Journal of Behavioral
Public Administration
Vol 9, pp. 1-25
DOI: 10.30636/jbpa.91.398

The Choice-Changing Effects of Post-Disclosure: Default Nudges Revisited

Kazumi Shimizu*, Akira Inoue†, Daisuke Udagawa, Yoshiki Wakamatsu††**

Abstract: This study explores whether the effectiveness of default nudges, which involve preselecting a preferred option, persists after disclosing the nudge to individuals. While numerous earlier studies have suggested disclosure did not diminish nudge efficacy, they have not explicitly situated the nudging agent's identity and motivations within specific political and social contexts. In this survey experiment, which incorporated such information into the disclosure, 2,791 Japanese respondents were presented with hypothetical choices to test post-disclosure effects across four policy areas: healthy eating, organ donation, COVID-19 vaccination, and seasonal influenza vaccination. The study examined three types of disclosure: neutral, public, and private. The findings reveal that post-disclosure decreases the likelihood of individuals adhering to the default choice, with effectiveness influenced by demographic factors and the nature of the policy. Contextual and cultural factors may account for why disclosure meaningfully affects nudge effectiveness. Further research is needed to explore cultural differences in the impact of nudge disclosure.

Keywords: default nudge, post-disclosure, policies, experiment.
Supplements: Open data, Open materials

Sunstein and Thaler, in their 2003 paper "Libertarian Paternalism Is Not an Oxymoron," laid the complementary normative foundations for nudging as a public and private policy under the concepts of "libertarian paternalism" (hereafter "LP") and "libertarian benevolence" (hereafter "LB"). LP is a paternalistic position that prompts, or "nudges," people to choose actions that will make them better off. LB proposes policies that nudge people to act in the interests of vulnerable third parties without making the choosers worse off.[1] The former corresponds to what has been termed "pro-self" nudges, while the latter corresponds to "pro-social" nudges or simply social nudges (Nagatsu, 2015), i.e., nudges that aim to benefit others without harming the chooser (about "pro-self" and "pro-social" nudges, see Hagman et al., 2015 & Congiu & Moscati, 2022). These labels help situate our discussion within a broader taxonomy of nudging practices. Following the introduction of these concepts, nudging has gained attention as an effective means of influencing people's choices in real-world policies without significantly altering economic incentives or limiting choices.

Among various forms of nudging, a default nudge capitalizes on status quo bias by pre-setting an option that policymakers or institutions consider optimal, while still allowing individuals to opt out.[2] For instance, organ donation systems that presume consent unless individuals explicitly opt out represent a default nudge. Default nudges raise ethical concerns because they may bypass reflective choice and thereby compromise individual autonomy (Bovens, 2008). To satisfy standards of ethical accountability, such choice architecture should be transparent and include explicit disclosure of both its presence and intended influence (Michaelsen, 2024).[3] These requirements align with broader principles of liberal legitimacy and informed consent in behavioral public policy design (Grüne-Yanoff, 2012; Sunstein, 2015). From this normative perspective, default

* Waseda University
† University of Tokyo
** Hannan University
†† Gakushuin University
Address correspondence to Kazumi Shimizu at [skazumi1961@gmail.com]

nudges that lack transparency at the point of choice risk being manipulative, even if their presence is disclosed after the initial decision (Hansen & Jespersen, 2013). These debates about the autonomy- and transparency-related implications of default nudges underscore that ethical concerns about autonomy and transparency remain fundamental to their legitimate use.

To address these ethical concerns, increasing the transparency of default nudges is essential. For example, nudgers could disclose the effects and intentions behind the default option to nudgees, either before or after implementation. However, this approach raises a problem: such disclosure may weaken the effectiveness of the nudge (Bovens, 2008; Grüne-Yanoff, 2012). Moreover, if the nudge is disclosed after the decision, some who accepted the default may, in some contexts, feel uneasy about having been nudged and revise their choice, thereby nullifying its intended effects.[4]

Motivated by these concerns, this study investigates whether the effectiveness of default nudges persists after post-disclosure by means of a preregistered online experiment with 2,791 participants. The experiment was conducted in Japan, where most nudge-based public policies have been implemented without disclosure to the public, making post-disclosure effects particularly relevant for understanding real-world implications. It was designed to test how different disclosure conditions—neutral, public, and private—affect individuals' willingness to maintain or revise their initial default choice.

Literature review

Previous experimental research has consistently reported that neither disclosure before nudging ("pre-disclosure") nor post-factum disclosure ("post-disclosure") reduces its effectiveness: the nudge continues to influence behavior even after disclosure (Bruns et al., 2018; Loewenstein et al., 2015; Michaelsen et al., 2020, 2021; Steffel et al., 2016). Pre-disclosure has been reported to enhance evaluations of choice architects (Paunov et al., 2019a, 2019b) and to have no significant effects on people's experiences of autonomy (Michaelsen et al., 2021; Wachner et al., 2020, 2021). Additionally, Michaelsen et al. (2020) revealed that pre-disclosure led to an increase in fairness ratings when an undisclosed version of default nudges was provided as a reference point (the increase vanished without a reference point).[5] In contrast, although Weijers et al. (2024) found no reduction in experienced autonomy, they did observe a decrease in expected autonomy when participants were presented with the same default nudge.[6] This suggests that post-disclosure may create the impression that choice architects undermine expected autonomy, even when experienced autonomy remains intact. Overall, prior studies generally support the robustness of nudge effects under both pre- and post-disclosure and affirm the ethical legitimacy of nudging under pre-disclosure, though the ethical standing of nudging after post-disclosure remains less certain (Weijers et al., 2024; Michaelsen et al., 2021).

