between reported convincingness of PRO/CONTRA arguments post treatment and reported prior attitude (before treatment). Both variables are normalized to values between [-1,1] with positive numbers indicating a shift in attitude toward the respective set of arguments.

p-values in parentheses unadjusted for multiple hypothesis testing unless specified otherwise; *Romano-Wolf p*-values in (2) and (5) corrected for 20 hypotheses (outcome variables *AvoidanceCONTRA, ReadOpposingAttitute, ΔIntendedVote, ΔReportedVote, ΔAgreementPRO*) for treatments *HIGH, LOW, BUBBLE* and *CONFRONT*) and in (1) and (4) corrected for 8 hypotheses (outcome variables *AvoidanceCONTRA, ΔIntendedVote, ΔReportedVote, ΔAgreementPRO*) for treatments *HIGH and LOW,* each based on 10,000 replications.

The treatment effect of HIGH is robust to both variations in the sample and to the correction for multiple hypothesis testing. Agreement with arguments against the initiative, however, is not affected by any of the treatments (regressions (6) and (7) in Table 7). In sum, we do find robust evidence for the use of biased information processing when morality becomes more salient, as in HIGH. If moral self-signaling is behind this bias, then HIGH will affect intended and reported voting behavior, too, as we hypothesized for pathway B. We hence now turn to investigating intended and reported votes.

3.4 Pathway B: voting behavior

Intended voting relative to participants' attitude towards the initiative before the intervention (*ΔIntendedVote*) exhibits the same pattern as the agreement with PRO arguments. The impact of HIGH on *ΔReportedVote* is only significant at the 10%-level. This gives some initial support of Hypothesis H2.B (a).

Table 8: Information Processing Non-Parametric Tests

Outcome Variable	**Treatments**			**Mann-Whitney**	**N**
ΔIntendedVote				(*p*-values)	
Sample					
Waves 1 & 2	HIGH	vs.	NEUTRAL	0.001	712
Wave 1	HIGH	vs.	NEUTRAL	0.025	1206
Waves 1 & 2	LOW	vs.	NEUTRAL	0.670	674
Wave 1	LOW	vs.	NEUTRAL	0.362	1144
ΔReportedVote					
	HIGH	vs.	NEUTRAL	0.064	529
	LOW	vs.	NEUTRAL	0.423	516

Figure 2 splits the share of reported *YES* votes by treatment and three categories of *PriorAttitude*. In all categories, the share of *YES* votes is higher in the HIGH treatment than in NEUTRAL but the difference is statistically significant (5%-level) only among those initially opposed to the initiative. The latter group contains more than half of all participants in the HIGH and NEUTRAL treatments that reported their vote in wave 2 of the survey. Note that the attitude categories in Figure 2 are based on measurements prior to exposure to treatments

and that the HIGH treatment induced a bias in favor of the PRO arguments relative to participants' prior attitude. Hence, the stronger impact on votes for those initially opposing the initiative is not surprising for two reasons. First, identifying a PRO-bias in a group that already reports agreement ex-ante is much harder than in a group that initially is more skeptical. Second, with a constant but small share of additional *YES* votes induced by the treatment, they are much more likely to occur and found significant, the larger the number of pre-treatment *NO* voters.[15]

Figure 2: Share of reported *YES* votes by categories of prior attitude

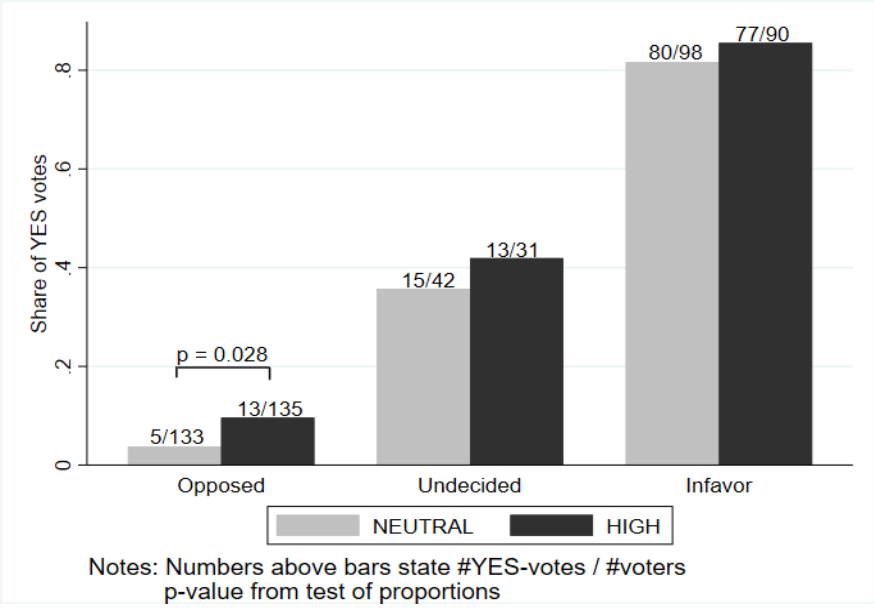


Table 9 presents additional results of OLS regressions on *ΔIntendedVote* and *ΔReportedVoting*. The sample of regressions (1) to (3) contains all participants that completed both waves of the survey whereas regression (4) also includes those that did not complete wave 2 of the survey which increases the sample size by 718 participants relative to (2). The sample in regressions (5) and (6) is smaller as it includes only those who have completed both waves of the survey and reported to have participated in the ballot. All specifications show a significant impact of HIGH on intended and reported voting. Significance is robust to multiple hypothesis testing. In sum, we clearly find evidence for moral self-signaling along pathway B: Exposure to the intervention that raised the salience of moral self-signaling by voting in favor of the initiative, while having no impact on information selection, did increase agreement with PRO arguments

