



**HAL**  
open science

# Norm from the top: a social norm nudge to promote low-practiced behaviors without boomerang effect

Alix Rouillé

► **To cite this version:**

Alix Rouillé. Norm from the top: a social norm nudge to promote low-practiced behaviors without boomerang effect. 2022. halshs-03673004

**HAL Id: halshs-03673004**

**<https://shs.hal.science/halshs-03673004>**

Preprint submitted on 19 May 2022

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

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

Norm from the top:  
a social norm nudge to promote low-practiced  
behaviors without boomerang effect\*

Alix Rouillé <sup>†</sup>

Ph.D. Student at ENS Paris-Saclay<sup>‡</sup>

May 19, 2022

---

\*Many thank to Anne Corcos, Francois Pannequin, Thomas Vendryes, Thibault Richard, Ninon Moreau-Kastler, Jose de Sousa, Andrea Guido, for their comments and advice.

<sup>†</sup>[alix.rouille@ens-paris-saclay.fr](mailto:alix.rouille@ens-paris-saclay.fr)

<sup>‡</sup>Université Paris-Saclay, ENS Paris-Saclay, Centre d'Economie de l'ENS Paris-Saclay, 91190, Gif-sur-Yvette, France.

## Abstract

Social norms have proven to be a powerful nudge to make people adopt prosocial behavior. Informing people that most of their peers behave virtuously encourages them to improve their own behavior. However, since feedback is based on the average behavior of the population, the targeted desirable behavior must already be practiced by a majority of the population in order to avoid the boomerang effect. The boomerang effect is defined as a deterioration in attitude towards prosocial behavior. This deterioration is by the people who contributed to prosocial behavior more than the average; once they are informed about this, they modify their action accordingly. In this study, our purpose is to create a norm that can be implemented as nudges in behaviors where current social norm nudges are inefficient. This novel implementation could increase the range of prosocial behaviors that can be enhanced by social norm nudges. Within a nudge framework, we build a new norm that provides information based on the most altruistic people in the population. By having participants fill out additional surveys related to environmental topics, we found that this new norm, i.e., "Norm from the top", acted as an efficient nudge, increasing the average contribution to prosocial behavior. In contrast, the standard norm does not have a significant effect due to the boomerang effect. These results show the potential of applying the "Norm from the top" to promote low-practiced prosocial behaviors.

**JEL CODES:** C91, D91.

**Key words:** Social norms, nudge, prosocial behavior, descriptive norm, injunctive norm.

# 1 Introduction

The interest in the impact of social norms on behaviors has been growing for the last century. Since the experiment of Salomon Asch in 1951 emphasizing the role of conformity in the decision process, numerous studies have improved our understanding of this concept (Deutsch and Gerard 1955; Reno et al. 1993; Bicchieri 2006). Social norms can influence both favorable behaviors (keeping public places clean, respecting lines in stores) and detrimental ones (criminality [Glaeser et al. 1996]; smoking [Christakis and Fowler 2008]). According to Bicchieri (2016), social norms can be defined as unofficial rules such that individuals tend to conform to it if: (1) most people in their reference network conform to it (empirical expectation) or (2) if most people in their reference network believe they ought to conform to it (normative expectation).

In contrast to the established work in social norms, nudge theory is more recent and has been highlighted by Sunstein and Thaler (2008). A nudge can be defined as a slight modification in the choice architecture that can switch individuals' decisions towards better, healthier, more virtuous outcomes for them, or for the society, without decreasing their own welfare.

For the last few years, the combination between nudges and social norms has proven to be a relevant tool for encouraging people to make prosocial choices (Frey and Meier 2004; Glazier and Mone 2019) or pro-environmental choices (Goldstein et al. 2008; Kurz 2018). To illustrate the application of this combination, in the context of energy conservation, nudges tended to encourage households to bring their energy consumption closer to that of their neighbors (Allcott 2011; Costa and Kahn 2013). A standard way to apply a social norm nudge is to inform people that most of their group acts in a given way (the desirable behavior) to incite them to adopt the norm. The standard structure of such information is as follows: "90% of people turn off the light before leaving the room". The objective is then two-fold, to convince the remaining 10% left to turn off the light, and to confirm to others that their behavior is correct. This concept is named "descriptive norm". A descriptive norm can also be associated with an injunctive norm that indicates

the approved behavior, in our example: "and avoid energy wastes, which are detrimental to the environment". A social norm nudge is comprised of a descriptive norm, i.e., what the behavior of other people is, and an injunctive norm, i.e., what other people approve of (Loschelder et al. 2019).

However, social norm nudges are efficient if the proportion of people who already act prosocially is above a threshold. Otherwise, the norm can potentially produce a "boomerang effect" (first documented by Schultz et al. 2007); namely, this incites a person who acts prosocially to lessen her behavior since she is informed that her prosocial contribution is above the average. Boomerang effect issues reduce the scope of social norm nudges to attitudes already practiced by most of the population, except in a few studies using "trending norms" (Mortensen et al. 2017; Sparkman and Walton 2017).

In this paper, we create and implement a new social norm nudge to solve the boomerang effect issue in the context of prosocial behaviors not practiced by the majority of people in a given group. The objective is to shape new mechanisms to extend the range of behaviors where nudges and social norms could be applied. In our experiment, we set up a new type of feedback involving descriptive and injunctive norms, and then we observe whether they are more efficient than the standard social norm feedback. This new type of feedback provides information that focuses only on the people who contribute the most to prosocial behavior in a population, i.e., "top contributors", instead of informing about the average behavior. We call that feedback "Norm from the top". This paper investigates whether a solution for solving the boomerang effect issue can be found in constructing norms based on the top contributors of a given population, thus increasing the average attitude towards prosocial behavior when compared to a standard norm.

The results of our experiment show that, in a context of prosocial behavior not entirely performed by a large part of the population, the standard norm had no impact on the average contribution because of the boomerang effect, consistently with the literature (Schultz et al. 2007; Richter et al. 2018). Furthermore, the standard norm generated a concentration of contributions towards the average behavior. In contrast, the norm from the top did not generate such a detrimental effect and increased contributions to prosocial

behavior. Therefore, the findings of this paper expand the nudging possibilities towards prosocial behaviors still practiced by a minority of people in a group. More generally, this paper advances the literature on nudges by testing whether nudges can affect consumers' prosocial behavior.

The paper is organized as follows: in the next section, we present the experimental design of the study, which includes, our strategy and the several parts that make up the experiment. Then, in the third section, we discuss the hypotheses. The fourth section is dedicated to our methods, where we describe how the recruitment of participants took place and what materials were used. We present the results in the fifth section, and finally, the conclusion and limitations of the study are discussed in the sixth and last section.

## **2 Experimental design**

### **2.1 Strategy of the experiment**

We chose a time donation task as prosocial behavior. Following Mortensen et al. (2017), we asked participants to fill out additional surveys at the end of the experiment without being rewarded. Participants were informed that these surveys concern environmental topics and that the data would be used for another study. The participants had to choose the number of surveys they wanted to fill out (between 0 and 8).

For this study, we created four treatments. The treatments are differentiated by the choice architecture proposed to the participants. Illustrations of the treatments are available in the appendix. Our approach was to first run the experiment without any feedback in order to observe how people behaved without information about their peers (treatment 1). Then, we use the data of this first group to create our three different types of feedback. Thus, the feedback we used are accurate and based on the same type of population as the rest of the sample. Agerström et al. (2016) stressed the importance of providing real

numbers when setting a descriptive norm, yet this is not the method used in all nudge studies. Therefore, by implementing real social norms, deception is avoided, and, we can establish our tool as a nudge (Sunstein and Thaler 2008).

In treatment 2, we provided these participants with the average number of surveys filled out by participants of the first treatment. This feedback corresponded to the standard descriptive norm message (Cialdini et al. 1990; Cialdini 2003).

In the last two treatments, the feedback focused on the most significant contributors of the population (i.e., norm from the top). In treatment 3, these participants received a feedback on the number of surveys previously filled out; here we calculated the average based on the participants who had completed the highest number of questionnaires. This feedback came from the twenty participants who filled out the most surveys, a standard that was set to keep the message credible. Yet we assume it should vary according to the framework.

