



Contents lists available at ScienceDirect

# Food Quality and Preference

journal homepage: [www.elsevier.com/locate/foodqual](http://www.elsevier.com/locate/foodqual)

## Immediate backfire? Nudging sustainable food choices and psychological reactance

Sanchayan Banerjee<sup>a,d,\*</sup>, Matteo M. Galizzi<sup>c</sup>, Peter John<sup>b</sup>, Susana Mourato<sup>d</sup><sup>a</sup> Department of Environmental Economics, Institute for Environmental Studies, Vrije Universiteit Amsterdam, The Netherlands<sup>b</sup> Department of Political Economy, King's College London, United Kingdom<sup>c</sup> Department of Psychological and Behavioural Science, London School of Economics and Political Science, United Kingdom<sup>d</sup> Department of Geography and Environment, London School of Economics and Political Science, United Kingdom

### ARTICLE INFO

#### Keywords:

Think  
Nudge  
Nudge+  
Reflection  
Backfire

### ABSTRACT

Nudging people towards sustainable diets can help mitigate agricultural emissions. Recent debate suggests “nudges” can have heterogeneous treatment effects in the population, including some backfire effects. In this paper, we present experimental evidence on backfiring effects to a nudge, after people pledge for sustainable food choices. The backfiring effect is driven by people’s short-term intention to eat sustainably. Specifically, we compare a purely reflective tool, the “think”, versus a hybrid “nudge-think” tool, “nudge+”, in a group of 611 participants in the United Kingdom. While the think prompted people to first reflect on a green pledge and then choose an appropriate nudge to comply with their pledge, the nudge+ altered this think by proposing a green default to people who took the pledge. In both experimental conditions, participants self-reported their short- and long-term intentions to eat sustainably. We do not find any evidence of heterogeneous treatment effects by unconditional intentions of participants. Yet nudging people, who already intend to eat sustainably in the short-term, after they have taken the pledge, leads to a backfire. This fades out when we control for people’s overall longer-term intentions. Our results suggest that policymakers should not be deterred by initial reactance to behavioural policies if such effects are driven by people’s short-term intentions.

### 1. Introduction

Increasingly “nudges” (that is, changes in the decision context, or, more precisely, changes in “the choice architecture that alter people’s behaviour in predictable way[s]”: Thaler & Sunstein, 2009, p. 8) are proposed as policy solutions to shift people’s food choices (Bergeron, Doyon, Saulais, & Labrecque, 2019; Vecchio & Cavallo, 2019). There is growing evidence that “green” nudges also help people adopt sustainable diets, both in online settings (Banerjee & Picard, 2023; Banerjee, Galizzi, John, & Mourato, 2022a; Prusaczyk, Earle, & Hodson, 2021) and field experiments (Gravert & Kurz, 2021; Kurz, 2018; Lohmann, Gsottbauer, Doherty, & Kontoleon, 2022). However, recent findings suggest some nudges can have heterogeneous treatment effects on people, including some backfiring effects (Banerjee, Hunter, John, Koenig, Lee-Whiting, Loewen, & Savani, 2022b; Maier et al., 2022; Mertens, Herberz, Hahnel, & Brosch, 2022). These concerns point towards the limited scalability of nudges in delivering systemic changes in the population (Chater & Loewenstein, 2022). To understand such

heterogeneity is an important first step in designing better nudges for the environment and elsewhere (Beshears & Kosowsky, 2020; Szasz, Palinkas, Palfi, Szollosi, & Aczel, 2018), especially for dietary behaviours (Bauer, Bietz, Rauber, & Reisch, 2021). Among the many possible sources of individual heterogeneity, intentions are among the most natural candidates: for example, according to the Transtheoretical Model of Change (TTM: Di Clemente et al., 1991; Prochaska, Wright, & Velicer, 2008) people may need to change their attitudes toward the behaviour, and their intentions to change it, prior to actually changing their behaviour (typically in what the TTM calls the ‘contemplation’ and the ‘preparation’ stages of behavioural change). In this paper, we present experimental evidence from a pre-registered, online randomised controlled trial that suggests that nudging respondents who have short-term intentions to eat sustainably, after they are made to think about a green pledge, can lead to backfires versus leaving people on their own to choose an appropriate way to comply with that pledge. But when we control for people’s long-term intentions, these backfire effects attenuate. Immediate backfire effects resonate with notions of psychological

\* Corresponding author at: Department of Environmental Economics, Institute for Environmental Studies, Vrije Universiteit Amsterdam, The Netherlands.  
E-mail address: [S.Banerjee@vu.nl](mailto:S.Banerjee@vu.nl) (S. Banerjee).

<https://doi.org/10.1016/j.foodqual.2023.104923>

Received 2 November 2022; Received in revised form 10 March 2023; Accepted 15 June 2023

Available online 24 June 2023

0950-3293/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

reactance where nudges can demotivate those who are already motivated, but as our results indicate, policy makers should account for people's long-term intentions before deciding against nudges.

There are many reasons for psychological reactance to occur. For example, some find it harder than others to self-regulate (Baumeister, Leith, Muraven, & Bratslavsky, 2002). When faced with motivational conflicts, citizens often under- or mis-regulate their behaviour (Baumeister & Heatherton, 1996) given their underlying biases such as lacking sufficient attention or self-control, or due to temporal inconsistencies in preferences. Here, prompts like persuasive messages and nudges can help some individuals to adopt sustainable diets. However, the same cannot be said for those who already intend to eat sustainably, for they may react to these prompts negatively (Sunstein, 2017b). Psychological reactance can also occur when citizens perceive a policy intervention to be liberty threatening. Humans usually want to be in control (Seligman, 1975); and often display observable actions to validate this, with hostility towards interventions that reduce their freedom to choose (Langer & Rodin, 1976).

Psychological reactance can manifest itself by backfiring (Osman et al., 2020), with citizens increasingly engaging in behaviours that they were specifically warned against. When prompted to act, underlying motives get blurred, inducing a negative reinforcement of such prompts in the long run; for instance, rewards or penalties can be counterproductive in the long-run (Bénabou & Tirole, 2003, 2006). Motivational crowding out through extrinsic monetary incentives have been thoroughly investigated in the spirit of Titmuss (1970): paying people to engage in pro-social behaviour deters altruistic motives (Ellingsen & Johannesson, 2008; Lacetera, Macis, & Slonim, 2012), with a reluctance to act when paid to do so. Motivational crowding out has been studied also in the context of pro-environmental behaviours, with mixed findings (e.g. D'Adda, 2011; Travers, Clements, Keane, & Milner-Gulland, 2011; Rode, Gómez-Baggethun, & Krause, 2015). Similarly, persuasion can lead to reactance when citizens reject welfare improving directives and prescriptions (Rosenberg & Siegel, 2018). For the same type of reasoning, nudges too can be subject to psychological reactance.

In this paper, we contribute to this growing debate on psychological reactance and backfiring effects by presenting and discussing findings from a pre-registered, online randomised experiment comparing a "think" and a "nudge+" behavioural interventions. A "think" policy is an intervention where participants are explicitly asked to reflect and deliberate on a sustainable eating pledge. A "nudge+" policy is a hybrid intervention combining a "nudge" and a "think": in particular, a nudge setting the default option in a decision context is used in combination with a pledge. In reducing the intended choice of carbon intensive meals, we find that one's self-reported intentions do not unconditionally moderate treatment effects of the nudge in the nudge+ intervention. Nonetheless, when the participants who already intend to eat sustainably reflect and take up on the offer of the pledge, their short-term intentions produce reactance to the nudge. We find that these backfiring effects increase with participants' intentions subject to accepting the pledge. However, such backfiring effects attenuate when we control for people's longer-term intentions conditional on their reflection on the pledge. This is also true when we control for both short- and longer-term intentions, conditional on taking the pledge. Thus, backfiring effects fade soon, as respondents are made to think about their choices preceding the nudge. Our results offer policy insights to researchers and policymakers in support of guiding citizen's sustainable food choices, by enabling them to take control of their decisions, even though there are backfires in the immediate short-term.