[15] For the sample presented in Figure 2, the difference in the share of *NO* votes between the NEUTRAL and the HIGH treatment is: 5.87 percent for those initially opposed, 6.23 percent for those initially undecided and 3.93 percent for those initially in favor.

and intended and reported actual voting in favor of the Horncow Initiative. Overall, this provides strong support for the modelling paradigm developed by Bénabou and Tirole (2002, 2006, 2011) and the entire literature building on them. In the present context, the preferred strategy for keeping up a positive moral self-image is not to entirely avoid information that would undermine moral self-signaling strategies but to assign higher weights to information that helps rationalizing such strategies.

In order to validate this interpretation, we test a potential alternative explanation for the effects of HIGH. This explanation hypothesizes that those exposed to the HIGH treatment regard the instigators of the Horncow Initiative as being driven by good intentions and try to reward them by voting in favor of the initiative. The variables used to test this idea are *GoodIntent* measuring the degree to which participants agree with the claim that good intentions rather than good consequences of actions should be rewarded and its interaction with HIGH (*HIGHxGoodIntent*). We do not find evidence for the alternative explanation of our results (see (3) in Table 8). Without controlling for *PriorAttitute* or any other control variables, the share of *Yes* votes in the HIGH treatment is 40.2 percent relative to 36.6 percent in the NEUTRAL treatment. If only 36.6 percent of the 256 participants (i.e. 94 participants) exposed to the HIGH treatment that reported their voting decision had voted in favor of the initiative, then we would have seen nine fewer *Yes* votes. If we assume the same impact for all 719 participants exposed to the HIGH treatment, i.e. including those that did not complete wave 2 of the survey or did not report their vote, then the number of *Yes* votes has increased by 26 due to the experimental intervention. Using the coefficient from regression (6) in Table 9, i.e. 0.07 percent, the number of *Yes* votes increased by 18 in the sample reporting their vote and by 50 in the full sample exposed to the HIGH treatment. For comparison, in the ballot the number of *No* votes exceeded the number of *Yes* votes by 239,182 out of 2.6 million votes cast.

There is a potential self-selection issue in the information processing analysis. Participants self-select their exposure to information and, by design, we only observe the *AgreementPRO* and *AgreementCON* variables for those that have chosen to read them. However, for both intended and reported voting we observe outcomes irrespective of information selection. We test (see Table A.6) whether intended and reported voting are affected differently by the HIGH treatment for those reading both PRO and CON arguments (ALL) and those reading neither (NONE) and whether the voting intentions and reports differ between ALL and NONE.

The HIGH treatment is (at least weakly) significant in the combined and the ALL samples but not in NONE. Whether participants expose themselves to information has no impact on voting intentions and reports ($p > 0.87$). Hence, the observed effect is indeed driven by treatment assignment and not by information selection. Further explorative correlations are presented in Appendix C.

Table 9: Impact on intended/reported voting (OLS)

	(1)	(2)	(3)	(4)	(5)	(6)
	ΔIntendedVote			*ΔInt.V./ wave 1*	*ΔReportedVote*	
HIGH	0.041	0.046	0.044	0.025	0.059	0.070
	(0.003)	(0.001)	(0.003)	(0.024)	(0.046)	(0.019)
Romano-Wolf	(0.012)	(0.005)			(0.092)	(0.046)
LOW	-0.004	-0.002	-0.002	0.003	0.018	0.031
	(0.783)	(0.888)	(0.885)	(0.777)	(0.555)	(0.312)
BUBBLE		0.023	0.022	0.013		0.015
		(0.114)	(0.125)	(0.245)		(0.616)
Romano-Wolf		(0.386)				(0.856)
CONFRONT		0.003	0.002	-0.015		0.025
		(0.858)	(0.873)	(0.205)		(0.420)
Informed		-0.009	-0.009	-0.010		0.013
		(0.030)	(0.041)	(0.001)		(0.138)
Farmer		0.068	0.069	0.037		0.146
		(0.339)	(0.335)	(0.491)		(0.294)
FarmHorn		0.096	0.095	0.060		0.255
		(0.385)	(0.387)	(0.460)		(0.291)
Female		0.009	0.011	-0.003		0.002
		(0.438)	(0.380)	(0.776)		(0.949)
GoodIntent			0.006			
			(0.157)			
HIGHxGoodIntent			-0.005			
			(0.538)			
Age categ.	*No*	*Yes*	*Yes*	*Yes*	*No*	*Yes*
_cons	-0.001	-0.034	-0.032	-0.011	-0.161	-0.241
	(0.907)	(0. 150)	(0.189)	(0.554)	(0.000)	(0.000)
N	1021	1017	985	1735	772	768
R^2	0.012	0.033	0.034	0.017	0.005	0.033
F	6.195	2.271	2.02	1.993	2.076	1.685
Aic	-525.4	-515.3	-489.2	-783.2	528.8	531.1
Bic	-510.6	-436.5	-401.1	-695.8	542.8	605.4