The feedback of treatment 4 was also made up of the norm from the top with the addition of an injunctive norm dimension (i.e., what one ought to do). In their study on energy consumption, Schultz et al. (2007) set an injunctive norm by displaying a happy/sad emoticon for the below/above average users. In our framework, the highest contributors previously defined are called “the most devoted participants”. As we assume that the term “most devoted” makes participants perceive the behavior as approved by their peers we believe this feedback was comprised of the combination between a descriptive norm (i.e., what the others do) and an injunctive norm.

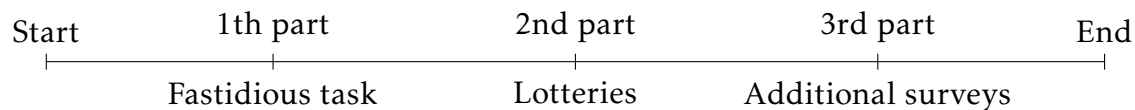
In the literature, several studies showed that this combination is efficient (Habib et al. 2021) and can prevent boomerang effects (Schultz et al. 2007).

## **2.2 Presentation of the experimental design**

The experiment was composed of three parts. During the first part, participants had to complete a fastidious task. They received an endowment once they finish the task. Then,

in the second part, they played four grids of lotteries. Participants were informed that one lottery would be randomly chosen and drawn once the experiment is over. The outcome of this lottery is added or subtracted from the endowment received in the first part. Before the end of the experiment, we asked participants to fill out additional surveys without being remunerated. They faced different frameworks depending on the treatment they were assigned. The timeline of the experiment is summarized in Figure 1 below.

Figure 1: Timeline of the experiment



The three parts are presented in detail in the next section.

### 2.2.1 Effort game and lotteries

During the first part of the experiment, subjects had to carry out an effort game. Indeed, to avoid the "gambling with the house money" effect (Thaler and Johnson 1990) when playing lotteries, beforehand, we asked them to perform a fastidious task: convert ten letter codes into numbers to earn an endowment equal to 16\$. We chose this game from a previous experiment in the literature (Lévy-Garboua et al. 2009). In addition, giving a flat endowment to all participants enabled us to design lotteries in the domain of losses and gains.

The second part was dedicated to lotteries. As discussed earlier, there were four grids of lotteries <sup>1</sup>. The first two grids of lotteries implied risk (Holt and Laury 2002), and the other grids of lotteries implied ambiguity (Chakravarty and Roy 2009). Each grid was made up of 10 pairs of lottery choices with gradual trade-offs to elicit subjects' attitudes

---

<sup>1</sup>the Grid 1 is illustrated in the Appendix



towards risk or ambiguity. Participants had to make a binary choice between each pair of lotteries. For each lottery, there were two potential gains (Grids 1 and 3) or losses (Grids 2 and 4). In the grids under risk, participants could choose a risky option, A, which was a lottery between two distant outcomes. On the other hand, they could choose a safe option, B, which was a lottery between two close outcomes. In the same vein, in the grids under ambiguity, participants could choose an ambiguous option, A, which was a lottery between ambiguous outcomes. Otherwise, they could choose a safe option, B, which was a lottery between two close outcomes.

The principal role of the lottery was first to implement a game that allowed participants to win a payoff and second to distract them to the core of our study to avoid experimenter demand effects (Zizzo 2010). Without the lotteries, the experiment would have been shorter, and we assume that participants could have identified prosocial behavior as the major interest of the experiment and therefore adapted their decisions according to their expectations of the experimenter's desire.

Thus, participants received a flat endowment for completing the tedious task and a potential reward according to their outcome in the lottery. The lottery draw took place a few hours after the experiment. By doing that, we avoided a potential effect of the payoffs obtained by the participants on their willingness to contribute to prosocial behavior.

### **2.2.2 Prosocial behavior**

Before ending the experiment (and the lottery draw), participants received on their screen a message (the message is in the appendix) that presents prosocial behavior. We asked participants whether they are willing to fill out additional surveys without being rewarded. We asked them for this service as a favor to trigger altruistic motivation. These surveys are related to environmental topics and are considered prosocial behavior (Mortensen et al. 2017). Environmental issues seem very suitable here because they should also trigger pro-environmental motivation. Participants had to choose the number of surveys they desire to fill out (between 0 and 8). Each survey is related to a different

environmental topic<sup>2</sup> and is made up of 8 questions. The surveys appeared in a random order for each participant.

The participants who chose not to answer any surveys were conducted directly to the end of the experiment. Before leaving the experiment, all the participants were informed that they would receive an email indicating their final payoff and the procedure to retrieve their gain.

### 3 Hypotheses

In this section, we discuss the hypotheses regarding each treatment of the experiment.

First, we expect to find the usual effect documented in the literature with a standard social norm when the behavior is not entirely performed by a large part of the population, i.e., a backfire from the norm or, as described in the literature, a boomerang effect (Schultz et al. 2007; Richter et al. 2018).

**H1:** A social comparison feedback (Treatment 2) encourages participants to adjust their prosocial contribution towards the average of others' contributions.

According to this hypothesis, fewer participants choose a contribution far from the average behavior than in the treatment without feedback. This consequence has both positive (people that would have contributed less tend to increase their contribution) and negative effects (people that would have contributed more tend to decrease their contribution, i.e., boomerang effect).

As discussed earlier, participants in the third group received the norm from the top feedback. We expect the norm from the top feedback to act as a high benchmark for

---

<sup>2</sup>global environmental sensitivity, waste, waste composting, sustainable purchasing, reuse of electronic goods and devices, energy consumption, transports, consumption of non-reusable plastics.

subjects who have intentions to behave prosocially, thus inciting them to increase their contributions.

**H2:** A feedback that provides a descriptive norm focusing on the greatest contributors (Treatment 3) improves the average attitude towards prosocial behavior.

Despite the expectation of an increase in the average contribution in the third treatment, compared to the second treatment, we expect also a boomerang effect in the third treatment. Indeed, boomerang effects appear when the feedback is only comprised of a descriptive norm without an injunctive norm.

**H3:** A descriptive norm from the top (treatment 3) reduces the number of individual contributions beyond the information provided by this norm, also resulting in a boomerang effect.

Note that this boomerang effect should appear towards a higher level than in treatment 2.

We expect that introducing an injunctive norm would erase the boomerang effect found in treatments 2 and 3 and, therefore, increase the average contribution towards prosocial behavior. This effect has been shown in the literature (Schultz et al. 2007).

**H4:** The combination of injunctive and descriptive norms (treatment 4) increases the contributions to prosocial behavior.

**H5:** This combination erases the boomerang effect (potentially founded in the third treatment [**H3**]).

# 4 Methods

## 4.1 Participants

We recruited 203 participants from Cirano (Montreal, Quebec). The Cirano is an interuniversity center, multidisciplinary and intersectoral. The center has conducted over 800 experimental economics with the participation of the students of environing Montreal universities. They allowed us to diffuse recruitment messages to their base of respondents. We first randomly sent the message to 25% of the base to recruit the respondents of the first treatment (since the feedback of other treatments is built from the responses of the first group). We then sent the message to the rest of the base to compose the three other groups.

The sample is composed of 92 women and 111 men (Average age = 32.1 ; SD = 10.5). Participants obtained an endowment equal to 16\$ for the fastidious task, then the payoffs increased (up to 32\$) or decreased (down to 0\$) according to the result of the lottery. The minimal remuneration was 5\$.

## 4.2 Materials and procedure

The experiment took place online. The 203 participants were randomly assigned to one of the four treatments:

- Control condition (T1; n=49): Participants had no information regarding other participants' actions.
- Standard norm (T2; n=51): Participants were informed that other participants filled out on average four surveys.
- Norm from the top (T3; n=52): Participants were informed that the average of surveys filled out is above seven among the participants who have filled out the most

surveys.

- Norm from the top + injunctive norm (T4; n=51): Participants were informed that the most devoted participants filled out on average more than seven surveys.

One participant of the control group did not finish the experiment, therefore his data had to be removed from the sample. We present the results of the experiment in the next section.

## 5 Results

The results of the experiment are presented in this section. Treatments and hypotheses are summarized below.