In the rest of the paper, we first present the experimental design by discussing the broader experimental set-up and the two interventions, the "think" and the "nudge+". Then, we outline our testable hypothesis and empirical strategy, following which we present our main experimental findings on these moderation effects. We discuss our findings and contextualise them in the broader literature on behavioural science and food policy, before concluding with our research limitations and future

directions.

## 2. Experimental design

We administered a pre-registered (available online here), incentive-compatible survey experiment to 3,074 residents in the United Kingdom in November 2020 to test the effectiveness of behavioural climate policies, broadly classified as "nudges", "thinks", or and "nudge+", in reducing intended orders over carbon-intensive foods. Respondents, recruited via Prolific, were randomly assigned to ten different experimental conditions, including a control condition. In each treatment arm, respondents were presented with food menus listing main course items. Consequently, they were tasked to choose an item on the menu to elicit their dietary preferences. Their choices had real consequences in that they had a chance to be rewarded with a £15 voucher that could be redeemed at numerous major high-street chains of restaurants in the UK to purchase their preferred food item. The menus presented to the respondents implemented the various behavioural interventions. An extensive analysis of this study is provided by Banerjee et al., (2022) and all treatments are listed in Table A.1. The full survey is available in the Online Appendix A7.

In this paper, we use the same experimental dataset to present our analysis on backfire effects pertaining to a subsample of 611 respondents randomly assigned to either a think or a nudge+ intervention from the broader experiment. The treatments and their design are explained below.

In the think intervention arm, respondents were offered a pledge to commit to a sustainable diet. The pledge read as follows:

Dear Participant,

To reduce the impact on the environment, one can consume an environmentally sustainable diet. An environmentally sustainable diet is one with a low environmental impact. Sustainable food items have low carbon emissions associated with their production and consumption. You can contribute to sustainability by pledging to choose an environmentally sustainable diet in order to reduce your carbon footprint. Please indicate if you would like to pledge towards this cause.

Thank you for your cooperation.

Following the offer of the pledge, respondents had to reflect if they wanted to either opt in or out of the pledge. They were also given the option to remain indifferent to the pledge. If they opted out, respondents were shown a baseline (control) menu of 36 main course items. Contrarily, those who pledged (or who were indifferent) were first asked to report<sup>1</sup> their intentions to consume sustainable diets, either for the next meal or over the next week. This allowed us to measure participants' self-reported intentions to eat sustainably, both for the short- and the longer-term. Following this, participants who opted in for the pledge were made to reflect further about the best way to comply with the pledge. They could place their order either from a restricted menu with environmentally sustainable items only; or from a regular menu which was colour labelled with traffic lights to indicate environmental sustainability of the different items. Respondents could also ask for a regular menu with no colour coding. Set out this way, the think intervention harnessed complete deliberation preceding food choices (in line with the general idea that there is a slow, deliberate, analytical reasoning,

<sup>1</sup> We used two questions to elicit these intentions. Specifically, participants were asked to "please indicate your willingness in favour of the following statements using the scale below" for the two statements: "I intend to consume an environmentally sustainable diet in my next meal" (short-term intention); and "I intend to reduce my consumption of environmentally sustainable food over the next week" (longer-term intention). Participants were given a choice of five options: definitely yes, probably yes, might or might not, probably not, and definitely not.

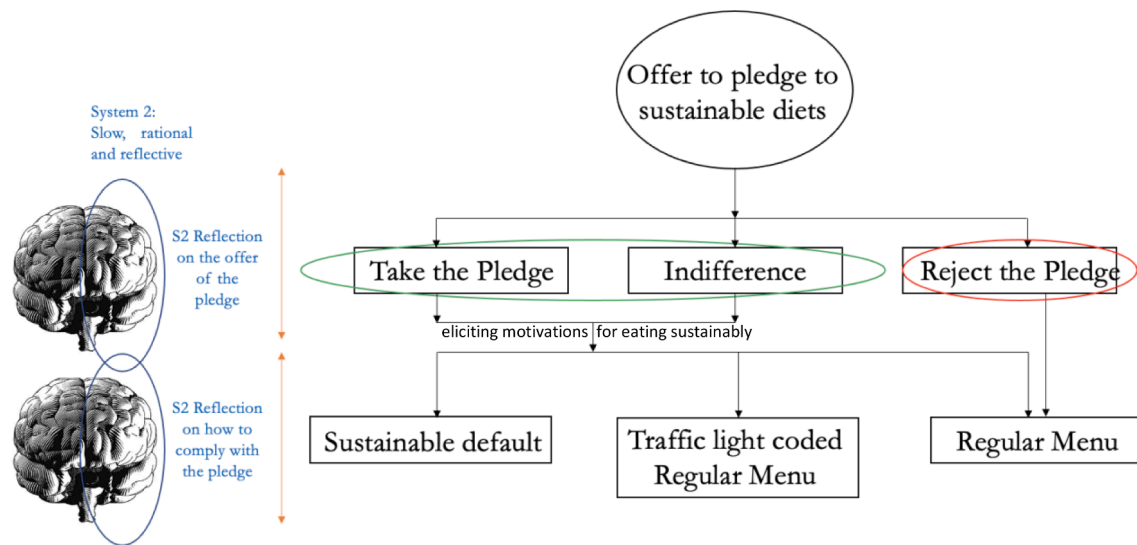


Fig. 1a. How did the think work?

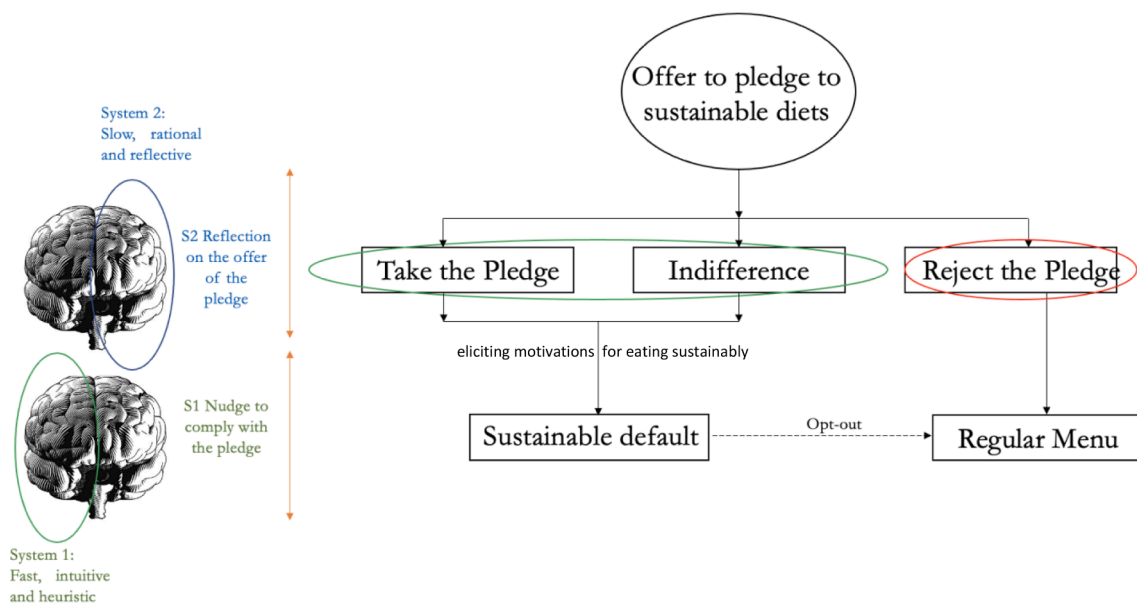


Fig. 1b. How did the nudge+ work?

“system 2”, operating alongside a fast, intuitive, and heuristic reasoning, “system 1”: Evans, 1984; Slovic, 1996; Kahneman, 2003). Enabling respondents to think about the pledge and their decisions to comply with the pledge embodied an active choice mechanism. Nonetheless, regardless of their commitment through the pledge, respondents could eventually order an item that was not environmentally sustainable, as their final choice set was unaltered. Fig. 1a is a graphical illustration of the proposed cognitive mechanism of the think intervention design.