Notes: Dependent variables: *ΔIntendedVote* and *ΔReportedVote* capture the difference between reported planned/actual vote and reported prior attitude (before treatment). Both variables are normalized to values between [-1,1] with positive numbers indicating a shift in attitude toward support of the horncow initiative. *p*-values in parentheses unadjusted for multiple hypothesis testing unless specified otherwise; *Romano-Wolf p*-

values in (2) and (6) corrected for 20 hypotheses (outcome variables *AvoidanceCONTRA, ReadOpposingAttitude, ΔIntendedVote, ΔReportedVote, ΔAgreementPRO*) for treatments *HIGH, LOW, BUBBLE* and *CONFRONT*) and in (1) and (5) corrected for 8 hypotheses (outcome variables *AvoidanceCONTRA, ΔIntendedVote, ΔReportedVote, ΔAgreementPRO*) for treatments *HIGH and LOW* each based on 10,000 replications.

4 Concluding remarks

We find experimental evidence of self-image concerns motivating voting behavior in a controversial Swiss ballot. This ballot had been initiated by farmers campaigning for more subsidies with the proclaimed purpose of increasing animal welfare. Participants exposed to scientific evidence establishing a correlation between kindness towards animals and kindness towards fellow humans assigned significantly more importance to the arguments that supported the initiative than those not exposed to such evidence. They were also more likely to vote in favor of the initiative. We do not find a treatment effect on participants' selection of which arguments to read (those in favor of the initiative, those against, both kinds, or the empty set). Thus, an increase in the self-signaling value of voting in favor of the initiative did not affect information selection prior to voting but did bias information processing instead. This identifies a precise channel through which individuals, in particular voters, generate motivated biases that shape their choices.

References

Bénabou, R., & Tirole, J. (2002). Self-confidence and personal motivation. Quarterly Journal of Economics, 117(3), 871-915.

Bénabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. American Economic Review, 96(5), 1652-1678.

Bénabou, R., & Tirole, J. (2011). Identity, morals, and taboos: Beliefs as assets. Quarterly Journal of Economics, 126(2), 805-855.

Bodner, R. & Prelec, D. (2003) Self-signaling and diagnostic utility in everyday decisionmaking, in Brocas, I., & Carrillo J. D. (eds), The psychology of economic decisions. Oxford: Oxford University Press.

Borah, A. (2019). Voting expressively. Economic Inquiry, 57(3), 1617-1635.

Cantoni, D., Yang, D. Y., Yuchtman, N., & Zhang, Y. J. (2019). Protests as strategic games: experimental evidence from Hong Kong's antiauthoritarian movement. Quarterly Journal of Economics, 134(2), 1021-1077.

Clarke, D., Romano, J. P., & Wolf, M. (2020). The Romano–Wolf multiple-hypothesis correction in Stata. The Stata Journal, 20(4), 812-843.

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. Economic Theory, 33(1), 67-80.

DellaVigna, S., List, J. A., Malmendier, U., & Rao, G. (2016). Voting to tell others. Review of Economic Studies, 84(1), 143-181.

Di Tella, R., Perez-Truglia, R., Babino, A., Sigman, M. (2015). Conveniently upset: avoiding altruism by distorting beliefs about others' altruism. American Economic Review 105 (11), 3416–3442.

Ehrich, K. R., & Irwin, J. R. (2005). Willful ignorance in the request for product attribute information. Journal of Marketing Research, 42(3), 266-277.

Eil, D., & Rao, J. M. (2011). The good news-bad news effect: asymmetric processing of objective information about yourself. American Economic Journal: Microeconomics, 3(2), 114-38.

Enke, B. (2020). Moral values and voting. Journal of Political Economy, 128(10), 3679-3729.

Falk, A. (2021). Facing yourself – A note on self-image. Journal of Economic Behavior & Organization, 186, 724-734.

Feddersen, T., Gailmard, S., & Sandroni, A. (2009). Moral bias in large elections: Theory and experimental evidence. American Political Science Review, 103(2), 175-192.

Feiler, L. (2014). Testing models of information avoidance with binary choice dictator games. Journal of Economic Psychology, 45, 253-267.

Freddi, E. (2019). Do people avoid morally relevant information? Evidence from the refugee crisis. Review of Economics and Statistics, 1-45.

Fryer Jr, R. G., Harms, P., & Jackson, M. O. (2019). Updating beliefs when evidence is open to interpretation: Implications for bias and polarization. Journal of the European Economic Association, 17(5), 1470-1501.

Gao, L., Wheeler, S. C., & Shiv, B. (2009). The "shaken self": Product choices as a means of restoring self-view confidence. Journal of Consumer Research, 36(1), 29-38.

Garrett, K. N., & Bankert, A. (2020). The moral roots of partisan division: How moral conviction heightens affective polarization. British Journal of Political Science, 50(2), 621-640.

Gerber, A. S., Green, D. P., & Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. American Political Science review, 102(1), 33-48.

Golman, R., Hagmann, D., & Loewenstein, G. (2017). Information avoidance. Journal of Economic Literature, 55(1), 96-135.

Grossman, Z. (2014). Strategic ignorance and the robustness of social preferences. Management Science, 60(11), 2659-2665.

Grossman, Z., & Van Der Weele, J. J. (2017). Self-image and willful ignorance in social decisions. Journal of the European Economic Association, 15(1), 173-217.