Table 1: Treatments

Treatment	Definition	Cue received
Treatment 1	Control condition	/
Treatment 2	Standard norm	4 surveys
Treatment 3	Norm from the top	7 surveys
Treatment 4	Norm from the top + injunctive norm	7 surveys

Table 2: Hypotheses

Hypothesis	Treatment	Definition	Treatment compared
H1	T2	Boomerang effect of behaviors towards the average	T1
H2	T3	Increase of the behaviors on average	T1, T2
H3	T3	Boomerang effect towards the norm from the top	T1, T2
H4	T4	Increase of the behaviors on average	T1, T2, T3
H5	T4	Avoid the boomerang effect	T3

## 5.1 Descriptive statistics

Table 3 and Figure 2 display the number of surveys chosen per treatment. It indicates that participants of treatment 1 chose, on average 4.73 surveys. In the end, the number of surveys they filled out was close to four. Thus, the feedback received by the participants of treatment 2 is four surveys. On the other hand, when we considered the twenty participants who filled out most questionnaires, we obtained an average of more than seven surveys. We chose that formulation because, even if the feedback is real, it had to be credible in the view of the participants. For instance, if we had chosen the first fifteen participants, the average would have been eight. Providing a social norm message informing that people, even the most prominent contributors, chose on average, all the surveys could have been suspicious. Participants of treatment 3 received the information that the greatest contributors completed on average, more than seven surveys each, whereas participants of treatment 4 received the same feedback with the naming "most devoted participants".

Then, when we look at the preliminary results, we see that the average number of questionnaires selected is higher in treatments 3 and 4 than in treatments 1 and 2. We conduct econometric tests later in the paper to evaluate the significance of these differences. Furthermore, the standard deviation is higher in the group without feedback than in other groups. We assume that, for some participants, the feedback might be interpreted as cues about how to behave (more specifically for T2 and T4).

Table 3: Average number of surveys per treatment

	Observations	Mean	Stv.Dev
Treatment 1	48	4.73	3.28
Treatment 2	51	4.84	2.58
Treatment 3	52	5.65	2.98
Treatment 4	51	6.06	2.60

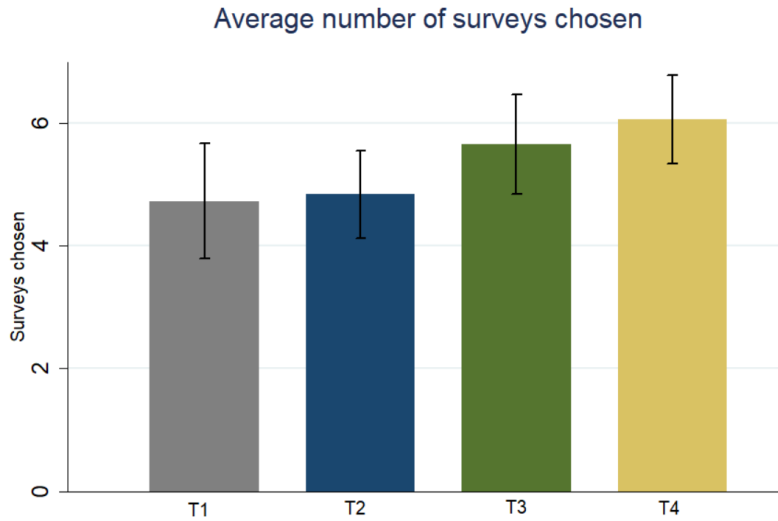


Figure 2: Average surveys per treatment

## 5.2 Sample distribution analysis

### 5.2.1 Graphics

Figures 3 and 4 illustrate the distribution of the number of surveys chosen per participant and per treatment. First, the modal answer is eight surveys (the maximum) for almost every treatment. Second, there is a peak at four surveys in the second treatment (blue line). This peak is worth noting since participants of this group received the information that four surveys correspond to the average behavior. The average behavior seemed to appear as a reference point for participants of this group. Indeed, fewer participants in T2 chose one, two, and eight surveys than in T1. It seems that people who would have selected one or two surveys increased their contributions to meet the average behavior, while the reverse effect applied to some people who would have chosen all of the surveys. These phenomena result in a concentration of contributions towards the average attitude. The dwindling of high contributions in T2 compared to T1 corresponds to the boomerang effect demonstrated in the literature.

In the same vein, we note that treatments 3 and 4 are the groups with the highest number of participants who chose a contribution equal to seven surveys, the norm from the top.

Besides, no boomerang effect seemed to appear in T3. Finally, the median value of the number of surveys chosen for each treatment corresponds to the number of surveys indicated in the feedback: four surveys in T2 and seven surveys in T3 and T4. The median in T1 is five surveys.

Figure 3: Distribution of the number of surveys chosen per participant

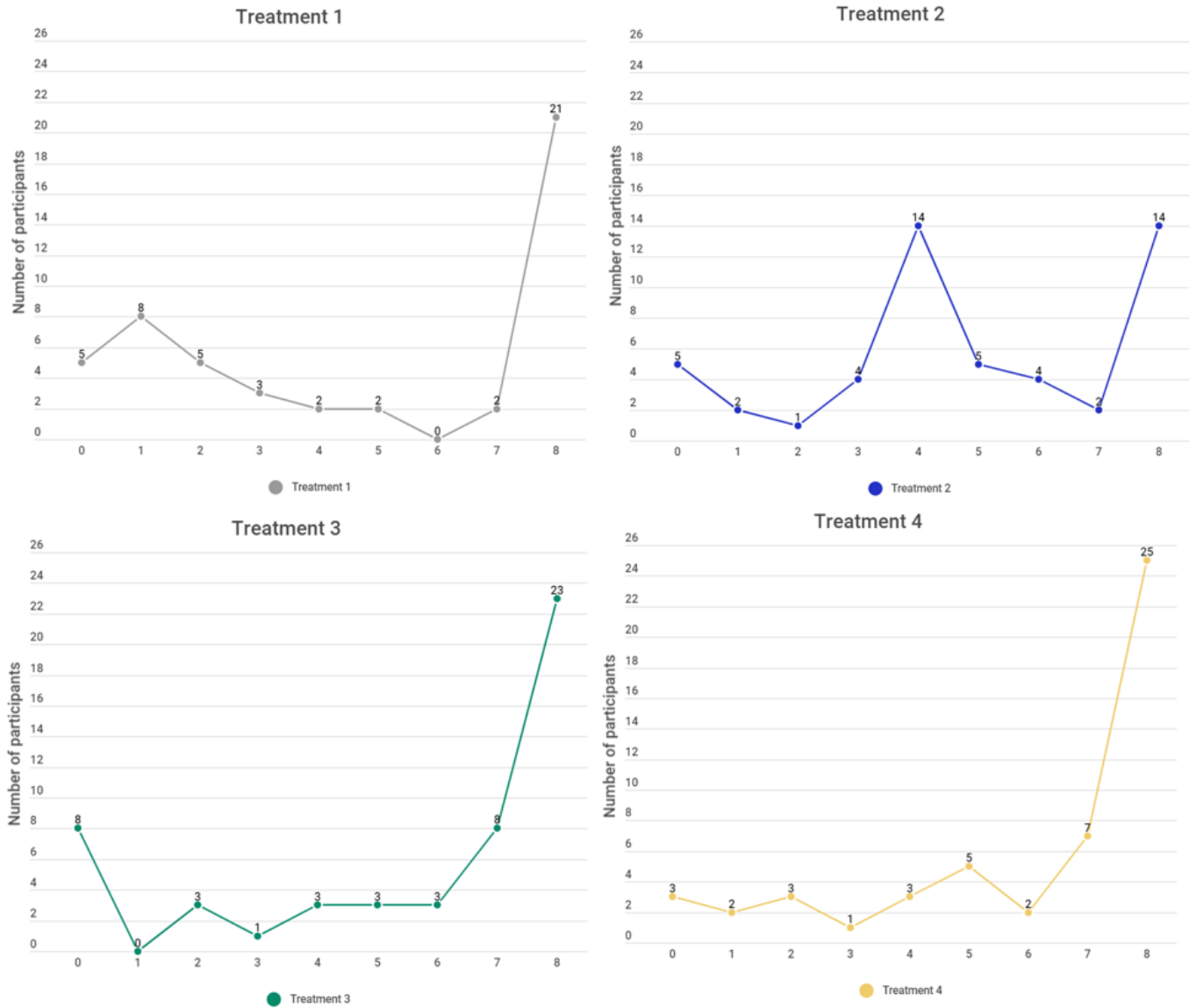
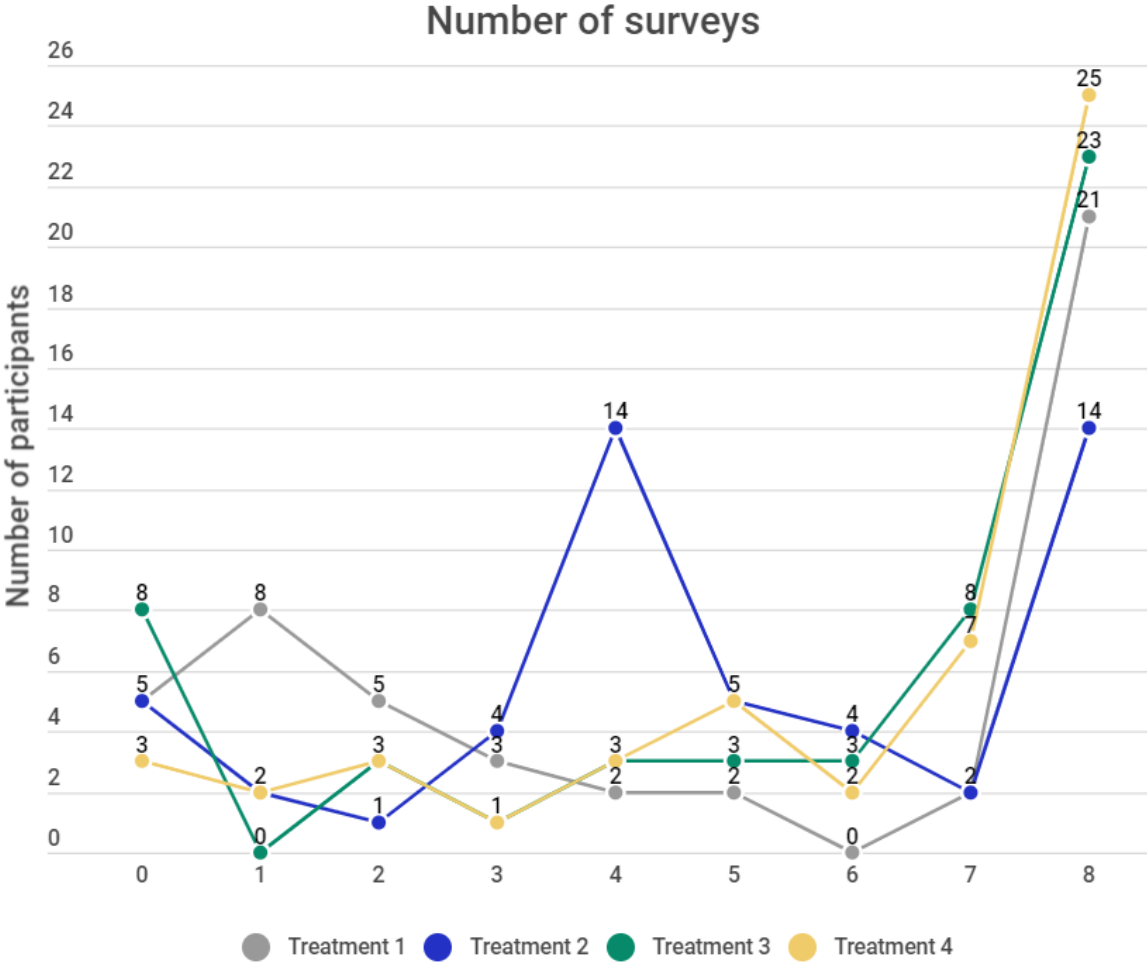