A review of the literature about the influence of disclosure on the effects of nudges reveals that many studies do not include the minimum disclosure content that Michaelsen (2024) deems necessary for nudges to ethical, namely: (1) the fact that a nudge is being used to influence their behavior, and (2) the direction of influence intended by the nudge.[7] In addition to the two necessary disclosure elements, information about who conducts the nudge and their intentions or motivations is also critical.[8] This assumption is supported by the large body of research showing that people's perceptions of identical content can vary significantly depending on the characteristics of the information provider (for a literature review in this area, see Michaelsen et al., 2021). If the disclosure of nudges affects nudgees' perceptions of the nudges and their autonomy, they are likely to reconsider the default option and possibly change it once they become aware of the type (identity) of nudging agent (e.g., government, private firm, or individual) and of that agent's intentions.

Our experimental study aims to design disclosure conditions that consider not only the effects and intentions of nudges themselves, as emphasized by Michaelsen (2024), but also the identity, specific intentions, and motivations of nudgers. We investigate whether the effects of nudging persist after such disclosures. Note that our study focuses exclusively on the use of *default* nudges, i.e., choice architecture that pre-selects a particular choice option. While default nudges have been shown to be effective across numerous domains, they represent only one among many types of nudges. Therefore, we do not claim to offer conclusions about nudging in general.

There are two important studies related to ours; we will now clarify the differences between our current study and these earlier analyses. Steffel et al. (2016) examined ethical concerns surrounding default nudges,

investigating how disclosure affects their effectiveness and perceived legitimacy by varying the nudger's identity (we, I, or research) and motivations (pro-self, pro-social, or neither) across seven scenarios. While their research shares some similarities with ours, it differs in three key respects. First, in our study, nudging agents were institutionally specified (public vs. private). Second, our study focused on public policy domains—i.e., healthy eating, organ donation, COVID-19 vaccination, and infectious disease vaccination. And third, in their study, participants were asked—under either a pre-disclosure or post-disclosure condition—what type of information they would share, what kinds of surveys they would complete, and whether the default setting was considered ethical. In contrast, in our study, participants made an initial choice under a default option and were later permitted to revise it under the post-disclosure condition.

Van Rookhuijzen et al. (2023) investigated whether nudge effects persist over time through post-disclosure transparency. Although their study and ours share a focus on the transparency, they also differ in two important ways. First, Van Rookhuijzen and colleagues did not manipulate disclosure conditions. Second, their study examined whether temporal spillover effects—namely, effects of non-transparent nudges that persist and influence later similar decisions made without nudges—are influenced when the nudges are made transparent. In contrast, the purpose of our study is not to investigate these temporal spillover effects, but to examine individuals' willingness to maintain or revise their initial default choice by making nudges transparent.

Drawing in part on Loewenstein et al. (2015) and Michaelsen et al. (2021), the present study extends existing research on this phenomenon in three ways. First, we focus on public policy contexts that encourage healthy food choice, organ donation and vaccination, where transparency standards differ substantially from private settings. Second, we systematically examine how the nudging agent's identity (public vs. private) and stated motivations (pro-self vs. pro-social) influence post-disclosure responses—dimensions unexplored previously. Third, we provide evidence from Japan, where institutional trust and sensitivity to perceived manipulation may differ from Western populations.

Methods

Purpose

Our study examines whether and the extent to which post-disclosure of opt-out defaults affects people's behavioral change in four important policy contexts: lifestyle intervention (toward healthy food choices in casual dining), organ donation, and two types of vaccines, i.e., COVID-19 vaccination and viral infectious disease vaccination.

We focus only on the effects of post-disclosure treatments. There are two reasons for this decision. First, the effects of post-disclosure on default nudge efficacy have been underexplored, despite their relevance for understanding experienced autonomy and potential feelings of manipulation (Michaelsen et al., 2024). As discussed earlier, while previous research has demonstrated that pre-disclosure has negligible or no effects on people's perceptions of nudges, there are indications that post-disclosure may influence such perceptions (Paunov et al., 2019a, 2019b; Weijers et al., 2024; Michaelsen et al., 2021). In light of these reports, thus, it would be reasonable to devote more attention to post-disclosure effects, with a view to addressing the ethical concerns in our experiment.