In the nudge+ treatment arm, the intervention design adapted a hybrid nudge-think set-up as outlined by Banerjee and John (2021). Since thinks are cognitively demanding and difficult to scale-up (John et al., 2019; John, 2011), our nudge+ design borrowed an initial reflective element from the think, when people had to think about their decision to pledge. While relying on the nudge next to offer some cognitive easing in the process. As such, it was designed to combine slow and fast thinking. To transform the think into a nudge+, respondents were first offered the pledge and were tasked to opt in or out of the pledge (or remain indifferent to it). Nonetheless, afterwards, they were

spared the cognitive hassle of deciding on the best way to comply with it by automatically defaulting all pledge takers (including those that were indifferent) into a restricted menu with environmentally sustainable items only. Similar to the think intervention, participants who took the pledge (or were indifferent to it) were also asked to self-report their intentions to eat sustainably before being defaulted into the sustainable menu. Combined this way, the nudge (opt-out default sustainable menu) followed the plus (pledge). Fig. 1b is a graphical illustration of the proposed cognitive mechanism of this nudge+ design.

### 3. Empirical methods

#### 3.1. Variables

Our main outcome variable is a continuous measure of greenhouse gas emissions corresponding to the main ingredient in the chosen food item (for more details on construction, see Banerjee et al., 2022). In particular, we measured the life cycle emissions of the main ingredient.

**Table 1**  
List of moderators used in different model specifications reported.

Model Specification	List of Moderators
Model 1	Short Term Intentions conditional on accepting or being indifferent to pledge
Model 2	Longer Term Intentions conditional on accepting or being indifferent to pledge
Model 3	Short- and Longer-Term Intentions conditional on accepting or being indifferent to pledge
Model 4	Short Term Intentions unconditional of decision on the offer of the pledge
Model 5	Longer Term Intentions unconditional of decision on the offer of the pledge
Model 6	Short- and Longer-Term Intentions unconditional of decision on the offer of the pledge

To do this, we first identified the primary food type and ingredient of each dish on our menu using the McCance and Widdowson’s Composition of Foods Integrated Dataset (CoFID), following which a carbon

the pledge are also measured in this sample as if they had no intentions. These additional results are reported in Table A.4 in the Appendix. The measure of “unconditional intention” is summarised below:

$$\text{Unconditional intention} = \begin{cases} 0, & \text{if respondent reported no intention to eat sustainably} \\ 1, & \text{if respondent reported unsure about eating sustainably} \\ 2, & \text{if respondent reported intention to eat sustainably} \end{cases}$$

score (in kgCO<sub>2</sub>e) was assigned using the UK Greenhouse gas emissions scale developed by Scarborough et al., (2014).

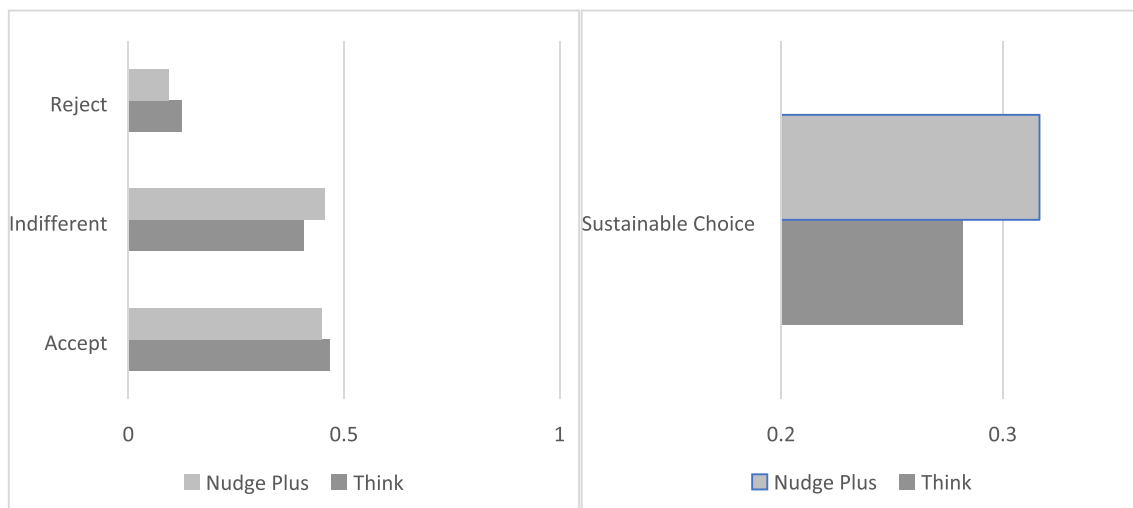
As discussed in the previous section, prior to selecting their preferred food item, respondents who accepted the pledge or were indifferent to it self-reported their intentions to consume sustainable diets, for in the short- and longer-term. Specifically, using a 5-point Likert scale (see footnote 1), respondents self-reported their intentions to consume an environmentally sustainable diet either in their next meal (short-term intention) or over the next week (longer-term intention). Thus, while the pledge was designed without any temporal dimension, the intentions were recorded for respondents’ present and future selves. We use these self-reported intentions to construct our moderating variables.

First, we construct measures of short- and longer-term intentions, unconditional on whether or not the respondent took the pledge in the previous step. We measure this by clubbing respondents’ choices on the 5-point Likert scale as yes (“definitely yes”, “probably yes”), no (“definitely no”, “probably no”), and do not know (“might or might not”). Since respondents were asked to self-report their intentions *if and only if* they took the pledge, our measure of “unconditional intention” has a total of 539 complete responses – 65 respondents did not take the pledge. Later, we include robustness checks where those who rejected

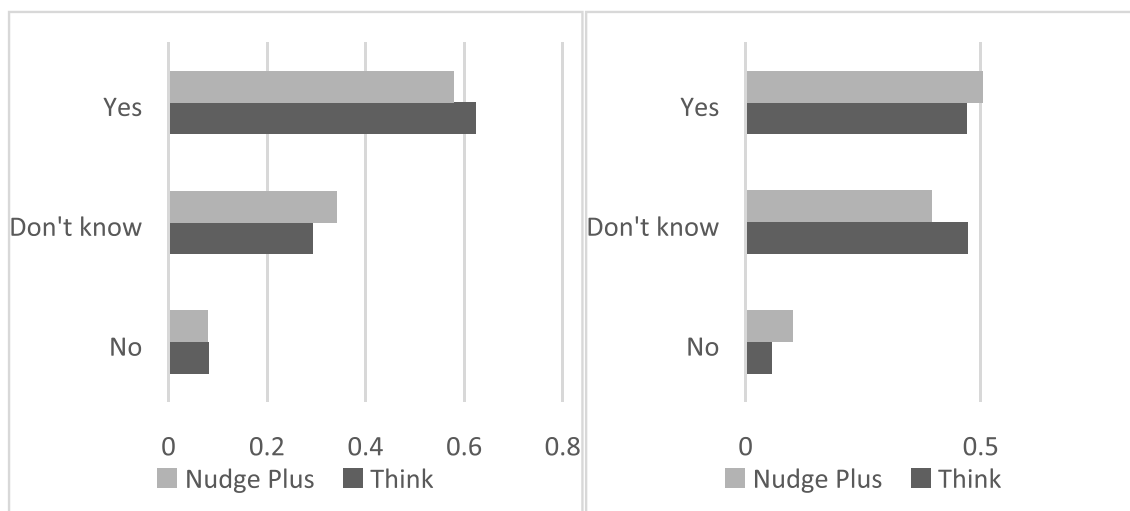
Second, we also construct conditional measures of short- and longer-term intention, where these intentions are measured conditionally on the respondent’s decisions about the pledge. This variable of “conditional motivation” was measured across four levels, namely “0” if the respondent rejected the pledge, “1” if the respondent did not reject the pledge and had no intention to eat sustainably, “2” if the respondent did not reject the pledge and was unsure about their intentions, and “3” if the respondent did not reject the pledge and had intentions to eat sustainably. Our measure of “conditional intention” includes all 604 respondents in the sample. These moderators are further outlined in Table 1 in the next section. The measure of “conditional intention” is summarised below:

$$\text{Conditional intention} = \begin{cases} 0, & \text{reject pledge} \\ 1, & \text{did not reject pledge and has no intentions} \\ 2, & \text{did not reject pledge but unsure intentions} \\ 3, & \text{did not reject pledge and has intentions} \end{cases}$$

Finally, we use a dummy variable “Treatment” to indicate if the respondent was assigned to the think (measured as “0” and considered as the baseline) or the nudge+ (measured as “1”) experimental condition.