Haidt, J. (2012). The righteous mind: Why good people are divided by politics and religion. Vintage.

Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. Journal of Personality and Social psychology, 60(6), 895-910.

Kuzmanovic, B., Rigoux, L., & Tittgemeyer, M. (2018). Influence of vmPFC on dmPFC predicts valence-guided belief formation. Journal of Neuroscience, 38(37), 7996-8010.

Lee, J. L., Nader, K., & Schiller, D. (2017). An update on memory reconsolidation updating. Trends in Cognitive Sciences, 21(7), 531-545.

Le Yaouanq, Y. (2021) Motivated cognition on a model of voting. Mimeo.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. Journal of Personality and Social Psychology, 37(11), 2098.

Matthey, A., & Regner, T. (2011). Do I really want to know? A cognitive dissonance-based explanation of other-regarding behavior. Games, 2(1), 114-135.

McCright, A. M., & Dunlap, R. E. (2011). The politicization of climate change and polarization in the American public's views of global warming, 2001–2010. Sociological Quarterly, 52(2), 155-194.

Morgan, J., & Várdy, F. (2012). Mixed motives and the optimal size of voting bodies. Journal of Political Economy, 120(5), 986-1026.

Munro, G. D., & Ditto, P. H. (1997). Biased assimilation, attitude polarization, and affect in reactions to stereotype-relevant scientific information. Personality and Social Psychology Bulletin, 23(6), 636-653.

Nickerson, D. W., & Rogers, T. (2010). Do you have a voting plan? Implementation intentions, voter turnout, and organic plan making. Psychological Science, 21(2), 194-199.

Niehaus, P. (2020). A theory of good intentions. Mimeo.

Nyborg, K. (2011). I don't want to hear about it: Rational ignorance among duty-oriented consumers. Journal of Economic Behavior & Organization, 79(3), 263-274.

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. Political Behavior, 32(2), 303-330.

Osborne, M. J., & Turner, M. A. (2010). Cost benefit analyses versus referenda. Journal of Political Economy, 118(1), 156-187.

Pomerantz, E. M., Chaiken, S., & Tordesillas, R. S. (1995). Attitude strength and resistance processes. Journal of Personality and Social Psychology, 69(3), 408-419.

Rabin, M., & Schrag, J. L. (1999). First impressions matter: A model of confirmatory bias. Quarterly Journal of Economics, 114(1), 37-82.

Romano, J. P., & Wolf, M. (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. Journal of the American Statistical Association, 100(469), 94-108.

Romano, J. P., & Wolf, M. (2005b). Stepwise multiple testing as formalized data snooping. Econometrica, 73(4), 1237-1282.

Romano, J. P., & Wolf, M. (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. Statistics & Probability Letters, 113, 38-40.

Sandel, M. J. (2005). Public philosophy: Essays on morality in politics. Harvard University Press.

Schneider, F. H. (2020). Signaling moral values through consumption. Working paper series/Department of Economics, (367).

Serra-Garcia, M., & Szech, N. (2019). The (in) elasticity of moral ignorance. CESifo Working Paper 7555

Sharot, T., & Garrett, N. (2016). Forming beliefs: Why valence matters. Trends in Cognitive Sciences, 20(1), 25-33.

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. Nature Neuroscience, 14(11), 1475-1479.

Shayo, M., & Harel, A. (2012). Non-consequentialist voting. Journal of Economic Behavior & Organization, 81(1), 299-313.

Sunstein, C. R., Bobadilla-Suarez, S., Lazzaro, S. C., & Sharot, T. (2016). How people update beliefs about climate change: Good news and bad news. Cornell Law Review, 102, 1431.

Yao, Z., Lin, X., & Hu, X. (2021). Optimistic amnesia: how online and offline processing shape belief updating and memory biases in immediate and long-term optimism biases. Social Cognitive and Affective Neuroscience, 16(5), 453-462.

Appendix A: Interventions (translated from German) **and additional tables**

A.I Information provided in treatments HIGH and LOW

The interventions HIGH and LOW had the following wording (translation from German):

Treatment	Text shown
HIGH	Did you know that according to a scientific study (Arluke and Madfis 2013, available on request) cruelty to animals and anti-social behaviour towards humans are correlated? The study reports that those being cruel to animals are more likely to conduct criminal acts against humans. Examples from the study: ● Someone torturing animals is much more likely to be violent against humans than someone who is kind towards animals. ● Someone torturing animals is much more likely to run amok than someone who is kind towards animals. ● Someone torturing animals is much more likely to disrespect property rights than someone who is kind towards animals. According to psychological research a common cause of anti-social behavior is a lack of compassion (empathy). Another study (Erlanger und Tsytsarev 2012, available on request) shows that: Compassionate people are much more likely to treat animals kindly than non-compassionate people. Compassionate people are much more opposed to cruelty to animals and animal testing than non-compassionate people. Being compassionate is a necessary condition for kind-hearted behavior. Overall this implies: Kind-hearted people who care about the wellbeing of others and the good rules of living together are also more caring towards animals!