Figure 4: Distribution of the number of surveys chosen per participant



Treatment / Surveys	0	1	2	3	4	5	6	7	8
Treatment 1	5	8	5	3	2	2	0	2	21
Treatment 2	5	2	1	4	14	5	4	2	14
Treatment 3	8	0	3	1	3	3	3	8	23
Treatment 4	3	2	3	1	3	5	2	7	25

**5.2.2 Two-sample Kolmogorov-Smirnov test**

We conducted statistical tests to confirm the insights illustrated in Figures 3 and 4. First, we continued to study the population distribution among treatments, notably

to validate the existence of the boomerang effect in treatment 2. To be able to do this, we performed two-sample Kolmogorov-Smirnov tests. This test is a non-parametric test that compares the cumulative distributions of two data sets. The two-sample Kolmogorov-Smirnov test does not assume that data are sampled from Gaussian distributions (or any other defined distributions) and is very useful to test whether two samples come from the same distribution. Before analyzing the results, we must admit that one of the main drawbacks is the small sample sizes of our treatments. Thus, we consider as significant the p values that are less than 0.1. P values are indicated for all results and are available in the tables.

Table 4 shows the results of the two-sample Kolmogorov-Smirnov (KS) test. In Table 4 a) we found a significant difference in the distributions of treatments 1 and 4 ( $D=0.2610$ ,  $p=0.055$ ). In the same way, we observe in tables 4b and 4c that the distribution of treatment 2 is significantly different compared to both treatment 3 ( $D=0.3104$ ,  $p=0.017$ ) and treatment 4 ( $D=0.3137$ ,  $p=0.012$ ). There are no significant differences between other treatments. Therefore, we interpret that replacing the standard norm with the norm from the top combined with the injunctive norm shifts the distribution of contributions. In addition, the presence of the norm from the top modifies the distribution of contributions compared to the control group.

Table 4: Two-sample Kolmogorov-Smirnov test

Table 4a: Treatment 1 versus Treatment 4			
Smaller group	D	P-value	Exact
Treatment 1	0.2610	0.034	
Treatment 4	0.000	1.000	
Combined K-S	0.2610	0.069	0.055*

Table 4b: Treatment 2 versus Treatment 3			
Smaller group	D	P-value	Exact
Treatment 2	0.3104	0.011	
Treatment 3	-0.0824	0.729	
Combined K-S	0.3104	0.023	0.017**

Table 4c: Treatment 2 versus Treatment 4

Smaller group	D	P-value	Exact
Treatment 2	0.3137	0.007	
Treatment 4	0.000	1.000	
Combined K-S	0.3137	0.013	0.012**

### 5.2.3 Identifying the effects of norms

#### 5.2.3.1 Boomerang effects

We built logit models to identify the source of differences in treatment distributions and confirm insights from the graph. First, we wanted to investigate whether the boomerang effect observed in the graph in treatment 2 is significant. A boomerang effect results in a decrease of the contributions higher than the average behavior (indicated by the feedback) and an increase of contributions towards this average, here four surveys. We created the dependent variable  $Surveys=4$ , which is a dummy variable = 1 if the number of surveys chosen by the participants is four, and = 0 if the number of surveys chosen is strictly higher than four. Therefore, contributions lower than four surveys were not considered in this analysis, leaving 148 observations. This dummy variable isolates contributions to four surveys from higher contributions and enables to study the presence of a boomerang effect in treatment 2. The equation of the model is specified here:

$$P(S = 4) = F(\alpha + \beta_1 Treatment2_i + \beta_2 Treatment3_i + \beta_3 Treatment4_i + \epsilon_i) \quad (1)$$

Where F is the cumulative distribution function of a logistic distribution of parameters 0 and 1. The dependent variable is the dummy variable  $Surveys=4$  defined above. The independent variables are dummy variables that correspond to the different treatments. First of all, note that Treatment 1 is the reference treatment in the model.  $\beta_1 Treatment2_i$  is a dummy variable = 1 for observations of Treatment 2, and = 0 for other treatments.

$\beta_2 Treatment3_i$  is a dummy variable = 1 for observations of Treatment 3, and = 0 for other treatments. Finally,  $\beta_3 Treatment4_i$  is a dummy variable = 1 for observations of Treatment 4, and = 0 for other treatments. These variables are the same for all logit models in the paper.

The results of the logit regression are reported in Table 5. We see that the standard norm significantly increases the contributions equal to four surveys ( $B=1.946$ ,  $SE=0.807$ ,  $p=0.016$ ) compared to the control group. Since the model analyzed only the contributions  $\leq 4$ , these findings mean a decrease in the high contributions generated by the standard norm compared to the control group, thus resulting in a boomerang effect. In addition, it confirms what we have seen in the graph in Figure 3. Thus, the identification of the boomerang effect in Treatment 2 validates the hypothesis H1.

Unsurprisingly, treatments 3 and 4 have no significant effect on the dummy variable  $Surveys = 4$ .