**Fig. 2.** (L) Proportion of respondents who accept the pledge in think v/s nudge+ (N<sub>think</sub> = 303; N<sub>nudge+</sub> = 301); (R): Proportion of pledgers who follow through with their pledge (N<sub>think | accept pledge</sub> = 142; N<sub>think | indifferent pledge</sub> = 123; N<sub>nudge+ | accept pledge</sub> = 136; N<sub>nudge+ | indifferent pledge</sub> = 138).



**Fig. 3.** Intentions of respondents in think ( $N_{think} \mid$  did not reject pledge = 265) and nudge+ ( $N_{nudge+} \mid$  did not reject pledge = 274) to consume a sustainable diet in their next meal (L) or over the week (R).

**Table 2**  
Regression results with conditional moderation effects of intentions.

Outcome: GHG emissions	OLS Model 1	OLS Model 2	OLS Model 3
Treatment	Baseline: Think		
Nudge+	-8.619** (2.789)	-7.477** (2.865)	-8.570** (2.941)
Short Term Intentions conditional on Pledge	-5.841*** (0.818)		-5.482*** (1.404)
Longer Term Intentions conditional on Pledge		-4.681*** (0.790)	-0.422 (1.335)
Moral spillovers (s = 1)	-0.362 (1.207)	-1.422 (1.242)	-0.488 (1.245)
Short Term Intentions conditional on Pledge #	Baseline: Think		
Nudge+	2.596* (1.186)		2.563 (1.967)
Longer Term Intentions conditional on Pledge#	Baseline: Think		
Nudge+		1.933 (1.168)	0.011 (1.904)
Constant	21.541*** (2.075)	20.304*** (2.180)	21.802*** (2.199)
N	604	604	604

Table notes: Columns 1–3 present linear regression models corresponding to different moderating variables, namely short-term conditional intention (“OLS Model 1”), longer-term conditional intention (“OLS Model 2”), and both short- and longer-term conditional intentions (“OLS Model 3”). Standard errors in parentheses. Legend: \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ . Corrected p-values in Table A.5.

**3.2. Estimation strategy**

We use a moderation-of-process design to look at the role of intentions as possible individual moderators through which the experimental interventions could affect the outcome variable. We directly measure intentions, as moderators, via a questionnaire, before the experimental conditions. We then use linear regression models to test if intentions, conditional or unconditional on the uptake of the pledge, moderate the average treatment effect of our experimental treatment, namely the nudge+. This is an adapted version of our pre-registered hypothesis 4a as available online in our pre-analysis plan here. More precisely, we use the following regression specification:

$$(GreenhouseGasEmissions)_i = \alpha + \beta(Treatment)_i + \delta(Moderator)_i + \gamma(Treatment * Moderator)_i + \theta.1(s = 1) + \epsilon_i$$

The list of moderators used in the different model specifications are listed below in Table 1.

We adjust for multiple hypotheses testing using a Benjamini-Hochberg p-values correction method for false discovery rates at 10 percent level of significance (available in Appendix A5).

**4. Results**

Our sample is composed of young adults ( $\mu = 28.8$  years,  $\delta = 10.4$  years, min = 18; max = 74) and 45 percent females. More than half the sample is employed in either full- or part-time employment and follow Christianity as their main religion. Our sample is predominantly white in ethnic background (75 percent). More than three-quarters of the sample respondents in these experimental conditions reported consuming balanced and/or flexitarian diets, with <10 percent reporting specific forms of vegetarian diets (such as “lacto-ovo”, “lacto”, “ovo”). Respondents were randomly assigned to the think and nudge+ treatments (for details on randomisation checks, see Table A.2). Across both conditions, fish and poultry were the most frequently chosen food items. The final food choices by respondents across the think and nudge+ experimental conditions are outlined in Table A.6 in the Appendix. Compared to the control condition in the broader experimental study (Banerjee et al., 2022), participants allocated to the think and the nudge+ conditions consumed meals with mean reductions in number of carbon-equivalent emissions of -14.505 (1.667) and -17.905 (1.669), respectively, both differences being highly statistically significant.

Based on simple descriptive statistics, while a greater proportion of respondents seem to have taken up the pledge in the think relative to the nudge+ condition, a smaller proportion was seen to follow through with it, as indicated by their final choice of a food item. However, these sample differences are not statistically significant ( $p = 0.241$ ). Furthermore, the engagement time in both these treatment categories did not differ significantly (k-Wallis  $\chi^2 = 0.049$ ; p-value = 0.8250). Fig. 2 (L) shows the proportion of people who accepted the offer of the pledge; and 2(R) shows the proportion of those who complied with the pledge conditional on accepting (or indifference to) it. In the final stage,



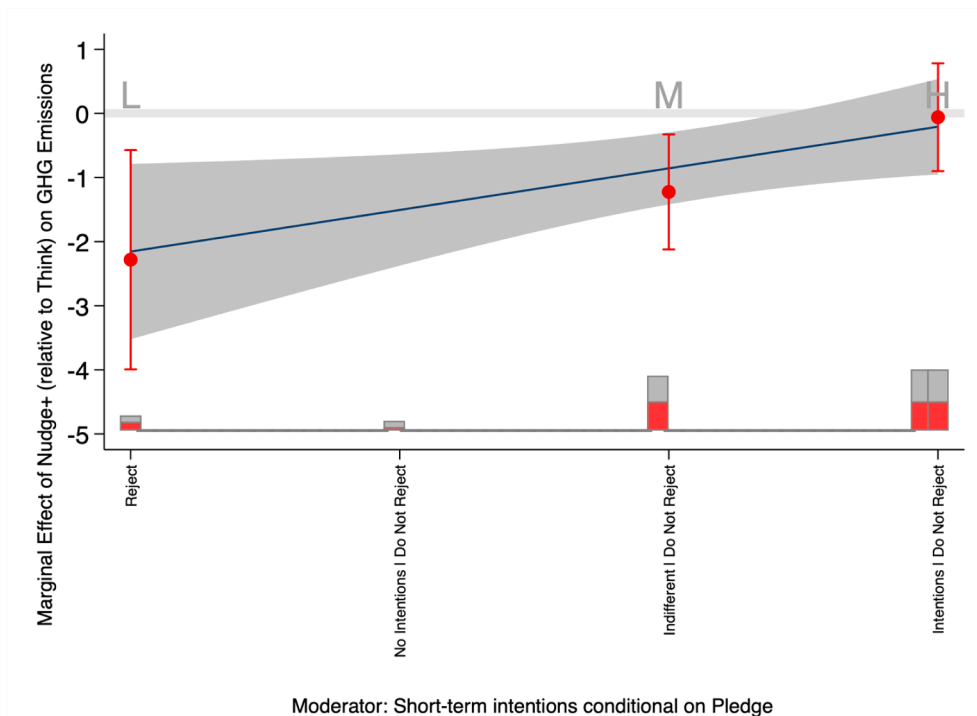


Fig. 4. Moderation effects for short-term intentions conditional on the pledge.

respondents had to select a food item from the list of all main course items.

We have missing values in the intention measures: 38 and 27 respondents in the think and nudge+ category, respectively, did not report these intentions as they had rejected the pledge in the first step. Nonetheless, we meet our sampling requirements in each case (for power analysis, see A.3).

Fig. 3 shows these short-term (L) and longer-term (R) intentions to consume sustainable diets. Participants assigned to the think condition tend to have stronger intentions, on average, to consume an environmentally sustainable diet in their next meal compared to those in the nudge+ treatment. However, this is not true for longer-term intentions (over the next week). Table 2 presents our findings from multiple linear regression models controlling for heterogeneity in treatment effects driven by the short- and longer-term intentions of respondents to consume sustainable diets, conditional on the decision made by respondents on the offer of the pledge. To see the effect of reflection on those who are already intrinsically motivated, we include an interaction term of self-reported short- or longer-term intentions with an indicator variable reflecting respondents' decision about the pledge as our moderator variable. This interaction term corresponds to the "conditional intention" variable described in Section 3.1. above. The indicator,  $i$ , is defined below.