LOW	Did you know that according to a scientific study (Levin, Arluke und Irvine 2017, available on request) care for animals and indifference towards humans can co-exist? The study reports that those helping animals might well ignore the suffering of other humans. Examples from the study: ● A call for donations to help a sickly dog motivated more people to donate than a call for donations of a sickly child. ● A dog that had been knocked out induced an emotional response in more people than an adult that had been knocked out. What is the reason for some people to be more indifferent towards other people than towards animals? According to the researchers, a possible reason is that such people believe humans but not animals to be responsible („at fault“) for their own hardship. The following true event provides further evidence for the possibility that compassion towards animals and indifference towards humans can co-exist: In a western industrialized country many people actively protested that a police officer who shot a dog out of an unfounded feeling of threat gets punished. The same people did not care whether a police officer who shot a mentally ill woman out of an unfounded feeling of threat gets punished. Being compassionate is a necessary condition for kind-hearted behavior. Overall this implies: People that care about animals are not necessarily kind-hearted people that care about the wellbeing of others and the good rules of living together!

A.II Additional Tables

Table A.1: Variable descriptions

Explanatory variables	
HIGH	Dummy variable that equals 1 if participant in HIGH treatment
LOW	Dummy variable that equals 1 if participant in LOW treatment
BUBBLE	Dummy variable that equals 1 if participant in BUBBLE treatment
CONFRONT	Dummy variable that equals 1 if participant in CONFRONT treatment
Female	Dummy variable that equals 1 if participant reports to be female (rather than male or other).
Age categ.	Dummies for eight age categories. Lowest bracket: '18-24 years', then in steps of ten years up to 84. Highest bracket: 'above 84'.
Farmer	Dummy variable that equals 1 if participant reports to work as a farmer
FarmHorn	Dummy variable that equals 1 if participant reports to keep horned farm animals in particular horned cows or goats
Informed	Categorial variable centered around zero with seven categories. -3 indicated that the participant reports to be 'not at all informed' and 3 that the participant reports to be 'very well informed' about the Horncow Initiative and the upcoming ballot
PriorAttitude	Categorial variable on a seven point Likert scale measuring the attitude towards the Horncow Initiative. 1 represents 'Certainly against' and 7 'certainly in favor'.
Emotions	Categorial variable centered around zero with seven categories. -3 indicated that the participant reports that (s)he does 'not at all' and 3 that the participant reports to 'very much' respond emotionally to the Horncow Initiative.
FreqMeat	Categorial variable on an eight point scale capturing the self-reported estimate of the frequency of eating red or white meat or meat products such as sausages, ham and entrails. Categories: 1: never; 2: only as an exception; 3: once a month; 4: several times a month; 5: once a week; 6: several times a week; 7: once a day; 8: several times a day.
Intensive	Dummy variable that equals 1 if participant reports to eat meat at least once a day. Constructed from *FreqMeat.*
Vegetarian	Dummy variable that equals 1 if participant reports never to eat meat. Constructed from *FreqMeat.*
Vegan	Dummy variable that equals 1 if participant reports to adhere to a vegan diet.
NoEggsMilk	Dummy variable that equals 1 if participant reports not to eat eggs and milk.
GoodEffects	Categorial variable on a seven point Likert scale measuring the attitude towards the statement that consequences are more important than intentions of someone's actions. 1 represents 'Certainly against' and 7 'certainly in favor'.

GoodIntent	Categorial variable on a seven point Likert scale measuring the attitude towards the statement that intentions are more important than consequences of someone's actions. 1 represents 'Certainly against' and 7 'certainly in favor'.
HIGHxGoodIntent	Interaction between variables *HIGH* and *GoodIntent*
Overconfident	Dummy variable that equals 1 if participant's self-reported degree of informedness (based on variable *Informed*) is above the median response (=0) but at the same time the participant's performance in the quiz is below the median performance (8 out of 10 questions correctly answered).

Outcome variables

AvoidanceCONTRA	Dummy variable that equals 1 if participant chooses not to read the arguments opposing the Horncow Initiative.
ReadOpposingAttitute	Dummy variable that equals 1 if participant chooses to read the arguments opposing his/her own *PriorAttitude* towards the Horncow Initiative.
ΔIntendedVote	Continuous variable bound to interval [-1,1] capturing the normalized difference between the self-reported anticipated voting at the end of the first wave and *PriorAttitude*. The variable is computed as follows: *ΔIntendedVote = [(AnticipatedVoting - 3)/2 - (PriorAttitude - 4)/3]/2* such that negative numbers indicate that the likelihood to vote in favor of the initiative has decreased relative to the attitude expressed before the exposition to the PRO and/or CONTRA arguments. Where *AnticipatedVoting* is a categorical variable that measures the participant's voting plan in the ballot: 1 'certainly vote against the initiative'; 2 'likely to vote against the initiative'; 3 'I have not yet formed an opinion on how to vote', 4 'likely to vote infavor of the initiative', 5 'certainly vote infavor of the initiative'.
ΔReportedVote	Continuous variable bound to interval [-1,1] capturing the normalized difference between the self-reported actual voting and *PriorAttitude*. The variable is computed as follows: *ΔReportedVote = ReportedVoting - PriorAttitude/7* such that negative numbers indicate that the likelihood to vote in favor of the initiative has decreased relative to the attitude expressed before the exposition to the PRO and/or CONTRA arguments. Where *ReportedVoting* is a dummy that equals 1 if the participant reports to have voted in favor of the initiative.
ΔAgreementPRO	Continuous variable bound to interval [-1,1] capturing the normalized difference between the self-reported agreement with arguments infavor of the initiative at the end of the first wave and *PriorAttitude*. The variable is computed as follows: *ΔAgreementPRO = [(AgreementPRO - 3)/2 - (PriorAttitude - 4)/3]/2* such that negative numbers indicate that the agreement with arguments in favor of the initiative has decreased relative to the attitude expressed before the exposition to the PRO and CONTRA arguments. Where *AgreementPRO* is a categorical variable