Second, in Japan, the Ministry of the Environment's "Guidelines for Implementing Nudges in Society" provides an overview of public policies based on nudging but does not specify the details of how information disclosure should be handled during and after implementation. In real-world policymaking, according to reports from various municipalities in Japan, almost all nudge choice architectures were implemented without any disclosure to the public, either prior to or following the nudge. In rare cases, individuals became aware that they had been nudged only after making inquiries to the municipalities. Therefore, in Japan, people may become aware that they have been nudged either by making their own inquiry afterward or by reading a report on the matter issued by the municipality. With these two considerations in mind, to examine the sustainability of nudging effects in real-world settings and enhance the practical relevance of this experiment, it was deemed appropriate to first implement post-disclosure rather than pre-disclosure in our experiment.[9]

We outline the four policy areas and the three disclosure conditions that were implemented in the present experimental design. The four key policy areas identified for nudging (healthy food choices in the casual

dining, organ donation, and two types of vaccination) were selected based on a detailed analysis of Sunstein and Thaler's normative discussions on LP/pro-self and LB/pro-social. In particular, our analysis led us to classify healthy food choices as LP/pro-self, organ donation as LB/pro-social, and the two types of vaccination as a combination of both LP/pro-self and LB/pro-social; vaccination is not only for the promotion of the interests (own well-being) of the vaccinated, but also for social benefits (societal well-being), especially for those vulnerable to viral infectious diseases from a public health perspective. This implementation enables an examination of the sustainability of the effects of nudging on targets that are ethically endorsed for nudging on the basis of Sunstein and Thaler's arguments.[10]

Three disclosure conditions were implemented to consider the effects of nudging. The first condition pertained solely to the consequences of nudges themselves. The second condition disclosed the effects of the nudge and indicated that nudging was conducted by a public entity for the benefit of the public. The third condition disclosed the effects of the nudge and indicated that the nudge was conducted by a private entity for the benefits of the private entity.[11] The second and third conditions differ from the first in that they include both the identity and intentions of the nudging agent, whereas the first does not.

The main dependent variable is the change in people's choice after post-disclosure. Rather than guiding an informed initial decision by pre-disclosure, post-disclosure may force individuals to reassess a choice they have already made, potentially producing reactance, frustration, or moral discomfort—especially if the disclosed nudger is perceived as self-serving. In this sense, post-disclosure may not simply reduce experienced autonomy, but may instead evoke negative affective responses, such as feeling manipulated or tricked, which drive choice revision. We anticipate stronger effects when the nudger's motivations appear exploitative (e.g., in the case of profit-seeking firms) and weaker effects when motivations appear benevolent or institutionally legitimate (e.g., public health agencies). However, notice that we cannot make a definitive prediction regarding the degree to which identity and intent impact the efficacy of default nudges, because prior evidence about the effects of post-disclosure on autonomy (Paunov et al., 2019a, 2019b; Weijers et al., 2024; Michaelsen et al., 2021) is both limited and mixed, with some studies reporting autonomy impairment and others not.[12] Hereafter, we use the term "disclosure" to indicate "post-disclosure" to simplify the description.

Experiment

Participants. A private research company (Rakuten Insight, Inc.) was asked to recruit respondents for our online experiment. These respondents had voluntarily applied to the research company to participate in experiments from their homes by answering questions via the Internet. The instructions were presented on their computer. After the experiment, the company paid them a fee of 500 yen (approximately US$3.50). The experiment took place from March 7th to 11th 2024, with 2,791 subjects (1,420 females and 1,371 males).[13] The age distribution was 15 respondents in their teens (specifically ages 18–19), 428 in their 20s, 414 in their 30s, 497 in their 40s, 498 in their 50s, 440 in their 60s, and 499 in their 70s. Our sample roughly corresponds to the age and gender distribution of the actual population in Japan. Although this is not a nationally representative sample in the strict statistical sense, the close correspondence allows us to cautiously generalize our findings to the broader Japanese population. However, it should be noted that there may exist a bias in that the participants were voluntarily registered with Rakuten Insight. To be clear, the experiment was conducted in Japan and was not monetarily incentivized. The scenario and disclosure statements used in our experiment are entirely hypothetical, and participants knew that no actual policy manipulation occurred.[14]

Design and materials. We constructed four policy contexts: (a) healthy food choices, (b) organ donation, (c) COVID-19 vaccination, and (d) viral infectious disease vaccination. Each policy choice comprised two cases: (1) an opt-out default (the selected option allows respondents to uncheck it) and (2) no default as the baseline (two available options: respondents must select one). The choice ratio of the two options in case (2) represents the ratio of choices people make when they are not nudged.

Taking the organ donation question as an example, the following is displayed in the answer column for (1):

"Please answer the question based on the following situation.
You are asked to fill out an organ donor card.
☑ I will donate my organs after my death.

(If you do not want to donate your organs, please uncheck the box.)"
And in the answer column for (2), the following is displayed:
"Please answer the question based on the following situation.
You are asked to fill out an organ donor card.
Which one will you choose?
□ I will donate my organs after my death.
□ I will not donate my organs after my death."

The respondents are first assigned either (1) or (2). In (1), i.e., the opt-out default case of each policy choice, respondents are exposed to one of the three types of disclosure statement about nudging after their initial choice as follows:

(a) Neutral (only the effects of the nudge are disclosed): In the previous question, "I will donate my organs after my death" was initially set as the selected option. In fact, from the viewpoint of psychology and cognitive science, it is well known that this way of asking the question leads many people to choose this option. Now we would like to ask you a question. Given this, which option would you choose?