Table 5: Logit regression on the contributions  $\geq 4$  surveys

Variable	Dummy variable: Surveys chosen = 4		
	Coef [Std. Dev]	p-value	Odds ratio [Std. Dev]
Standard norm	1.946[.807]	0.016**	7 [5.650]
Norm from the top	.013[.949]	0.989	1.013[.962]
Norm from the top + injunctive norm	-.039[.948]	0.967	.962[.912]

We built a second logit model to study the presence of boomerang effects when the norm from the top is implemented. We followed the same method and set the dependent variable as a dummy variable named  $Surveys=7$ . This variable takes for value 1 if the number of surveys chosen by the participants is seven, which corresponds to the feedback sent by the norm from the top, and 0 if the number of surveys chosen is strictly higher than seven (i.e., eight). This analysis did not consider contributions lower than seven surveys, leaving 102 observations. This dummy variable isolates contributions to seven surveys from higher contributions and enables to study the presence of a boomerang effect in treatments 3 and 4. The equation of the model is specified here:

$$P(S = 7) = F(\alpha + \beta_1 Treatment2_i + \beta_2 Treatment3_i + \beta_3 Treatment4_i + \epsilon_i) \quad (2)$$

Where F is the cumulative distribution function of a logistic distribution of parameters 0 and 1. The dependent variable is the dummy *Surveys=7* defined earlier. Following the method of the first model, the independent variables are dummy variables that correspond to the different treatments, with Treatment 1 as the reference treatment in the model.

The logit regression results are reported in Table 6 and show no significant effect of Treatment 3 on reducing the contributions higher than the feedback. Thus, no boomerang effect has been generated by the norm from the top alone. The hypothesis H3 is not validated. As expected, Treatment 4 does not produce the boomerang effect. As there is no boomerang effect in Treatment 3, we cannot identify the addition of the injunctive as the element that removes the boomerang effect. The hypothesis H5 is not validated.

Variable	Dummy variable: Surveys chosen = 7		
	Coef [Std. Dev]	p-value	Odds ratio [Std. Dev]
Standard norm	.405 [1.058]	0.702	1.5 [1.587]
Norm from the top	1.295[.846]	0.126	3.652[3.091]
Norm from the top + injunctive norm	1.078[.855]	0.207	2.94[2.513]

### 5.2.3.2 Positive impacts of the norms

Now that we have identified the potential boomerang effects, we are interested in studying whether the norms lead to a decrease in the contributions lower than the average behavior. For instance, the concentration of contributions at four surveys in Treatment 2 could have also been triggered by a reduction of contributions equal to 0, 1, 2, and 3 surveys compared to Treatment 1. To investigate the potential positive impact of the standard norm, we set another dummy variable *Surveys=4'*, which = 1 when the number of surveys corresponds exactly to four and = 0 for contributions lower than four surveys.

Conversely to the first logit model, the contributions higher than four surveys are not included in this model, leaving 76 observations. The equation of the model is written in the same way as the equation (1), although the set of observations of the model is different. As in previous logit models, the independent variables correspond to dummy variables set according to the different treatments. Treatment 1 is the reference treatment in the regression.

Table 7 reveals the results of the logit regression. Treatment 2 has a positive and significant effect ( $B=2.506$ ,  $SE=0.838$ ,  $p=0.003$ ) on the dependent variable  $Surveys = 4'$ , which means that the standard norm tends to decrease the number of low contributions contributions (between 0 and 3 surveys) compared to the control group. Table 7 indicates that there is no effect of Treatment 3 and Treatment 4 on  $Surveys = 4'$ .

Table 7: Logit regression on the contributions  $\leq 4$  surveys

Variable	Dummy variable: Surveys chosen = 4		
	Coef [Std. Dev]	p-value	Odds ratio [Std. Dev]
Standard norm	2.506[.838]	0.003***	12.25 [10.266]
Norm from the top	.965[.982]	0.326	2.625[2.578]
Norm from the top + injunctive norm	1.253[.996]	0.208	3.5[3.486]

We now turn to the analysis of the positive effects of the norm from the top, alone and combined with the injunctive norm. We set the dummy variable  $Surveys=7'$ , with  $Surveys=7' = 1$  when the number of surveys corresponds exactly to seven, and  $= 0$  for contributions lower than seven surveys. Contributions to eight surveys are excluded from the model, leaving 119 observations. As for the previous model, the equation of the model is written in the same way as the equation (2), yet the set of observations of the model is different. Independent variables still correspond to the different treatments.

The results of the model are indicated in Table 8. The table shows a positive and significant effect of Treatment 3 ( $B=1.561$ ,  $SE=0.844$ ,  $p=0.064$ ) and Treatment 4 ( $B=1.527$ ,  $SE=0.858$ ,  $p=0.075$ ) on the contributions equal to seven surveys, at a level 10%. These results indicate that the norm from the top decreases the number of contributions lower than the feedback

indicated to the participants compared to the control group.

Note that the results of this subpart are all robust to the addition of sex and age variables in the model.

Table 8: Logit regression on the contributions  $\leq 7$  surveys

Variable	Dummy variable: Surveys chosen = 7		
	Coef [Std. Dev]	p-value	Odds ratio [Std. Dev]
Standard norm	-.336 [1.033]	0.745	.714 [.738]
Norm from the top	1.561 [.844]	0.064*	4.762 [4.020]
Norm from the top + injunctive norm	1.527 [.858]	0.075*	4.605 [3.950]

### 5.2.3.3 Global effects of the norm

As seen in the previous part, the standard norm leads to a concentration of contributions to the average by reducing lower and higher values. In order to validate these findings, we built a logit model with the dependent variable  $Surveys=4$ , yet this model includes all observations, contrary to the previous models. The equation of the model is similar to (1). Table 14a reports the result of the model (Appendix D). In line with the previous results, the standard norm significantly affected participants' decision towards choosing four surveys ( $B=2.16$ ,  $SE=0.788$ ,  $p=0.006$ ) compared to the control group. Thus, this result illustrates the global impact of the standard norm to concentrate the contributions in four surveys. Other treatments are not significant.

By contrast, the norm from the top (with or without the injunctive norm) reduces only values lower than its feedback. We performed a logit regression similar to the model in Table 14b with  $Surveys=7$  as the dependent variable and found no significant effect of Treatment 3 nor Treatment 4. Therefore, since the effects of the standard norm are mitigated and, in contrast, treatments implementing the norm from the top affect the contributions only positively, hypotheses H2 and H4 seem to be corroborated.

#### 5.2.3.4 Effect of the norm on expanded concentrations

To conclude, we aimed to study whether the concentration of behaviors at the average is still significant when contributions near four are included. Indeed, we can wonder whether the standard norm only encouraged the participants who would have been close to the average (three or five surveys) to adjust their contribution to four surveys. Thus, we tested the effects of treatments on several levels of expansion of the concentrations of contributions. The first level of expansion is represented by the dummy variable  $Surveys=[3, 5] = 1$  for contributions of three, four, or five surveys and  $= 0$  for other contributions. We built a logit model similar to the model used in Table 5. The equation of this second model is specified here:

$$P(S = [3, 5]) = F(\alpha + \beta_1 Treatment2_i + \beta_2 Treatment3_i + \beta_3 Treatment4_i + \epsilon_i) \quad (3)$$

Where  $F$  is the cumulative distribution function of a logistic distribution of parameters 0 and 1. The dependent variable is the dummy  $Surveys=[3, 5]$  with 1 when the number of surveys = 3, 4 or 5 and 0 otherwise. Independent variables correspond to the different treatments.

The results are reported in Table 14b (Appendix D). The logit regression reveals a positive and significant effect of treatment 2 on the dummy variable ( $B=1.571, SE=0.496, p=0.002$ ) compared to the control group. This effect indicates that the impact of the standard norm is still significant when the interval of concentration is expanded to contributions  $\pm 1$  the average contribution. Other treatments have no significant effect.

The second level of the expansion includes contributions from two surveys to six. We built a new logit model, and we set a dummy variable  $Surveys=[2, 6]$ , which  $= 1$  for contributions included between two and six surveys and  $= 0$  otherwise. The equation of this third model is specified here:

$$P(S = [2, 6]) = F(\alpha + \beta_1 Treatment2_i + \beta_2 Treatment3_i + \beta_3 Treatment4_i + \epsilon_i) \quad (4)$$



Where  $F$  is the cumulative distribution function of a logistic distribution of parameters 0 and 1. The dependent variable is the dummy  $Surveys=[2, 6]$  with 1 when the number of surveys = 2, 3, 4, 5 or 6 and 0 otherwise. Independent variables correspond to the different treatments.