The indicator,  $i$ , is defined below.

$$i = \begin{cases} 1, & \text{if respondent was indifferent to or accepted the pledge when offered} \\ 0, & \text{if respondent rejected the pledge when offered} \end{cases}$$

We find that for the respondents who express intentions to consume sustainably and sign up for the pledge, there is reactance driven by the short-term intentions. Thus, conditional on taking the pledge, intending to eat sustainably in the short run attenuates the treatment effects corresponding to the nudge+. Conditional on whether a respondent take the pledge or not, a one-point increase in the intentions of consuming a

sustainable diet in the next meal significantly attenuates the treatment effect by 2.596 kgCO<sub>2</sub>e. That is equivalent to say the treatment effect is attenuated by 30 percent. However, when we consider their long-term intentions, conditional on the uptake of the pledge, we do not find any significant moderation effects. When both short- and longer-term intentions, conditional on taking the pledge, are controlled for, we do not find any reactance.

To test if simply having intentions generates reactance to the interventions, we re-run our regression analysis with unconditional short and longer-term intentions as moderators. We find that intentions to consume sustainably, regardless of the decision on the pledge, do not moderate the treatment effects of any of the behavioural interventions. These results are presented in Table A.4 in the Appendix. Thus, it is only when the motivated take the pledge that we find reactance. Nonetheless, unconditional intentions of consuming a sustainable diet, either in the next meal or over the entire week, are negatively associated with greenhouse gas emissions. Thus, while intending to consume a sustainable diet is significantly correlated with reduced emissions through the choice of less carbon-intensive foods, they do not moderate the treatment effects. These main effects of intentions on reducing emissions are stronger for the short-term relative than for the longer-term. This reactance increases emissions over orders of meal choices, on average, by 30 percent. In other words, when participants with intentions to eat sustainably in the short-term are nudged, they choose food items with 30 percent more emissions compared to respondents who lack definitive intentions. If we compare participants who have short-term intentions conditional on taking the pledge to participants who reject the pledge and have no intentions, nudging the former leads to a backfire that is strong enough to nullify the positive treatment effects ("reduced emissions") of the nudge, on average (see Fig. 4).

As a robustness check, we follow Hainmueller, Mummolo, and Xu (2019) in testing the validity of the linear interaction effect assumption by using the four categories (labelled "0" to "3") of our moderator

variable measuring conditional intention. We find that across different levels of the moderator, we have common support in the treatment conditions – in other words, we can meaningfully compare the “think” and the “nudge+” conditions with each other across these different levels of unconditional intention. We also find no significant non-linear effects (p-value of Wald test statistic is 0.26 and 0.14 for short- and long-run intentions, conditional on taking up the pledge), suggesting that a linear interaction between the treatment and our moderator variable captures the heterogeneity in the average treatment effects. Fig. 4 shows the marginal effects of the treatment on greenhouse gas emissions in these different categories of the short-run intentions. As we have discussed before, we find that short-term intentions, conditional on the uptake of the pledge, significantly attenuates the treatment effect of nudge+, relative to the thinks, in a linear fashion. Overall, in comparison to the think, defaulting participants into a green menu after they have taken the pledge decreases the effectiveness of the intervention particularly for people who have stronger intentions to eat sustainably (Cohen’s  $d = 0.247$ ;  $\sigma = 0.019$ ).

Further, as an exploratory analysis, we also check for the presence of “moral licensing” behavioural spillovers (Dolan & Galizzi, 2015). Digressing from our pre-registration plan, we later anticipated that respondents can self-report incoherent intentions which can lead to behavioural spillovers due to their tendency to compensate for good intentions in the present or future by giving themselves “moral license” by choosing unsustainable items from the food menu; for instance, if Mary reports strong intentions to consume sustainably, in her next meal and/or over the next week, she is also likely to deviate, just for once, when it comes to choosing from the actual menu because she feels morally licensed to do so. To assess for such moral licensing behavioural spillovers, if any, we included an indicator variable,  $s$  such that,

$$s = \begin{cases} 1, & \text{if } \text{respondent has similar short- and long-term intentions} \\ 0, & \text{otherwise} \end{cases}$$

We also do not find any evidence for the presence of moral licensing behavioural spillovers effects. Thus, temporal incoherence in intentions, conditional or unconditional on the offer of the pledge, does not stimulate non-environmental (carbon-intensive) choices from the menu.

Finally, we assess if respondents’ self-reported levels of control over their life decisions are affected by reflective interventions. To do this, we rely on respondents self-reported beliefs on control over their life-decisions, using a 5-point Likert-scale, pre- and post-treatment. Using these responses, we find that difference in perceived self-control does not vary across the treatments. We correct for all these multiple comparisons using the Benjamini and Hochberg (1995) correction (with a false discovery rate of 10 percent). We find that moderation effects of short-term intentions, conditional on the uptake of the pledge, are still statistically significant at 5 percent level of significance. These corrections are reported in A.5.

## 5. Discussion and conclusion

In this paper, we find that nudging people who intend to shift their food choices towards a sustainable diet, after that they have initially reflected on a pledge, can lead to backfiring effects. This is true only when we consider intentions in the short-term for people who have taken the sustainable diet pledge. Longer-term intentions do not moderate these nudge effects. Our findings relate to the ones by Bruns and Perino (2019) who show that intrinsic levels of motivation can moderate treatment effects, but only when they are conditional on reflective engagement with behavioural interventions. Reactance fades when we consider the temporality in the intention of respondents. Thus, reactance

to policies at the beginning must not necessarily deter policymakers and must not be perceived as a failure of the intervention. Rather policymakers must build in reflection in the tool that respects the agency of citizens to make welfare-improving choices for themselves. Respondents who already intend to adopt sustainable food choices, when steered towards those choices, might react to it but only through immediate intentions. Nonetheless, this reactance can be sometimes necessary for transforming overall perspectives, as we see that sustainable food choices are reconciled when longer-term intentions, realised through reflection, are considered by the policymakers. Persistent behavioural transformation, therefore, needs patience: on the policymakers’ side in tolerating some initial reactance to behavioural stimulus, and on the citizens’ side to come to terms with their momentary cognitive reactance to the stimulus. Thinking about short-term and longer-term intentions keeps perceived self-control over decisions unchanged, and as such, overcomes psychological reactance.

Our findings contribute to a growing literature on psychological reactance towards nudges—such as when nudges are incongruent with people’s inherent values, beliefs, and motivations (Sunstein, 2017a), or when receivers do not trust the source of the nudge (Sunstein, Reisch, & Kaiser, 2019). Findings from psychological reactance theory resonate with resistance to nudges, in that nudges seem to backfire when citizens view them as agency reducing, with one’s intrinsic motivation moderating these differences (Bruns & Perino, 2019). Although we are unable to test further the causal mechanisms underlying these backfiring effects in our experiment, we conjecture that when participants, who pledge first (so retain self-control) and are nudged later towards sustainable diets, versus those who are not, can experience a loss of agency – their freedom to choose and act for themselves. Overriding people’s intentions, which are true sources of human agency (see Anscombe, 1957, Davidson, 1963), can therefore be problematic for steering behaviour change.

Consequently, when nudges are targeted to many people, who inherently differ in their intentions or motivations, they can cause the inevitable harm by partially or completely crowding out good decisions, in this case, sustainable food choices (Bruns & Perino, 2021). An attempt to steer welfare-improving choices, therefore, may run the risk of demotivating the already motivated.