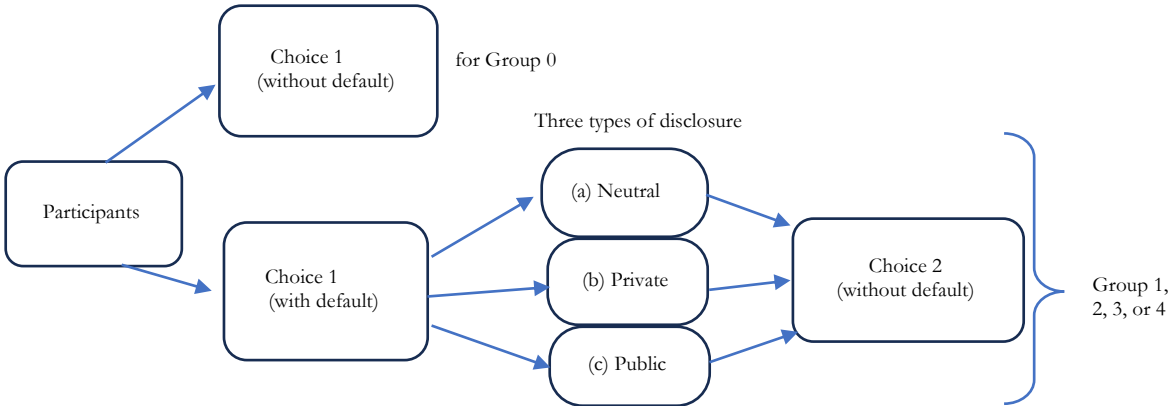
(b) Public (the effects of the nudge and that a public entity conducted it for public benefit are disclosed). Respondents are presented with the following statement: In the previous question, "I will donate my organs after my death" was initially set as the selected option. In fact, from a psychological or cognitive science point of view, it is well known that this way of asking the question makes many people tend to choose this option. And the reason the question was asked this way is as follows. In order to promote the overall social benefits, the government wanted you to choose this option, to increase the number of people who choose it, since everyone could one day need an organ donation. Now we would like to ask you a question. Given this, which option would you choose?

(c) Private (the effects of the nudge and that a private entity conducted it for private benefit are disclosed). Respondents are presented with the following statement: In the previous question, "I will donate my organs after my death" was initially set as the selected option. In fact, from a psychological or cognitive science point of view, it is well known that this way of asking the question makes many people tend to choose this option. And the reason the question was asked this way is as follows. People on the waiting list for organs wanted you to choose this option, in order to increase the number of people who choose it, and they lobbied politicians to make this happen. Now we would like to ask you a question. Given this, which option would you choose?

The statements quoted above are the actual texts used in the organ donation task. For all four policy areas and for each of the three disclosure types (neutral, public, private), the full text of the treatments is available in **Appendix A**. The wording changes depending on the presented policy context. While (a) discloses only general effects (how default affects behavior in general), (b) and (c) disclose not only the general effects but also the purpose (why the nudger wants people to choose the default option), and the agent identity (who is nudging). In this experiment, the public entity (agency) was consistently assumed to be the government across all policy areas, whereas the private entity varied by policy area. Specifically, in the healthy food choice context, the private entity was assumed to be restaurants; in the organ donation context, people on the waiting list for organs; and in the COVID-19 vaccination and seasonal influenza vaccination contexts, pharmaceutical companies.

Each case is distinguished by the actors involved, particularly the nudgers (government or private agency), and the nature of their interests (pro-social or selfish). To this end, respondents were then asked to make a second choice between two available options (without a default setting) after the disclosure of nudges. This allows us to observe how these differences influence their choice change. In (2), of course, the disclosure and second choice process were not included. See Figure 1 for a flow chart of the experiment and the note below Figure 1 for clarification of terminology.

Figure 1. Flow Chart of the Experiment



Note: Group 0 makes a yes/no decision at Choice 1 under four policy scenarios—healthy food choices, organ donation, COVID-19 vaccination, and viral infectious disease vaccination—**respectively, without a default**. Groups 1, 2, 3, and 4 make a yes/no decision at Choice 1 under the same four scenarios, **respectively, with a default**, and at Choice 2 they either retain or revise their initial Choice-1 decision after disclosure.

Note that our disclosure statements do not include either a *public entity with a pro-self motive* or a *private entity with a pro-social motive*. We acknowledge that this design choice constrains our ability to isolate the independent effects of the information source and its stated motivation. However, we felt that disclosure statements pairing a public entity with a pro-self motive or a private entity with a pro-social motive might appear somewhat unnatural to participants and might be less plausible in realistic policy settings. For example, it would be difficult to imagine a company disclosing, "In order to promote social welfare, we wanted you to choose this option to increase the number of people selecting it." Given these considerations, we chose to restrict the disclosure statements in our experiment to combinations that seemed more natural and policy-relevant—namely, a public entity with a pro-social motive and a private entity with a pro-self motive, corresponding to the "public" and "private" treatments. Of course, future research should explore policy contexts in which such combinations would appear plausible and should experimentally disentangle the source (public vs. private) and motive (pro-self vs. pro-social) components to assess their separate contributions.