The results of the logit regression are reported in Table 14c (Appendix D). Table 14c shows that treatment 2 has a significant and positive effect on the contributions between two and six surveys ( $B=1.295$ ,  $SE=0.436$ ,  $p=0.003$ ), compared to the control group. The impact of the standard norm on the concentration of surveys chosen towards the average is, therefore, still significant when the interval of concentration is expanded to contributions  $\pm 2$  the average contribution. Note that there is no longer a significant result when we extend the interval to contributions  $\pm 3$ .

The logit regressions results reveal that the concentration of contributions in the second treatment was not compensated by a decrease in the contributions near the average. The impact of the standard norm on the contributions is quite global, leading to fewer contributions far from the norm provided, as we preliminary saw in the graphs of Figures 3 and 4.

### 5.3 Sample differences analysis

In this subpart, we seek to analyze the differences between treatments. We performed non-parametric tests since the data are not normally distributed (it might be due to the small sample sizes). First, we used the Kruskal-Wallis test. This test is considered the non-parametric equivalent of the one-way analysis of variance (ANOVA). The Kruskal-Wallis test is a non-parametric method for testing whether samples originate from the same distribution. It is used for comparing two or more independent samples of equal or different sample sizes. Table 9 shows the result of a Kruskal-Wallis test, corrected for tied rank, among the four treatments ( $\chi^2 = 6.711$ ,  $p = .0817$ ,  $df = 3$ ). Following this

result, we can reject the hypothesis that the populations are the same at any level below 10%.

Table 9: Number of surveys: Kruskal-Wallis test

Treatment	Observations	Rank Sum
Treatment 1	48	4522.50
Treatment 2	51	4544.50
Treatment 3	52	5576.00
Treatment 4	51	5860.00

$\chi^2 =$	6.217	probability =	0.1015
$\chi^2$ with ties =	6.711	probability =	0.0817*

Table 10: Treatment 1 versus Treatment 4: Mann-Whitney U test

Treatment	Observations	Rank Sum	Expected
Treatment 1	48	2177	2400
Treatment 4	51	2773	2550
Combined	99	4950	4950

unadjusted variance	20400.00
adjustment for ties	-2113.80
adjusted variance	18286.20

H0: Treatment 1 = Treatment 4

$z = -1.649$

Prob  $|z| = 0.0991^*$

Now that we have examined the differences among all treatments, we aim to determine which treatments these differences are. Thus, we used the Mann-Whitney U test (also called Wilcoxon rank-sum test, or Wilcoxon-Mann-Whitney test). This non-parametric test is based on the null hypothesis that, for randomly selected values X and Y from two populations, the probability of X being greater than Y is equal to the probability of Y

being greater than X. The Mann-Whitney test compares the population of two treatments, contrary to the Kruskal-Wallis test.

Table 10 displays the Mann-Whitney test results regarding treatment 1 and treatment 4. We observe a significant difference between T1 and T4 at a level 10% ( $z=-1.649$ ,  $p=0.0991$ ). Therefore, this result added to the findings in Figure 1 demonstrates that the feedback of treatment 4 (norm from the top + injunctive norm) positively affected the number of surveys chosen compared to the group without feedback.

Table 11: Treatment 2 versus Treatment 3: Mann-Whitney U test

Treatment	Observations	Rank Sum	Expected
Treatment 2	51	2384	2652
Treatment 3	52	2972	2704
Combined	103	5356	5356

unadjusted variance	22984.00
adjustment for ties	-1256.08
adjusted variance	21727.92

H0: Treatment 2 = Treatment 3

$z = -1.818$

Prob  $|z| = 0.0690^*$

In Table 11, we ran a Mann-Whitney test to compare the population of treatment 2 and treatment 3. According to this table, the Mann-Whitney test demonstrates a significant difference between T2 and T3 at a level 10% ( $z=-1.818$ ,  $p=0.069$ ). Therefore, given this result and previous findings, we can conclude that the feedback of T3 (norm from the top) had a positive effect on the number of surveys chosen compared to the standard norm in T2, which validates the hypothesis H2.

Table 12: Treatment 2 versus Treatment 4: Mann-Whitney U test

Treatment	Observations	Rank Sum	Expected
Treatment 2	51	2357.5	2626.5
Treatment 4	51	2995.5	2626.5
Combined	102	5253	5253

unadjusted variance	22325.25
adjustment for ties	-1406.29
adjusted variance	20918.96

H0: Treatment 2 = Treatment 4

$$z = -2.551$$

$$\text{Prob } |z| = 0.0107^{**}$$

Table 12 reports the results of the Mann-Whitney test comparing treatment 2 and treatment 4. The Mann-Whitney test (Table 12) shows a significant difference between T2 and T4 at a level 5% ( $z=-2.551$ ,  $p=0.0107$ ). Thus, given this result and the result of Figure 1, the feedback of T4 (norm from the top + injunctive norm) has a positive effect on the number of surveys chosen compared to the standard feedback in T2. The hypothesis H4 is partially validated. Indeed, T4 is more efficient than T1 and T2, nonetheless, we found no significant difference between T3 and T4.

Also, there are no significant differences between T1 and T2, nor between T1 and T3. The results of non-parametric tests are summarized in the tables below<sup>3</sup>.

Table 13a: Two-sample Kolmogorov-Smirnov test

	Treatment 1	Treatment 2	Treatment 3	Treatment 4
Treatment 1	.			0.055*
Treatment 2		.	0.017**	0.012**
Treatment 3		0.017**	.	
Treatment 4	0.055*	0.012**		.

<sup>3</sup>Note that  $p^* < .10$ ,  $p^{**} < .05$ ,  $p^{***} < .01$

	Treatment 1	Treatment 2	Treatment 3	Treatment 4
Treatment 1	.			0.0991*
Treatment 2		.	0.0690*	0.0107**
Treatment 3		0.0690*	.	
Treatment 4	0.0991*	0.0107**		.

## 5.5 Summarize of the results

The results show that, while the standard norm had no impact on the contributions to prosocial behavior, the combination of the norm from the top and the injunctive norm increased the contributions compared to the control group. The standard norm generated the boomerang effect found in the literature (Schultz et al. 2007), meaning that it reduced the number of contributions above the average behavior. Furthermore, the standard norm led to a concentration of the contributions towards the average behavior. By contrast, the norm from the top (and the combination with the injunctive norm) did not have this issue and conserved an important part of high contributions (even increasing the seven surveys' contributions), while reducing low contributions. These results allow us to argue that the norm from the top applied as a nudge can be more adapted in contexts where a large part of the population does not adopt the desired behavior. In addition, note that the modal answer was eight surveys in each treatment, meaning that the most popular choice was to contribute all the surveys, except in T2. The presence of a completion bias might justify this phenomenon. Completion bias is the tendency of people to feel compelled to complete a task once they have started it. When a person engages in prosocial behavior, she obtains greater satisfaction if she performs it entirely rather than leaving it unfinished. Thus, we can assume that at least some people perceived their behavior as prosocial only if the task was fully completed.

## 6 Discussion

This study reports the results of an online experiment involving prosocial behavior (filling out surveys without monetary incentives) with the implementation of different social norms that act as nudges. In the literature, several studies applied the "standard" social norm as a nudge that incites people to adopt prosocial behavior. However, the implementation of such tools can backfire (i.e., the boomerang effect Schultz et al. 2007) when a sufficient part of the population does not already adopt the behavior. In this study, we implemented new social feedback to extend the contribution of social norms as nudges in attitudes where the standard feedback would not be efficient. This feedback, called "Norm from the top", focuses on the action of the most altruistic people in a given population.