	capturing how much the participant the PRO arguments he/she has just read convince him/her: 1 'not at all convincing'; 2 'more unconvincing than convincing, 3 'neither convincing nor unconvincing', 4 'more convincing than unconvincing', 5 'fully convincing'.
ΔAgreementCON	Continuous variable bound to interval [-1,1] capturing the normalized difference between the self-reported agreement with arguments against of the initiative at the end of the first wave and *PriorAttitude*. The variable is computed as follows: *ΔAgreementCON = [(AgreementCONTRA - 3)/2 - (PriorAttitude - 4)/3]/2* such that negative numbers indicate that the agreement with arguments in favor of the initiative has decreased relative to the attitude expressed before the exposition to the PRO and CONTRA arguments. Where *AgreementCONTRA* is a categorical variable capturing how much the participant the CONTRA arguments he/she has just read convince him/her: 1 'not at all convincing'; 2 'more unconvincing than convincing, 3 'neither convincing nor unconvincing', 4 'more convincing than unconvincing', 5 'fully convincing'

Table A.2: Attitudes and Voting in Sample vs. Ballot

	Infavor	**Opposing**	**Neutral**	*N*
Participants completing wave 1				
PriorAttitude	40.4	42.8	16.8	1,825
AnticipatedVoting	36.6	40.6	22.8	1,954
Participants completing both waves				
PriorAttitude	45.3	37.2	17.5	1,057
AnticipatedVoting	43.9	36.7	19.4	1,021
ReportedVoting	38.5	61.5		772
Ballot Result				
all of Switzerland	45.3	52.9		2.62 million
German speaking cantons	43.8	53.5		1.93 million

Note: Source of ballot results: Bundesamt für Statistik, Statistik der eidg. Volksabstimmungen (Abst.-Nr. 6230)

Table A.3: Avoidance of CONTRA arguments (Logit)

	(1) Logit	(2) OLS	(3) Logit	(4) OLS	(5) OLS/wave 1
HIGH	0.048	0.049	0.045	0.046	0.031
	(0.206)	(0.199)	(0.229)	(0.236)	(0.273)
LOW	0.000	0.000	0.004	0.003	-0.027
	(0.999)	(0.999)	(0.915)	(0.933)	(0.359)
BUBBLE			-0.052	-0.046	0.031
			(0.222)	(0.252)	(0.298)
CONFRONT			0.023	0.025	0.025
			(0.534)	(0.526)	(0.393)
Informed			0.026	0.025	0.001
			(0.030)	(0.036)	(0.880)
Farmer			-1.624	-0.150	-0.218
			(0.985)	(0.577)	(0.265)
FarmHorn			1.799	0.435	0.320
			(0.984)	(0.255)	(0.249)
PriorAttitude			0.004	0.004	0.004
			(0.790)	(0.790)	(0.742)
Female			0.008	0.007	0.002
			(0.794)	(0.836)	(0.919)
Age categ.	*No*	*Yes*	*No*	*Yes*	*Yes*
_cons		0.160		0.162	0.128
		(0.000)		(0.109)	(0.099)
N	578	578	574	574	1079
R^2		0.004		0.032	0.015
F		1.064		1.229	1.036
aic		529.3		520.2	1035.4
bic		542.4		589.8	1120.1

Note: Dependent variable: Dummy variable whether CONTRA arguments have been avoided (1 = avoided, 0 = read). For the logit regressions (1) and (3) marginal effects are presented. *p*-values in parentheses unadjusted for multiple hypothesis testing. *Romano-Wolf p*-values not reported as even unadjusted *p*-values do not allow to reject null hypothesis.

Table A.4: Reading arguments opposing own attitude

	(1) Logit	(2) OLS	(3) Logit	(4) OLS	(5) OLS/wave 1
BUBBLE	0.051	0.049	0.049	0.047	0.003
	(0.102)	(0.102)	(0.115)	(0.119)	(0.909)
Romano-Wolf				(0.366)	
CONFRONT	0.013	0.014	0.011	0.011	-0.012
	(0.652)	(0.645)	(0.715)	(0.726)	(0.619)
HIGH			-0.032	-0.033	-0.015
			(0.268)	(0.266)	(0.517)
LOW			-0.006	-0.006	0.004
			(0.851)	(0.842)	(0.873)
Informed			-0.015	-0.014	0.000
			(0.087)	(0.092)	(0.945)
Farmer			0.087	0.077	0.154
			(0.598)	(0.582)	(0.159)
FarmHorn			-0.201	-0.239	-0.146
			(0.353)	(0.289)	(0.383)
PriorAttitude			0.007	0.008	0.013
			(0.237)	(0.223)	(0.005)
Female			-0.014	-0.013	-0.016
			(0.565)	(0.588)	(0.416)
Age categ.	*No*	*Yes*	*No*	*Yes*	*Yes*
_cons		0.791		0.810	0.801
		(0.000)		(0.000)	(0.000)
N	1057	1057	1049	1052	1814
R^2		0.003		0.019	0.014
F		1.339		1.248	1.540
aic		1041.4		1035.1	1884.7
bic		1056.3		1119.4	1978.3