Following the aforementioned process, all respondents were requested to respond to the following: (i) numeracy: three numeracy questions, (ii) political trust: an inquiry regarding their trust in the government on a four-point Likert scale and their views on partisanship, (iii) self-respect: six questions concerning the importance of autonomous decision-making (based on the Rosenberg Self-Esteem Scale)[15], (iv) interest in health on a five-point Likert scale , (v) attitude toward organ donation: an inquiry regarding their interest in organ donation on a five-point Likert scale, (vi) experience with COVID-19: a binary question asking whether the respondent has ever had COVID-19, (vii) experience with seasonal influenza: a binary question asking whether the respondent has ever had seasonal influenza, and (viii) questions pertaining to standard personal attributes (age, gender, and income). The objective of these questions was to identify potential covariates that may have influenced individuals' reactions to the disclosure of nudges.

In total, there are 4 policies × [3 disclosure treatments + 1 control (no default nudge)] = 16 experimental conditions. The experiment employs a *mixed-subjects* design: while respondents were randomly assigned to one of the 16 cases, the choice 1 (the first choice) vs. choice 2 (the second choice) aspect was measured within subjects. It is important to note that the values of the covariates (i) through (iii) above are not statistically different in these 16 (=4+3*4) cases. Post hoc analyses of gender distribution and mean age across the 16 experimental conditions confirmed no significant differences. The proportion of male and female participants was balanced ($\chi^2 = 14.0$, $p = 0.527$), and the average age did not differ significantly between conditions ($F(15, N\text{-}16) = 0.416$, $p = 0.975$). Hence the random assignment of experimental treatments was balanced in terms of sociodemographic (results of the post hoc balance tests are shown in **Appendix B**).

It should be noted that the sample sizes differ across treatments. For example, the healthy food control group has 84 participants, whereas the treatment groups often exceed 600 participants. This was done intentionally in order to ensure sufficient statistical power in each disclosure condition, particularly for conducting the logit regression analysis. A power analysis (G*Power 3.1) comparing two independent proportions (α=0.05, power = 0.80; one-sided) indicated a required sample size of 77 per subgroup of group 0 to detect the effect of nudges in inducing people to choose the default option.[16] A power analysis (G*Power 3.1) for a one-sided logistic-regression coefficient test (α = 0.05, power = 0.80; target OR = 1.8) indicated a required sample size of 319 for participants subject to logistic regression in groups 1–4 to detect the factors that prompt individuals who initially chose the default option to revise their choice after disclosure.[17]

Procedure. The respondents completed the questions in an online format at their own convenience. Prior to commencing the study, participants were required to read and sign a consent form and were assured of the anonymity of their data. Once consent had been granted, respondents were asked to respond to a question on their computer screen. Our experiment was approved by and conducted in accordance with the guidelines of the Waseda University Ethical Review Board (application no. 2023-458, March 7, 2024).

Findings

Empirical findings on default nudges and post-disclosure

Tables 1, 2, 3 and 4 provide descriptive statistics of pooled treatment: Table 1 presents the distribution of participants' choices without default nudge treatment; Table 2 reports the proportion with default nudge treatment; Table 3 shows the proportion of participants who initially selected the default option in Table 2 but changed their choice after disclosure. Finally, Table 4 summarizes the final choice proportions.[18]

By comparing the columns "Proportion of A" of Table 1 and Table 2, we can see, first, there were clear effects of nudges in three of four policy choices: participants selected the default choice significantly more often, except in the case of organ donation (while the Chi-squared values were 10.336, 5.145, 4.611, and 3.32, p values are < 0.05, $= 0.70$, < 0.05, and < 0.10, respectively).[19] Second, and more importantly, Table 3 shows that in all policy areas, the disclosure of nudges reduced the number of people who chose the default option under the opt-out condition and who kept their choice in the second opportunity. In particular, the comparison of the columns "Proportion of A" of Table 1 and Table 4 shows that making the default transparent may have resulted in a low proportion of participants who chose option A under the disclosure condition relative to the no-default treatment (Group 0).

For a more detailed look at the conclusions in the previous paragraph, Table 5 compares the proportions of default choices across the three types of disclosure in all policies. Notably, in the vaccination cases, disclosure in the "neutral" manner weakened the nudging effects; however, this weakening effect was smaller compared to disclosure in the "public" or "private" manner.

Table 1, 2, 3 and 4. Distribution of Participant Choices (Pooled Data)

Table 1 Distribution of choices (Group 0)

Policies	Choice without default and disclosure		
	Choose A	Choose N	Prop. of A
Total	200	147	0.58
Healthy food choice	37	47	0.44
Organ donation	47	43	0.52
COVID-19 vaccination	59	31	0.66
Viral infectious disease vaccination	57	26	0.69

Table 2 Distribution of choices (Group 1, 2, 3, 4)

Policies	Choice with default		
	Choose the default (A)	Choose N	Prop. of A
Total	1627	817	0.67
Healthy food choice (Group 1)	355	259	0.58
Organ donation (Group 2)	333	272	0.55
COVID-19 vaccination (Group 3)	479	146	0.77
Viral infectious disease vaccination (Group 4)	460	140	0.77

Table 3 Distribution of choices (Group 1, 2, 3, 4)

Policies	Post-disclosure choice revision		
	Do not change their choice	Change their choices from A to N	Prop. of change
Total	1282	345	0.21
Healthy food choice (Group 1)	240	115	0.32
Organ donation (Group 2)	259	74	0.22
COVID-19 vaccination (Group 3)	389	90	0.19
Viral infectious disease vaccination (Group 4)	394	66	0.14

Table 4 Distribution of choices (Group 1, 2, 3, 4)

Policies	Final decision (default A vs. alternative N)		
	Choose A	Choose N	Prop. of A
Total	1322	1122	0.54
Healthy food choice (Group 1)	255	359	0.42
Organ donation (Group 2)	274	331	0.45
COVID-19 vaccination (Group 3)	394	231	0.63
Viral infectious disease vaccination (Group 4)	399	201	0.66

Note: While A indicates the affirmative sentence, N indicates the negative sentence.