The results of this experiment reveal that the combination of a descriptive norm from the top, and an injunctive norm increases contributions to prosocial behavior, while the standard social norm was inefficient in enhancing the average number of surveys chosen. Indeed, the standard norm, which informs participants about the average behavior, reduces high contributions and generates a concentration of contributions towards this average attitude. By contrast, the norm from the top, alone or associated with an injunctive norm, does not produce a boomerang effect; therefore, we propose that it is more efficient than the standard social norm in this context. Since it provides information about people who succeed in aligning their intentions and their actions, we believe the norm from the top is more suitable than the standard norm to illustrate how a nudge can act as a bridge for a potential discrepancy between intention and action. Consequently, applying social norms built on the behavior of the greatest contributors in a group, can be a solution (as trending norms can be) in situations where standard social norms are not.

Whether the addition of the injunctive norm to the norm from the top has a significant impact compared with the norm from the top alone, has yet to be observed. Further experiments in that direction could indicate whether an injunctive norm is relevant when there is no boomerang effect issue. Also, future work could expand the number

of participants per treatment in order to strengthen the validity of these results. To expand generalizability, replications of the experiment could be run with populations from different countries to observe if the norm from the top produces different effects. In addition, the prosocial behavior we implemented in our study is related to environmental topics. As the environment could play a role in the participants' willingness to contribute, replacing it with other topics could reveal whether the effects of nudges are consistent. In fact, it has been shown that nudges are, in part context-dependent (Hummel and Maedche 2019). Thus, implementing a norm from the top in different prosocial behaviors, particularly in field experiments, would allow for the observation of its impact on real-world behaviors and, therefore, assess the external validity of this new tool. For example, in charitable giving, Kumru and Vesterlund (2010) argued that first soliciting donors with a high social ranking leads to increased donations from other donors. This demonstrate that being able to first ask for the contributions of people with the expected highest prosocial attitudes could be a solution to generating a norm from the top. A final consideration when using the norm from the top is for nudged people to identify themselves as part of the group of top contributors. If nudged people do not feel that they belong to this group, whether socially, economically, or geographically, they may not feel concerned by the norm and hence not modify their attitude. In this sense, further experiments could be conducted to examine this social identity issue.

## Acknowledgment

We are very thankful to the CIRANO Experimental Economics Laboratory that allowed us to recruit participants among their base and took care of the payments.

# Bibliography

Allcott, H. 2011. Social Norms and Energy Conservation. *Journal of Public Economics* 95(9–10): 1082–95.

Allcott, H., Rogers, T. 2014. The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation. *American Economic Review*, 104 (10): 3003-37.

Agerström, J., Carlsson, R., Niclasson, L., Guntell, L. 2016 Using descriptive social norms to increase charitable giving: the power of local norms. *J. Econ. Psychol.* 52, 147 – 153

Asch, S. E. 1951. Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men*. Pittsburgh: Carnegie Press. Pp 177-190.

Bicchieri, C. 2006. *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.

Bicchieri, C. 2016. *Norms in the Wild. How to Diagnose, Measure, and Change Social Norms*. Oxford: Oxford University Press ISBN 978-0190622053.

Chakravarty, S., Roy, J. 2009. Recursive expected utility and the separation of attitudes towards risk and ambiguity: An experimental study. *Theory and Decision*. 66. 199-228. 10.1007/s11238-008-9112-4.

Christakis N., Fowler, J. 2008. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study *BMJ*.

Costa, D.L., Kahn, M.E. 2013. Energy conservation "nudges" and environmentalist ideology: evidence from a randomized electricity field experiment. *Journal of the European Economic Association*, 11: 680–702.

Cialdini R. 2003. Crafting Normative Messages to Protect the Environment. *Current Directions in Psychological Science*. 12(4):105-109.



Cialdini, R., Kallgren, C., Reno, R. 1990. A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior. *Advances in Experimental Social Psychology*.

Demarque, C., Charalambides, L., Hilton, D.J., Waloquier, L. 2015. Nudging sustainable consumption: the use of descriptive norms to promote a minority behavior in a realistic online shopping environment. *J. Environ. Psychol.* 43, 166–174.

Deutsch, M., Gerard, H.B. 1955. A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology* 51(3), 629–636.

Frey, B.S., Meier, S. 2004. Social Comparisons and Pro-social Behavior: Testing "Conditional Cooperation" in a Field Experiment. *American Economic Review*, 94 (5): 1717-1722.

Glazier, A., Mone T. 2019. Success of Opt-In Organ Donation Policy in the United States. *JAMA*.;322(8):719–720.

Glaeser, E., Sacerdote, B., Scheinkman, J. 1996. Crime and Social Interactions, *The Quarterly Journal of Economics*, 111, issue 2, p. 507-548

Goldstein, N.J., Cialdini, R.B., Griskevicius, V. 2008. A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels, *Journal of Consumer Research*, Volume 35, Issue 3, October 2008, Pages 472–482.

Gonçalves, D., Coelho, P., Martinez, L.F., Monteiro, P. Nudging Consumers toward Healthier Food Choices: A Field Study on the Effect of Social Norms. *Sustainability* 2021, 13, 1660.

Habib, R., White, K., Hoegg, J. 2021. Everybody Thinks We Should but Nobody Does: How Combined Injunctive and Descriptive Norms Motivate Organ Donor Registration. *Journal of Consumer Psychology*. Volume 31, July 2021, Pages 621-630.

Holt, C.A., Laury, S.K. 2002. Risk Aversion and Incentive Effects. *American Economic Review*, 92(5): 1644–55.

Hummel, D., Maedche, A. 2019. How Effective Is Nudging? A Quantitative Review

on the Effect Sizes and Limits of Empirical Nudging Studies. *Journal of Behavioral and Experimental Economics*. 80. 10.1016.

Kalnikaitė, V., Rogers, Y., Bird, J., Villar, N., Bachour, K., Payne, S., Todd, P. M., Schöning, J., Krüger, A., Kreitmayer, S. 2011. How to nudge in situ: Designing lambent devices to deliver salience information in supermarkets. *Paper presented at 13th ACM International Conference on Ubiquitous Computer, Beijing, China*.

Kurz, V. 2018. Nudging to reduce meat consumption: Immediate and persistent effects of an intervention at a university restaurant. *Journal of Environmental Economics and Management*. Volume 90, Pages 317-341.

Kumru, C., Vesterlund, L. 2010. The Effect of Status on Charitable Giving, *Journal of Public Economic Theory*, 12, (4), 709-735.

Lévy-Garboua, L., Masclet, D., Montmarquette, C. 2009. A behavioral Laffer curve: Emergence of a social norm of fairness in a real effort experiment. *Journal of Economic Psychology*, Volume 30, Issue 2.

Loschelder, D.D., Siepelmeyer, H., Fischer, D., Rubel, J.A. 2019. Dynamic Norms Drive Sustainable Consumption: Norm-based Nudging Helps Café Customers to Avoid Disposable To-Go-Cups, *Journal of Economic Psychology*,

Momsen, K., Stoerk, T. 2014. From intention to action: Can nudges help consumers to choose renewable energy? *Energy Policy, Elsevier* vol. 74(C), pages 376-382.

Mortensen, C., Neel, R., Cialdini, R., Jaeger, C., Jacobson, R., Ringel, M. 2017. Trending Norms: A Lever for Encouraging Behaviors Performed by the Minority. *Social Psychological and Personality Science*.

Reno, R., Cialdini, R., Kallgren, C. 1993. The Transsituational Influence of Social Norms. *Journal of Personality and Social Psychology*. 64. 104-112.

Richter, I., Thøgersen, J., Klöckner, C.A. 2018. A Social Norms Intervention Going Wrong: Boomerang Effects from Descriptive Norms Information. *Sustainability* 2018.

Schultz, PW., Nolan, JM., Cialdini, RB., Goldstein, NJ., Griskevicius V. 2007. The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science*. 2007;18(5):429-434.

Sparkman, G., Waltman, G. M. 2017. Dynamic norms promote sustainable behavior, even if it is counternormative. *Psychological Science*, 28, 1663–1674.

Thaler, RH., Benartzi, S. 2004. Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving. *Journal of Political Economy*, 112(S1), S164-S187.

Thaler, R. H., Johnson, E. J. 1990. Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice. *Management Science*, 36(6), 643–660.

Thaler, RH. and Sunstein, CR. 2009. Nudge: Improving Decisions about Health, Wealth, and Happiness. *Penguin Books*.

Zizzo, D.J. 2010. Experimenter demand effects in economic experiments. *Exp Econ* 13, 75–98 .