Note: Dependent variable: Dummy variable whether arguments opposing ones' own prior attitude have been read (1 = read, 0 = avoided). For the logit regressions (1) and (3) marginal effects are presented. *p*-values in parentheses unadjusted for multiple hypothesis testing unless specified otherwise; *Romano-Wolf p*-values in (4) corrected for 20 hypotheses (outcome variables *ReadOpposingAttitute, AvoidanceCONTRA, ΔIntendedVote, ΔReportedVote, ΔAgreementPRO*) for treatments *HIGH, LOW, BUBBLE* and *CONFRONT*) based on 10,000 replications**.**

Table A.5: Using prior attitude as control

	AgreementPRO		*IntendedVote*		*ReportedVote*	
	(1)	(2)	(3)	(4)	(5)	(6)
HIGH	-0.130	-0.183	0.057	0.155	0.035	0.064
	(-1.54)	(-2.75)	(0.56)	(2.83)	(0.81)	(2.20)
LOW	-0.058	-0.025	0.021	-0.014	0.022	0.017
	(-0.68)	(-0.37)	(0.19)	(-0.25)	(0.51)	(0.56)
BUBBLE	0.170*	0.098	-0.047	0.071	-0.030	0.020
	(2.01)	(1.46)	(-0.45)	(1.25)	(-0.69)	(0.68)
CONFRONT	0.139	0.111	-0.014	-0.001	0.006	0.017
	(1.60)	(1.62)	(-0.14)	(-0.02)	(0.14)	(0.56)
PriorAttitude		-0.317		0.586		0.170
		(-22.24)		(50.55)		(29.15)
_cons	2.587	3.813	2.870	0.619	0.372	-0.271
	(37.60)	(49.26)	(34.63)	(9.86)	(10.88)	(-8.40)
N	825	825	1021	1021	772	772
R^2	0.010	0.383	0.001	0.716	0.002	0.527
aic	2345.8	1958.0	3542.9	2260.8	1087.1	512.8
bic	2369.3	1986.3	3567.5	2290.4	1110.3	540.7

Note: t statistics in parentheses. Adding PriorAttitude as a control substantially changes both coefficients and significance levels of treatment variables. Moreover, t-values of PriorAttitude coefficients are exceptionally high, suggesting endogeneity due to an omitted variable bias. Hence, estimated coefficients, including those for treatment variables are not reliable.

Table A.6: Voting Non-Parametric Tests for Self-Selection Concern

Outcome Variable	**Treatments**			**Mann-Whitney** (*p*-values)	*N*
ΔIntendedVote (all participants completing Wave 1)					
ALL or NONE	HIGH	vs.	NEUTRAL	0.020	1134
ALL	HIGH	vs.	NEUTRAL	0.020	952
NONE	HIGH	vs.	NEUTRAL	0.625	182
ALL or NONE	ALL	vs.	NONE	0.926	1134
ΔIntendedVote (all participants completing Wave 2)					
ALL or NONE	HIGH	vs.	NEUTRAL	0.042	505
ALL	HIGH	vs.	NEUTRAL	0.097	435
NONE	HIGH	vs.	NEUTRAL	0.137	70
ALL or NONE	ALL	vs.	NONE	0.871	505

Appendix B: Theory background

We present here the simplest model we can think of to illustrate the impact of our experimental manipulation on information processing and voting.

There are two states of nature, $x = 0$ and $x = 1$. If $x = 1$, the initiative, if passed, would improve animal welfare. If $x = 0$, animal welfare would remain unchanged even if the initiative is passed. If the initiative is not passed, animal welfare remains unchanged, too.

Individual i observes the state of the world x. But he can pay a cost c_i to bias his belief so that he believes the opposite state of the world to obtain. In particular, if x = 0, he can pay c_i to get belief $x' = 1$. If he does not pay c_i, then $x' = x$.

His action taken after belief formation is $v = 1$ (YES vote) or $v = 0$ (NO vote). The agent derives utility from voting according to his subjective preferences: $u(v|x' = 1) = \{\mu$ if $v = 1$, 0 else$\}$ and $u(v|x' = 0) = 0$. The parameter $\mu > 0$ is discussed below.

It is easy to see from this utility function that there is no incentive to bias one's belief when $x = 1$, but there is an incentive to do so when $x = 0$. The agent biases his belief, moving from true $x = 0$ to false $x'=1$ (and naively forgetting that x' was forged) if and only if $\mu > c_i$.

How to interpret μ? It can be conceived as $\mu = m + \pi$, where $m \geq 0$ is the hedonic utility derived by the individual from the real consequence of the initiative on animals (possibly accounting for the probability of being pivotal), and π the "ego utility" if the voter believes that if he improves animal welfare then he is a good person.

In the experiment, we manipulate this ego utility π. Since voters are indexed by c_i, by increasing π in WARM, we thus increase the share of Yes-voters by increasing the share of those who bias themselves.

We refer to Le Yaouanq (2021) for a more general model of voting with biased beliefs. Building on the memory management model of Bénabou and Tirole (2002, 2011), Le Yaouanq shows in his Proposition 1 that, in any equilibrium, an increase in μ increases the probability to bias beliefs, consistent with the prediction above.