Table 5. Sixteen Treatments Subdivided into Four Policy Areas and Disclosure Conditions

Healthy food choice (Group 1)	Choice 1 with default			Post-disclosure choice revision			Final decision (default A vs. alternative N)		
Disclosure type	Choose the default (A)	Choose N	Prop. of A	Do not change choice	Change choices from A to N	Prop. of change	Choose A	Choose N	Prop. of A
Neutral	121	80	0.60	84	37	0.31	88	113	0.44
Public	114	96	0.54	76	38	0.33	83	127	0.40
Private	120	83	0.59	80	40	0.33	84	119	0.41
Sum	355	259	0.67	240	115	0.32	255	359	0.42
Organ donation (Group 2)	Choice 1 with default			Post-disclosure choice revision			Final decision (default A vs. alternative N)		
Disclosure type	Choose the default (A)	Choose N	Prop. of A	Do not change choice	Change choices from A to N	Prop. of change	Choose A	Choose N	Prop. of A
Neutral	99	101	0.50	77	22	0.22	83	117	0.42
Public	123	77	0.62	96	27	0.22	100	100	0.50
Private	111	94	0.54	86	25	0.23	91	114	0.44
Sum	333	272	0.55	259	74	0.22	274	331	0.45
COVID-19 vaccination (Group 3)	Choice 1 with default			Post-disclosure choice revision			Final decision (default A vs. alternative N)		
Disclosure type	Choose the default (A)	Choose N	Prop. of A	Do not change choice	Change choices from A to N	Prop. of change	Choose A	Choose N	Prop. of A
Neutral	171	44	0.80	151	20	0.12	153	62	0.71
Public	158	50	0.76	117	41	0.26	118	90	0.57
Private	150	52	0.74	121	29	0.19	123	79	0.61
Sum	479	146	0.77	389	90	0.19	394	231	0.63
Viral infectious disease vaccination (Group 3)	Choice 1 with default			Post-disclosure choice revision			Final decision (default A vs. alternative N)		
Viral infectious disease vaccination	Choose the default (A)	Choose N	Prop. of A	Do not change choice	Change choices from A to N	Prop. of change	Choose A	Choose N	Prop. of A
Neutral	152	44	0.78	141	11	0.07	142	54	0.72
Public	150	47	0.76	127	23	0.15	129	68	0.65
Private	158	49	0.76	126	32	0.20	128	79	0.62
Sum	460	140	0.77	394	66	0.14	399	201	0.66

Note: In every table, A means the affirmative choice and N the negative one.

Note: Group 0 makes a yes/no decision at Choice 1 under four policy scenarios—healthy food choices, organ donation, COVID-19 vaccination, and viral infectious disease vaccination—**respectively, without a default**. Groups 1, 2, 3, and 4 make a yes/no decision at Choice 1 under the same four scenarios, **respectively, with a default**, and at Choice 2 they either retain or revise their initial Choice-1 decision after disclosure.

Determinants of choice revision after disclosure

To examine determinants of choice revision after disclosure, we conducted a logistic regression with choice revision as the dependent variable (1 = changed decision after disclosure from A to N; 0 = did not), analyzed separately by policy domain: healthy food choice, organ donation, COVID-19 vaccination, and viral infectious disease vaccination. A significant positive coefficient in Table 6 indicates the variables that drive participants to revise their default choice. We included the following as control variables: gender (1 = male, 0 = female), age, self-respect, political trust, numeracy, interest in health, attitude toward organ donation, experience with COVID-19, and experience with seasonal influenza. The results can be summarized as follows.

First, in the vaccination cases in columns 3 and 4 of Table 6 (a mix of LP/pro-self and LB/pro-social), both the "public" and "private" disclosure conditions had statistically significant positive effects on the revision of choice, whereas the "neutral" disclosure condition (which did not include information about the nudger) showed smaller effects (see the Neutral rows and Post-disclosure choice revision column of Table 5). In the healthy food choice (LP/pro-self) and organ donation (LB/pro-social) cases, the "public" and "private" disclosure conditions did not produce statistically significant effects. This suggests that whether or not the disclosure condition includes information about the nudger (i.e., who the nudger is and his or her intentions or motivations) can significantly affect the effectiveness of nudging, depending on whether the target is LP (pro-self), LB (pro-social), or a combination of both.

Second, our regression analysis consistently shows that the older and more numerically literate people are, the *less likely* they are to revise their choice from the default option after disclosure (in other words, the nudging effects tend to persist after disclosure). This may suggest that the findings of previous studies that nudging effects persist after disclosure in samples with older and/or more educated participants (e.g., Bruns et al., 2018; Loewenstein et al., 2015) were also observed in our experiments, where the demographic attributes of the sample were controlled.