These shortcomings in nudges have been attempted to be reconciled in different ways, for example by making them more transparent (Loewenstein, Bryce, Hagmann, & Rajpal, 2016). To this aim, backfiring effects have been absent when nudges have been made explicit to the receivers (Gråd, Erlandsson, & Tinghög, 2021). Recently, Schumpe, Bélanger, and Nisa (2020) have also shown that deploying reactance decoys prior to target messages can increase the uptake of the interventions. Rather than attempting to minimise reactance, citizen’s engagement with decoys enables them to utilise it and boost the effects of persuasive messages. Including our findings, it is possible that psychological reactance, which is expressed by an initial resistance to the nudge, can become the pre-cursor to a more sustainable behaviour change, especially when we respect people’s agency to begin with.

If so, it might be possible to lever reactance in the short run to secure a future change in the desired direction, like the pearl fisher who prefers some grit in the oysters to encourage a defensive reaction that leads to the pearl. Too easy compliance with a nudge might easily fade and show an unthinking reaction that could easily reverse. The initially resistant subject might summon up the intellectual energy to engage with a message, as a prelude to long lasting behavioural changes, as some cognitive dissonance might be the stimulant for more profound thoughts. The effect is something of a Pauline conversion where initial resistance and hostility is the preparation for a more profound change of

views. One of the impacts of reactance is the need to return to freedom which can be the prompt for more thought. For this reason, restoring autonomy and agency of citizens prior to being nudged, in respecting their cognitive capacities.

Through our study, we also contribute to the literature testing the external validity of behavioural interventions, as outlined by the “nudge+” concept proposed by Banerjee and John (2021) as an enhancement of nudges. These reformed nudging tools can be used to engage citizens in owning behaviour change through deliberation. Nudge+ builds on John and Stoker’s (2011) work on “thinks”, purely reflective strategies, which are difficult to scale-up. This is reconciled in the nudge+ which embeds reflection in the nudge to switch citizens from “thinking fast to slow”; in turn facilitating a better uptake of the nudge without disparaging their self-control. Enabling citizens to reflect on the construct of the nudge or their life decisions, before they are presented with the nudge itself, increases the transparency of choice architecture and restores the autonomy of the individual. In this way, citizens undergo a process of perspective transformation that manifests in longer-term behaviour change. When citizens are presented with an initial stimulus, the nudge, along with an active trigger to think, the plus, individuals assess the relevance of the intervention to their personal beliefs. They react to it on deliberation, either overriding the initial stimulus, or updating their beliefs as they gravitate towards the new behaviour. The nudge+ framework proposes intentions as sufficient to deliver behavioural change. Here, we contribute to the vast literature on heterogeneous treatment effects of nudges (e.g. Bolderdijk, Gorsira, Keizer, & Steg, 2013; van den Broek, Bolderdijk, & Steg, 2017; Ling, Xu, & Chu, 2023) by testing if intentions moderate the effect of the reflection as embedded in the nudge by generating some reactance or backfires in the process.

We also note some limitations to our current study. First, we use a moderation-of-process design to look at intentions as moderators, where we directly measure intentions via a questionnaire (before the experimental conditions). An alternative, and arguably cleanest, approach to prove that intentions are the behavioural mechanisms explaining why the interventions changed the outcomes would have been to directly manipulate intentions using an experimental-causal-chain design. This alternative design was, unfortunately, not viable to us because of financial and logistic constraints. Such an approach is left for future research, together with other alternative designs such as measurement-of-mediation, for example. Second, while participants were given vouchers reporting the food items they picked as their preferred choices that could be redeemed at several major high-street chains of restaurants in the UK, we cannot completely exclude the possibility that, when actually ordering their meals from those chains of restaurants, some respondents at the end converted those vouchers into some other food items that were worth the same monetary amount (£15) in the restaurants. Third, it is fair to note that participants randomly allocated to the think or the nudge+ conditions did not necessarily know which food choices were sustainable and which were not when exposed to the pledge: this aspect should be kept in mind when interpreting the results. Finally, respondents in the survey experiment faced one-off decisions and we were unable to follow up with them after. Therefore, we rely on the stated reports of their longer-term intentions. We do not know if

these longer-term intentions reverse when they respondents arrive at the end of the week, when such longer-term intentions become short-term once again, and new longer-term intentions are developed. Consequently, our results need to be externally validated in the field with sequential or repeated decision making. However, we provide an initial insight to behavioural researchers and policymakers to design interventions that engage citizens in their food choices. Human decision-makers need to be given due credit for their agency; for even if they react, they are simply engaging with the process of behavioural transformation, which can be initially repulsive but eventually leads one down the pathway of sustainable food choices. Our findings from this pre-registered, online experiment reaffirms these notions as outlined by Banerjee and John (2021).

To summarise, we show that short- and long-term intentions are significantly correlated with reduced emissions through sustainable food choices. However, these intentions do not attenuate the treatment effects of the interventions independently from any reflection on the pledge. It is only by engaging in the pledge that the respondents who already intend to eat sustainably generate reactance. The effect is limited to self-reported short-run intentions of consuming sustainable diets with reactance disappearing when longer-term intentions, conditional on the pledge, are accounted for. Furthermore, respondents’ level of perceived self-control over their life decisions, measured before and after the treatment, also remain unaffected.

#### CRediT authorship contribution statement

**Sanchayan Banerjee:** Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Matteo M. Galizzi:** Conceptualization, Formal analysis, Writing – review & editing, Funding acquisition. **Peter John:** Conceptualization, Formal analysis, Writing – review & editing, Funding acquisition. **Susana Mourato:** Conceptualization, Formal analysis, Funding acquisition.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Data availability

The dataset and the do file for replication have been attached to the submission.

#### Acknowledgement

We thank Manu Savani and Julien Picard for their feedback on an earlier drafts of our manuscript. MG acknowledges financial support from Department of Psychological and Behavioural Science (LSE), PJ acknowledges financial support from Departmental of Political Economy (KCL), and SM acknowledges financial support from Department of Geography and Environment (LSE).

#### Appendix A1. . Overall treatment design (source: Banerjee et al., 2022)

In [Table A1](#).



**Table A1**

List of interventions. Source: Banerjee et al., (2021).

Sr No	Intervention Type	Intervention Sub-Type	Intervention Vehicle	Intervention Description
1.	Control	Do Nothing	Regular Menu (RM)	Respondents received an ala-carte menu with 36 items, including 18 vegetarian and 18 non-vegetarian items.
2.	Classic Nudge	Default	Default (D)	Respondents received a default set-menu with 18 low emission intensity items only. Of these, 12 were vegetarian and 6 were non-vegetarian. Participants could opt-out for the regular menu.
		Labelling	Traffic Lights (TL)	Respondents received an ala-carte menu with 36 items. These were colour coded as follows: red (high emissions), amber (medium emissions) and green (low emissions).
3.	Thinks	Active mechanism design	Pledge	Respondents received a pledge for an environmentally sustainable diet. Following their decision to pledge or not, they could choose from a regular, default menu or traffic lighting menu.
4.	Boosts	Uncertainty management	Quick Rules	Respondents received a set of three food rules along with the regular menu.
		Motivational boost	Implementation Intentions	Respondents were asked to design six 'if-then' implementation plans (three for lunch and three for dinner). They were provided with a regular menu afterwards to choose from.
5.	Nudge+	Two-part, simultaneous	Default with information disclosure	Respondents received a default set-menu with an explicit information disclosure about construct of the default menu.
			Traffic Lights with information disclosure	Respondents received a traffic lighting menu with an explicit information disclosure about the labelling scheme.
		Two-part, sequential	Opt-out default plus Pledge	Respondents received a default set-menu menu. Post-menu choice, they received a pledge for an environmentally sustainable diet. A choice to revisit their online order was also provided to those who took the pledge.
		Pledge plus Opt-out default	Respondents received a pledge for an environmentally sustainable diet. Post-pledge decision, they received a default set-menu.	

**Appendix A2. . Randomisation checks**

Table A2 below lists all parametric and non-parametric comparisons of think and nudge+ by age, gender, and education (binary) of participants.

**Table A2**

Balance of means of age, gender, and education by treatment conditions.