# Appendix:

## Appendix A:

Pour chacune des 10 situations suivantes (situation = ligne), choisissez l'urne (A ou B) dans laquelle vous souhaitez qu'on tire aléatoirement une bille.  
Attention, le nombre figurant sur la bille indique votre gain en \$.

	Urne A	Je choisis	Urne B
1	4 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$	<input type="radio"/> <input type="radio"/>	0 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$
2	4 \$ 4 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$	<input type="radio"/> <input type="radio"/>	0 \$ 0 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$
3	4 \$ 4 \$ 4 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$	<input type="radio"/> <input type="radio"/>	0 \$ 0 \$ 0 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$
4	4 \$ 4 \$ 4 \$ 4 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$	<input type="radio"/> <input type="radio"/>	0 \$ 0 \$ 0 \$ 0 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$
5	4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 6 \$ 6 \$ 6 \$ 6 \$ 6 \$	<input type="radio"/> <input type="radio"/>	0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 10 \$ 10 \$ 10 \$ 10 \$ 10 \$
6	4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 6 \$ 6 \$ 6 \$ 6 \$	<input type="radio"/> <input type="radio"/>	0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 10 \$ 10 \$ 10 \$ 10 \$
7	4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 6 \$ 6 \$ 6 \$	<input type="radio"/> <input type="radio"/>	0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 10 \$ 10 \$ 10 \$
8	4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 6 \$ 6 \$	<input type="radio"/> <input type="radio"/>	0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 10 \$ 10 \$
9	4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 6 \$	<input type="radio"/> <input type="radio"/>	0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 10 \$
10	4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$ 4 \$	<input type="radio"/> <input type="radio"/>	0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$ 0 \$

Valider

Figure 5: Lottery, Grid 1


Translation: "For each of the following 10 situations (situation = line), choose the urn (A or B) from which you want a ball to be randomly drawn. Note that the number on the ball indicates your gain in \$."

## Appendix B:

**Fin de l'expérimentation**

Questionnaires additionnels

Combien de questionnaires souhaitez-vous remplir ?

0  8

Cliquez sur la barre pour faire apparaître le curseur

Figure 6: Treatment 1: no information regarding others' people behavior


Translation: "How many surveys do you want to complete ?"

**Fin de l'expérimentation**


Questionnaires additionnels

Lors des sessions précédentes, les participants ont rempli en moyenne chacun 4 questionnaires.

4

0  8

Combien de questionnaires souhaitez-vous remplir ?

0  8

Cliquez sur la barre pour faire apparaître le curseur

Figure 7: Treatment 2: standard social norm nudge

Translation: "During previous sessions, participants completed an average of four questionnaires.

How many surveys do you want to complete ?"

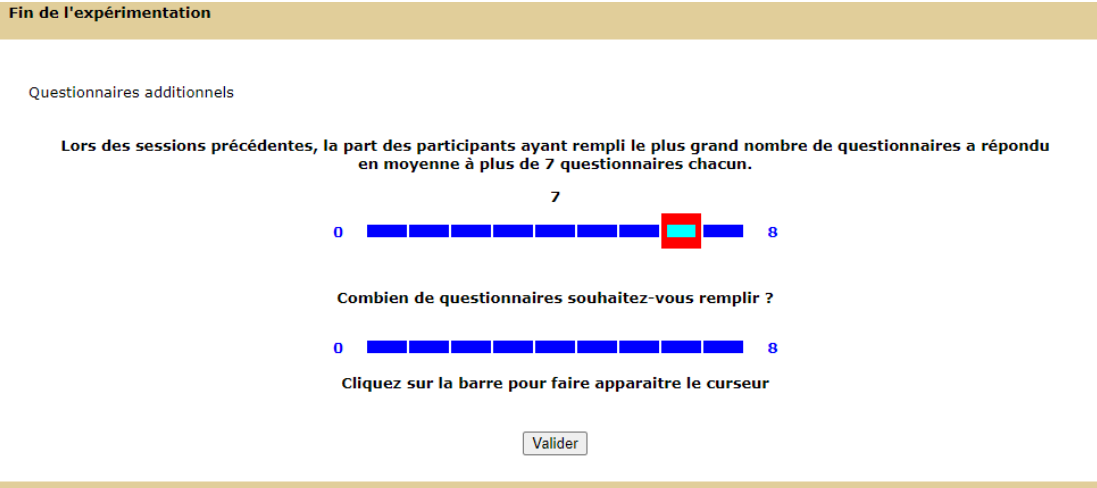


Figure 8: Treatment 3: norm from the top

Translation: "During previous sessions, the share of participants who completed the most questionnaires completed on average over seven questionnaires each.  
How many surveys do you want to complete ?"



Figure 9: Treatment 4: norm from the top and injunctive norm

Translation: "During previous sessions, the most devoted participants completed on average over seven questionnaires each.  
How many surveys do you want to complete ?"

## Appendix C:

The authentic message received by the participants asking to fill out additional surveys (in french):

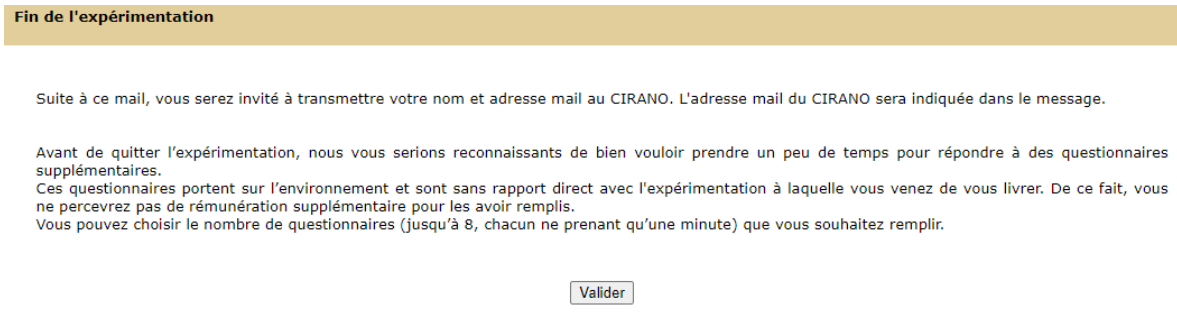


Figure 9: Message introducing prosocial behavior

Translation:

"Before finishing the experiment, we would be grateful if you could take some time to fill out additional surveys.

These surveys are related to the environment and are not linked with the experiment you have just carried out.

Therefore, you will not receive any reward for filling it out.

You can choose the number of surveys (up to eight, one minute each) you wish to fill out."

## Appendix D:

Table 14a below reports the logit regression results investigating the relation between the treatments and contributions equal to four surveys. The dependent variable is  $Surveys=4$ , with = 1 when the number of surveys chosen is precisely four and = 0 for all other numbers of surveys.

Variable	Dummy variable: Surveys chosen = 4		
	Coef [Std. Dev]	p-value	Odds ratio [Std. Dev]
Standard norm	2.163 [.788]	0.006***	8.703 [6.854]
Norm from the top	.342[.936]	0.715	1.408[1.318]
Norm from the top + injunctive norm	.363[.936]	0.698	1.438[1.345]

Tables below report the results of the logit regressions investigating the relation between the treatments and a set of values of the numbers of questionnaires around the mean. Those sets of values represent expanded concentrations of the contributions towards the average behavior, which is four surveys. Table 14b reports the results of the logit regression with a dependent variable = 1 if the number of surveys is between to three and five and = 0 for other contributions. Table 14c reports the results of the logit regression with a dependent variable = 1 if the number of surveys is between two and six and = 0 for other contributions.

Variable	Dummy variable: Surveys chosen $\in [3, 5]$		
	Coef [Std. Dev]	p-value	Odds ratio [Std. Dev]
Standard norm	1.571 [.496]	0.002***	4.811[2.388]
Norm from the top	-.093[.576]	0.872	.911[.525]
Norm from the top + injunctive norm	.227[.550]	0.671	1.255[.690]



Table 14c: Logit regression on surveys chosen

Variable	Dummy variable: Surveys chosen $\in [2, 6]$		
	Coef [Std. Dev]	p-value	Odds ratio [Std. Dev]
Standard norm	1.295 [.436]	0.003***	3.652[1.593]
Norm from the top	1.14e-16[.462]	1.000	1[.462]
Norm from the top + injunctive norm	0.127[.458]	0.782	1.135[.520]

$p^* < .10, p^{**} < .05, p^{***} < .01$