Appendix C: Further results

In this section, we report a number of interesting correlations between variables elicited in the survey as detailed in the pre-registration. However, these relationships cannot be interpreted causally, and several variables were elicited after the treatment intervention (GoodIntent, GoodEffect, FreqMeat, Intensive, Vegetarian, Vegan, NoEggsMilk, Overconfident) and hence can correlate due to past exposure to these interventions. Table C.1 shows that anticipated as well as reported votes in favor of the initiative decrease in the frequency of meat eating but not with other dietary habits related to animal products.
In addition, we test whether the self-reported level of prior informedness on the initiative correlated to specific ethical attitudes toward consequentialism. These attitudes are expressed by, first, the degree to which participants report to agree with a claim stating that

rewards should be given to those whose actions result in good consequences regardless of his or her intentions (GoodEffects), and, second, the degree to which they agree with a claim stating that rewards should be given to those with good intentions regardless of the consequences of these actions (GoodIntent). If looked at separately, we find a negative correlation. In a joint analysis (regression (3) in Table C.2), only GoodIntent remains significant. Hence, participants that reported to care for intentions behind actions also report to be less well informed. This is consistent as knowing the consequences of one's actions (and votes) is less relevant if one's focus is on intentions rather than consequences.

Table C.1: Voting (ordered Logit)

	(1)	(2)	(3)	(4)	(5)
	IntendedVote		*Int.V./wave 1*	*ReportedVote*	
FreqMeat	-0.096	-0.147	-0.174	-0.189	-0.206
	(0.018)	(0.075)	(0.001)	(0.001)	(0.038)
Intensive		0.063	0.103		-0.003
		(0.728)	(0.413)		(0.991)
Vegetarian		-0.436	-0.237		-0.188
		(0.390)	(0.500)		(0.782)
Vegan		-0.361	-0.407		-1.166
		(0.719)	(0.580)		(0.489)
NoEggsMilk		0.502	0.312		0.809
		(0.285)	(0.366)		(0.295)
_cons				0.469	0.545
				(0.107)	(0.245)
N	1021	1018	1949	772	770
chi2	5.612	7.539	23.455	11.328	12.671
Pseudo R^2	0.002	0.002	0.004	0.011	0.012
Aic	3272.2	3269.2	6246.4	1021.5	1024.3
Bic	3296.9	3313.5	6296.6	1030.8	1052.2

Notes: Dependent variable: *IntendedVote* and *ReportedVote*, reported are coefficients, *p*-values in parentheses

Table C.2: Prior information (ord. Logit)

	(1)	(2)	(3)
GoodEffects	-0.070		-0.038
	(0.048)		(0.303)
GoodIntent		-0.095	-0.073
		(0.005)	(0.043)
N	1015	1019	1001
chi2	3.90	7.89	7.10
Pseudo R^2	0.001	0.002	0.002
Aic	3760.5	3764.3	3703.9
Bic	3795.0	3798.8	3743.2

Notes: Dependent variable: *Informed*; coefficients reported, *p*-values in parentheses

Furthermore, both variables measuring information selection that we used in the previous section are not significantly correlated with proxies of ethical schools of thought (Table C.3). However, they are highly significantly correlated with both how emotionally touched participants are by the initiative (*Emotions*) and how much their self-assessed prior informedness (*Informed*) with respect to the initiative differs from their performance in a quiz about the initiative and horned animals (*Overconfident*). The latter is a dummy that equals one if a participant is above the median with respect to self-reported informedness but below the median in terms of quiz performance. Emotional involvedness is associated with more and overconfidence with less information selection.

Table C.3: Information selection (Logit)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		AvoidanceCONTRA				*ReadOpposingAttitute*		
GoodEffects	-0.088			-0.053	0.086			0.049
	(0.106)			(0.348)	(0.110)			(0.387)
GoodIntent	0.048			0.029	-0.101			-0.088
	(0.362)			(0.597)	(0.052)			(0.107)
Emotions		-0.150	-0.170	-0.168		0.149	0.147	0.171
		(0.001)	(0.000)	(0.001)		(0.001)	(0.000)	(0.001)
Overconfident		1.070	0.895	1.155		-1.127	-0.989	-1.189
		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
_cons	-1.500	-1.809	-1.539	-1.883	1.467	1.809	1.610	1.866
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	1002	1031	1964	979	1002	1031	1783	979
chi2	2.84	50.40	80.65	55.66	4.95	55.11	79.34	60.70
Pseudo R^2	0.003	0.050	0.040	0.060	0.005	0.055	0.044	0.065
Aic	955.8	955.6	1961.4	882.7	962.6	959.6	1747.6	883.7
Bic	970.5	970.4	1978.2	907.1	977.3	974.4	1764.1	908.1

Notes: Dependent variable: *AvoidanceCONTRA* (regressions (1) – (4)) and *ReadOpposingAttitute* (regressions (5) – (8)). Regressions (3) and (5) based on all participants that completed wave 1 of the survey, all other regressions based on sample completing both waves. *p*-values in parentheses.

Information acquisition in wave 1 is consistent with participants suffering from confirmation bias. Those ex-ante supporting (opposing) the initiative are more likely to only read the arguments supporting (opposing) the initiative.[16]

[16] Both Chi²-tests and univariate logit regressions yield $p < 0.01$. Results available upon reported.