Variable	Pairwise <i>t</i> -test	Kruskal Wallis test
Age	$t = 0.6467$ ; $p = 0.442$	$\chi^2 = 0.379$ ; $p = 0.5380$
Gender	$t = 0.0299$ ; $p = 0.476$	$\chi^2 = 0.552$ ; $p = 0.4576$
Education (university or not)	$t = -0.0428$ ; $p = 0.292$	$\chi^2 = 0.832$ ; $p = 0.3617$

**Appendix A3. . Power analysis**

In our pre-analysis plan (see Banerjee, John, Galizzi, and Mourato (2020)), we adopted a minimum detectable effect size of Cohen's  $d = 0.25$  (or, 25 percent) across the ten treatment arms, on an average. Using a 2-groups independent means comparison *t*-test design, the a priori computed total sample size requirement corresponded to  $n = 253$  for Cohen's  $d = 0.25$ , Power = 0.8 and Type-I error = 0.05. Since the cost of recruiting subjects in each of the treatment arms was equal, each treatment arm had an equal required sample size, resulting in a total sample size requirement of  $N = 2530$ . The critical *t* associated with this analysis was 1.963 and the non-centrality parameter was 2.807. All sample size calculations were carried out using G\*Power 3.1. The sampling requirements and rules were pre-registered on OSF (January 2020).

**Appendix A4. . Moderation effects of unconditional self-reported short- and long-term intentions to consume sustainably.**

In Table A4.

**Table A4**  
Regression results with unconditional moderation effects of intrinsic motivation.

Outcome: GHG emissions	OLS Model 4	OLS Model 5	OLS Model 6
Treatment	Baseline: Think		
Nudge+	-6.457* (2.550)	-3.235 (2.711)	-6.546* (2.978)
Short Term Intentions	-7.055*** (1.269)		-6.516*** (1.403)
Long Term Intentions		-3.559** (1.170)	-1.104 (1.261)
Moral spillovers (s = 1)	0.300 (1.122)	-1.630 (1.101)	0.181 (1.127)
Short Term Intentions#	Baseline: Think		
Nudge+	2.852 (1.648)		2.879 (1.865)
Long Term Intentions#	Baseline: Think		
Nudge+		0.584 (1.642)	0.009 (1.820)
Constant	17.182*** (1.899)	13.819*** (2.026)	18.192*** (2.216)
N	539	539	539

Table notes: Columns 4–6 present linear regression models corresponding to different moderating variables, namely short-term unconditional intention (“OLS Model 4”), longer-term unconditional intention (“OLS Model 5”), and both short- and longer-term unconditional intentions (“OLS Model 6”). Standard errors in parentheses. legend: \* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001. Corrected p-values in Table A.5.

**Appendix A5. . Benjamini-Hochberg correction**

OLS Model	Variable	Raw p-values	Rank	Benjamini Hochberg Adjusted p-value (FDR = 10%)	Significant (1 = Yes, 0 = No) if raw p-val < BH adjusted p-val
Model 1	Short Term motivation (conditional)	0	1.25	0.005681818	1
Model 2	Long Term motivation (conditional)	0	1.25	0.005681818	1
Model 4	Short Term motivation (unconditional)	0	1.25	0.005681818	1
Model 6	Short Term motivation (unconditional)	0	1.25	0.005681818	1
Model 3	Short Term motivation (conditional)	0.0001	5	0.022727273	1
Model 1	Treatment effect	0.0021	6	0.027272727	1
Model 3	Treatment effect	0.0037	7	0.031818182	1
Model 2	Treatment effect	0.0093	8	0.036363636	1
Model 4	Treatment effect	0.0116	9	0.040909091	1
Model 6	Treatment effect	0.0284	10	0.045454545	1
Model 1	Interaction effect with Short term motivation (conditional)	0.029	11	0.05	1
Model 4	Interaction effect with Short term motivation (unconditional)	0.0841	12	0.054545455	0
Model 2	Interaction effect with Long term motivation (conditional)	0.0983	13	0.059090909	0
Model 6	Interaction effect with Short term motivation (unconditional)	0.1232	14	0.063636364	0
Model 3	Interaction effect with Short term motivation (conditional)	0.1931	15	0.068181818	0
Model 5	Treatment effect	0.2333	16	0.072727273	0
Model 5	Long Term motivation (unconditional)	0.3818	17	0.077272727	0
Model 6	Long Term motivation (unconditional)	0.3818	18	0.081818182	0
Model 3	Long Term motivation (conditional)	0.7519	19	0.086363636	0
Model 3	Interaction effect with Long term motivation (conditional)	0.9954	20	0.090909091	0
Model 5	Interaction effect with Long term motivation (unconditional)	0.9963	21	0.095454545	0
Model 6	Interaction effect with Long term motivation (unconditional)	0.9963	22	0.064705882	0

**Appendix A6. . Meal choices by respondents in the “think” and “nudge+” experimental conditions.**

In Table A6.

Table A6

Types meals ordered across “think” and “nudge+” experimental conditions.

Food Type Chosen	Think	Nudge+	Total
Beans and Lentils	19	25	44
Pasta	51	61	112
Vegetable Dishes	23	30	53
Nuts and Seeds	16	25	41
Duck/White Fish/Chicken	146	141	287
Pork	23	6	29
Cheese	2	4	6
Lamb	2	2	4
Beef	21	7	28
Total	303	301	604