Third, it is important to acknowledge the potential role of ceiling effects, particularly in the vaccination scenarios. As shown in Table 2, while the proportion of participants adhering to the default choice in the COVID-19 and infectious disease vaccination conditions was already quite high (around 77%) before disclosure, the corresponding proportions in the food choice and organ donation scenarios were quite low (58% and 55%, respectively). The default adherence might allow for a wider range of change caused by post-disclosure in the vaccination scenarios.

Table 6. Determinants of Choice Change after Disclosure

term	(1) Healthy food choice		(2) Organ donation		(3) Covid-19 vaccination		(4) Viral infectious disease vaccination	
(Intercept)	−0.850 ***	−0.06	−1.30 ***	−0.31	−2.01 ***	0.10	−2.63 ***	−1.78 +
	−0.200	(0.91)	(0.25)	(1.24)	(0.24)	(0.95)	(0.33)	(1.05)
ByPublic	0.150	0.31	0.03	0.13	0.97 **	1.05 **	0.95 *	1.14 **
	(0.28)	(0.31)	(0.33)	(0.41)	(0.30)	(0.32)	(0.40)	(0.43)
ByPrivate	0.15	0.2	0.06	0.21	0.55 +	0.42	1.28 ***	1.56 ***
	(0.28)	(0.31)	(0.34)	(0.42)	(0.32)	(0.34)	(0.38)	(0.41)
Gender		0.54 *		0.17		0.37		0.47
		(0.25)		(0.33)		(0.26)		(0.30)
Age		−0.02 **		−0.01		−0.02 *		−0.01
		(0.01)		(0.01)		(0.01)		(0.01)
Self_respect		0.08		0.53 +		−0.08		0.18
		(0.21)		(0.30)		(0.22)		(0.24)
Political_trust		−0.05		−0.13		−0.16		−0.43 *
		(0.15)		(0.21)		(0.15)		(0.19)
Numeracy		−0.41 ***		−0.42 **		−0.46 ***		−0.46 **
		(0.12)		(0.16)		(0.13)		(0.15)
Health_interest		0.05						
		(0.17)						
Organ_opinion				−0.97 ***				
				(0.20)				
Covid_exp						−0.36		
						(0.27)		
Flu_exp								−0.54
								(0.53)
Num.Obs.	354	327	332	283	474	435	452	422
AIC	450.5	398.4	355.7	257.4	452.8	408.7	365.2	328.8
BIC	462.1	432.5	367.1	290.2	465.3	445.4	377.5	365.2
Log.Lik.	−222.248	−190.215	−174.865	−119.703	−223.395	−195.370	−179.592	−155.401
F	0.202	3.019	0.018	4.026	5.251	3.978	5.569	3.594
RMSE	0.47	0.44	0.41	0.36	0.39	0.38	0.35	0.34

Discussion

The present study examined whether the effectiveness of default nudges persists after individuals are informed that they were nudged, and whether the identity and motivation of the nudging agent influence post-disclosure choice revision. Building on previous experimental research, we aimed to clarify how transparency shapes behavioral responses. Our results reveal that post-disclosure, particularly when it specifies the nudger's identity and intent, substantially reduces adherence to the default option. This finding contrasts with earlier evidence suggesting that disclosure does not meaningfully diminish nudge effectiveness, indicating that contextual and cultural factors may play an important role. In what follows, we discuss the theoretical and practical implications of these findings, consider their limitations, and suggest directions for future research.

Theoretical and practical implications

Although a large number of studies found no or only negligible evidence that the disclosure of default nudges has a negative impact on default effectiveness (Bruns et al., 2018; Loewenstein et al., 2015; Michaelsen et al., 2020; Steffel et al., 2016; Wachner et al., 2020), our results may suggest otherwise: when individuals are informed of the fact that they were nudged—along with the nudger's identity and intent—*after* their initial choice, they are more likely to revise their original choice. By examining this intriguing result, we can suggest several theoretical and practical implications of the study.

First, we examine why the effects of the nudge were diminished or disappeared even in our "neutral" disclosure. The "neutral" disclosure was designed to align with a treatment typically employed in prior studies

that disclosed only the effects of nudges. The typical wording used in such studies is limited to "We would like to inform you that we intentionally preselected X to encourage more people to choose X." For example, in Michaelsen et al. (2021), the disclosure statement is "to ensure that you fully understood the way the choice to complete the additional questionnaire was set up, we would like to inform you that we intentionally preselected "I AGREE" to encourage more people to help the students." In Paunov, et. al (2019b), it is "Please note the following: we would want you to choose Category C. Therefore, we have preselected this category." In contrast, the "neutral" disclosure in our study was as follows: "In the previous question, 'I will donate my organs after my death' was initially set as the selected option. In fact, from the viewpoint of psychology and cognitive science, it is well known that this way of asking the question leads many people to choose this option." While our initial intention was merely to offer an informative account of the nudge mechanism, our version of the typical phrasing may have inadvertently triggered a different psychological reaction. Rather than merely providing information, the invocation of "psychology and cognitive sciences" may have been perceived as an appeal to scientific authority—suggesting that experts endorse the default option (Milgram, 1963; Cialdini & Goldstein, 2004; Götz et al., 2023). If this inference is correct, it suggests that individuals are highly sensitive to the perceived identity of the agent attempting to influence their decision-making, even when that agent is hypothetical and lacks concrete institutional presence.