## References

- Anscombe, G. E. M. (1957). *Intention*. Oxford: Basil Blackwell.
- Banerjee, S., & Picard, J. (2023). Thinking through norms can make them more effective. Experimental evidence on reflective climate policies in the UK. *Journal of Behavioral and Experimental Economics*, 102024. <https://doi.org/10.1016/j.joecoc.2023.102024>
- Banerjee, S., John, P., Galizzi, M. M., & Mourato, S. (2020). What works best? A comparative analysis of nudges, boosts, thinks and nudges+ in sustainably altering dietary choices. <https://osf.io/8ftu3>.
- Banerjee, S., Galizzi, M. M., John, P., & Mourato, S. (2022a). What works best in promoting climate citizenship? A randomised, systematic evaluation of nudge, think, boost and nudge+. A Randomised, Systematic Evaluation of Nudge, Think, Boost and Nudge+ (April 26, 2022). Research Square pre-prints.
- Banerjee, S., Hunter, A., John, P., Koenig, R., Lee-Whiting, B., Loewen, P., ... & Savani, M. (2022b). Citizen reflection reduces support for default vaccination appointments. Evidence from G7 countries.
- Banerjee, S., & John, P. (2021). Nudge+: Incorporating reflection into behavioural public policy. *Journal of Behavioural Public Policy*. <https://doi.org/10.1017/bpp.2021.6>
- Bauer, J. M., Bietz, S., Rauber, J., & Reisch, L. A. (2021). Nudging healthier food choices in a cafeteria setting: A sequential multi-intervention field study. *Appetite*, 160, 10510. <https://doi.org/10.1016/j.appet.2021.105106>
- Baumeister, R. F., & Heatherton, T. F. (1996). Self-Regulation Failure: An Overview. *Psychological Inquiry*, 7(1), 1–15. [https://doi.org/10.1207/s15327965pli0701\\_1](https://doi.org/10.1207/s15327965pli0701_1)
- Baumeister, R. F., Leith, K. P., Muraven, M., & Bratslavsky, E. (2002). Self-Regulation as a Key to Success in Life. In D. Pushkar, W. M. Bukowski, A. E. Schwartzman, D. M. Stack, & D. R. White (Eds.), *Improving Competence across the Lifespan: Building Interventions Based on Theory and Research* (pp. 117–132). US: Springer. [https://doi.org/10.1007/0-306-47149-3\\_9](https://doi.org/10.1007/0-306-47149-3_9).
- Bénabou, R., & Tirole, J. (2003). Intrinsic and Extrinsic Motivation. *The Review of Economic Studies*, 70(3), 489–520. <https://doi.org/10.1111/1467-937X.00253>
- Bénabou, R., & Tirole, J. (2006). Incentives and Prosocial Behavior. *American Economic Review*, 96(5), 1652–1678. <https://doi.org/10.1257/aer.96.5.1652>
- Benjamin, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300.
- Bergeron, S., Doyon, M., Saulais, L., & Labrecque, J. (2019). Using insights from behavioral economics to nudge individuals towards healthier choices when eating out: A restaurant experiment. *Food Quality and Preference*, 73, 56–64.
- Beshears, J., & Kosowsky, H. (2020). Nudging: Progress to date and future directions. *Organizational behavior and human decision processes*, 161, 3–19.
- Bolderdijk, J. W., Gorsira, M., Keizer, K., & Steg, L. (2013). Values determine the (in) effectiveness of informational interventions in promoting pro-environmental behavior. *PLoS One*, 8(12), e83911.
- Bruns, H., & Perino, G. (2021). Point at, nudge, or push private provision of a public good? Economic Inquiry, n/a(n/a). Doi: 10.1111/ecin.12981.
- Bruns, H., & Perino, G. (2019). The Role of Autonomy and Reactance for Nudging—Experimentally Comparing Defaults to Recommendations and Mandates (SSRN Scholarly Paper ID 3442465). *Social Science Research Network*. <https://doi.org/10.2139/ssrn.3442465>
- Chater, N., & Loewenstein, G. (2022). The i-frame and the s-frame: How focusing on the individual-level solutions has led behavioral public policy astray. Available at SSRN 4046264.
- D’Adda, G. (2011). Motivation crowding out in environmental protection: Evidence from an artefactual field experiment. *Ecological Economics*, 70(11), 2083–2097.
- Davidson, D. (1963). *“Actions, Reasons, and Causes”*, reprinted in *Davidson 1980 Essays on Actions and Events* (pp. 3–20). Oxford: Clarendon Press.
- Di Clemente, C. C., Prochaska, J. O., Fairhurst, S., Velicer, W. F., Rossi, J. S., & Velasquez, M. (1991). The process of smoking cessation. *Journal of Consulting and Clinical Psychology*, 59, 295–304.
- Dolan, P., & Galizzi, M. M. (2015). Like Ripples on a Pond: Behavioural Spillovers and Their Consequences for Research and Policy. *Journal of Economic Psychology*, 47, 1–16.
- Ellingsen, T., & Johannesson, M. (2008). Pride and Prejudice: The Human Side of Incentive Theory. *American Economic Review*, 98(3), 990–1008. <https://doi.org/10.1257/aer.98.3.990>
- Evans, J. S. B. T. (1984). Heuristic and analytical processes in reasoning. *British Journal of Psychology*, 75, 451–468.
- Gråd, E., Erlandsson, A., & Tinghög, G. (2021). Do nudges crowd out prosocial behavior? *Behavioural Public Policy*, 1–14. <https://doi.org/10.1017/bpp.2021.10>
- Gravert, C., & Kurz, V. (2021). Nudging à la carte: A field experiment on climate-friendly food choice. *Behavioural Public Policy*, 5(3), 378–395.
- Hainmueller, J., Mummolo, J., & Xu, Y. (2019). How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice. *Political Analysis*, 27(2), 163–192. <https://doi.org/10.1017/pan.2018.46>
- John, P. (2011). Nudge, nudge, think, think: Experimenting with ways to change civic behaviour. *Bloomsbury Academic*.
- John, P., Cotterill, S., Moseley, A., Richardson, L., Smith, G., Stoker, G., & Wales, C. (2019). *Nudge, nudge, think, think: Experimenting with ways to change citizen behaviour, second edition* (1st edition). Manchester University Press.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58, 697–720.
- 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*, 90, 317–341.
- Lacetera, N., Macis, M., & Slonim, R. (2012). Will There Be Blood? Incentives and Displacement Effects in Pro-social Behavior. *American Economic Journal: Economic Policy*, 4(1), 186–223. <https://doi.org/10.1257/pol.4.1.186>
- Langer, E. J., & Rodin, J. (1976). The effects of choice and enhanced personal responsibility for the aged: A field experiment in an institutional setting. *Journal of Personality and Social Psychology*, 34(2), 191–198. <https://doi.org/10.1037//0022-3514.34.2.191>
- Ling, M., Xu, L., & Chu, X. (2023). Heterogeneous effects of other-regarding interventions on household recycling: A field experimental study. *Journal of Environmental Management*, 329, Article 117102.
- Loewenstein, G., Bryce, G., Hagmann, G., & Rajpal, G. C. D. S. (2016). Warning: You are about to be nudged. *Behavioral Science & Policy*, 1, 35–42.
- Lohmann, P. M., Gsottbauer, E., Doherty, A., & Kontoleon, A. (2022). Do carbon footprint labels promote climatarian diets? Evidence from a large-scale field experiment. *Journal of Environmental Economics and Management*, 114, Article 102693.
- Maier, M., Bartoš, F., Stanley, T. D., Shanks, D. R., Harris, A. J., & Wagenmakers, E. J. (2022). No evidence for nudging after adjusting for publication bias. *Proceedings of the National Academy of Sciences*, 119(31).
- Mertens, S., Herber, M., Hahnel, U. J., & Brosch, T. (2022). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences*, 119(1).
- Osman, M., McLachlan, S., Fenton, N., Neil, M., Löfstedt, R., & Meder, B. (2020). Learning from behavioural changes that fail. *Trends in Cognitive Sciences*, 24(12), 969–980.
- Prochaska, J. O., Wright, J. A., & Velicer, W. F. (2008). Evaluating theories of health behavior change: A hierarchy of criteria applied to the transtheoretical model. *Applied Psychology*, 57(4), 561–588.
- Prusaczyk, E., Earle, M., & Hodson, G. (2021). A brief nudge or education intervention delivered online can increase willingness to order a beef-mushroom burger. *Food Quality and Preference*, 87, Article 104045.
- Rode, J., Gómez-Baggethun, E., & Krause, T. (2015). Motivation crowding by economic incentives in conservation policy: A review of the empirical evidence. *Ecological Economics*, 117, 270–282.
- Rosenberg, B. D., & Siegel, J. T. (2018). A 50-year review of psychological reactance theory: Do not read this article. *Motivation Science*, 4(4), 281–300. <https://doi.org/10.1037/mot0000091>
- Schumpe, B. M., Bélanger, J. J., & Nisa, C. F. (2020). The Reactance Decoy Effect: How Including an Appeal Before a Target Message Increases Persuasion. *Journal of Personality and Social Psychology*, 119(2), 272–292. <https://doi.org/10.1037/pspa0000192>
- Seligman, M. E. P. (1975). Helplessness: On depression, development, and death. WHFreeman.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3–22.
- Sunstein, C. R. (2017a). People Like Nudges (Mostly). In C. R. Sunstein (Ed.), *Human Agency and Behavioral Economics: Nudging Fast and Slow* (pp. 17–39). Springer International Publishing. [https://doi.org/10.1007/978-3-319-55807-3\\_2](https://doi.org/10.1007/978-3-319-55807-3_2)
- Sunstein, C. R. (2017b). Nudges that fail. *Behavioural Public Policy*, 1(1), 4–25. <https://doi.org/10.1017/bpp.2016.3>

- Sunstein, C. R., Reisch, L. A., & Kaiser, M. (2019). Trusting nudges? Lessons from an international survey. *Journal of European Public Policy*, 26(10), 1417–1443. <https://doi.org/10.1080/13501763.2018.1531912>
- Szaszi, B., Palinkas, A., Palfi, B., Szollosi, A., & Aczel, B. (2018). A systematic scoping review of the choice architecture movement: Toward understanding when and why nudges work. *Journal of Behavioral Decision Making*, 31(3), 355–366.
- Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth and happiness* [revised edition]. Penguin.
- Titmuss, R. M. (1970). *The gift relationship: From human blood to social policy*. Allen & Unwin.
- Travers, H., Clements, T., Keane, A., & Milner-Gulland, E. (2011). Incentives for cooperation: The effects of institutional controls on common pool resource extraction in Cambodia. *Ecological Economics*, 71, 151–161.
- van den Broek, K., Bolderdijk, J. W., & Steg, L. (2017). Individual differences in values determine the relative persuasiveness of biospheric, economic and combined appeals. *Journal of Environmental Psychology*, 53, 145–156.
- Vecchio, R., & Cavallo, C. (2019). Increasing healthy food choices through nudges: A systematic review. *Food Quality and Preference*, 78, Article 103714.