From this perspective, two possibilities arise. The "neutral" disclosure may trigger adverse affective responses—such as feeling manipulated or tricked—which can prompt individuals to revise their choices. Alternatively, it may evoke positive affective responses—such as feeling informed or respected—which can reinforce commitment to the initial choice rather than prompt revision. In the prevailing Japanese social climate in March 2024 (the time of the experiment), in which distrust and skepticism toward experts continued to persist in the post-COVID era, we conjecture that the neutral disclosure was more likely to elicit the former reaction, resulting in participants revising their initial choice. [20]

Together with the absence of distinct effects of the "public" and "private" conditions across all four policy areas, this may suggest not only the need to explore what is entailed by the three contexts of disclosure—"neutral," "public," and "private," but also an important policy implication. If, as our findings imply, even "neutral" post-disclosure *alone* induces a substantial proportion of individuals to reconsider their initial (default) choice, it follows that, for individuals who have previously experienced post-disclosure, the spillover effects of nudges identified in prior research (e.g., Van Rookhuijzen et al., 2023)—namely, the persistence of their influence after removal of nudges—may diminish in subsequent similar choice contexts. From a policy efficacy standpoint, our finding suggests that disclosure should preferably be implemented *ex ante*.

Lastly, the fact that most of the participants in key previous studies were Westerners, whereas our experiment involved Japanese participants, may have contributed to the different results regarding the robustness of nudging effects. Japanese people may be more likely than Westerners to feel that their autonomy inherently internalized by their culture is compromised, making them more aware of an interpersonal relationship with the nudger as someone who is manipulating them, and motivating them to alter their choices when the presence of a nudger is disclosed (for a review of this area, see Chirkov et al., 2003). In addition, the clear difference in the transition of choice for organ donation compared to the other three policies - contrary to findings in the other policy areas, we observed no statistically significant effects of default nudges in the organ donation condition- may suggest that choices about organ donation are influenced by culturally different views of life and death. [21]

Limitations

One important limitation concerns the external validity of the post-disclosure and second-choice setup in our experimental design, which was the same as the post-disclosure treatment of Loewenstein et al. (2015). The rapid progression in our design—from the initial decision under default nudges to the second decision after disclosure—may amplify participants' awareness of being nudged, potentially eliciting a negative reaction to the nudge as they are suddenly exposed to manipulation without prior transparency.

Arguably, such rapid progression from an initial default-based decision to a subsequent post-disclosure decision is realistic only in limited domains—such as automatic subscription renewals, where individuals are notified that a default has been applied and are given short windows to revise their choices. However, with the rapid expansion of digital administration and AI-assisted service platforms, revisable choice environments

are likely to become increasingly common, as seen in online tax-filing systems, digital public-benefit applications, and AI-based recommendation interfaces in municipal service portals. These systems also make it easier to implement default nudges, thereby increasing the importance of clear disclosure. We believe that our experimental paradigm is highly relevant to these emerging contexts.

At the same time, as highlighted in recent meta-analyses (e.g., Bruns, 2024), many behavioral experiments rely on online environments and overrepresent default nudges, raising concerns about ecological validity. We realize that our own design shares these characteristics. Because our study employed a hypothetical scenario in an online environment, the "public" and "private" conditions did not entail actual public and private visibility. Consequently, the absence of significant differences between the neutral and public/private disclosures in the context of healthy food choices or organ donation and the absence of distinct effects of the "public" and "private" conditions across all four policy areas may reflect insufficient verisimilitude in the manipulation. Notably, while our findings help isolate people's psychological responses to the disclosure of the nudger's identity and intent, we do not claim that these effects would be readily replicated in the field without further empirical testing. Future research should therefore complement our results with field experiments in applied domains where default nudges are implemented under real decision-making conditions.

Conclusion

Our study demonstrated that post-disclosure, especially when it specifies the identity and motivations of nudgers, can substantially erode default effects, even under neutral framing. Theoretically, this refines existing accounts of nudge transparency by identifying disclosure content as a key stimulus. Practically, it suggests that policymakers should carefully design the timing and substance of disclosure to preserve policy effectiveness while maintaining ethical standards. Because nudges are receiving increased attention as an effective device for changing people's choices without significantly altering economic incentives or limiting options, it is crucial to gather evidence on how the effects of nudges in the form of defaults are influenced by different disclosure conditions that variously reveal the purpose and effects of nudges, and/or the identity of nudgers, in general samples from culturally diverse countries.

Acknowledgment

We would like to express our sincere appreciation to four anonymous reviewers for their insightful and constructive feedback, as well as to the editor for thoughtful guidance and support during the review process. This work was partly supported by JSPS KAKENHI Grant Number JP23K21870 and JP23H00802.

Notes

1. Regarding the distinction between "libertarian paternalism" and "libertarian benevolence," see Inoue et al., 2016. The paper argues that LP- and LB-relevant cases are less common in policy-oriented nudge than might be imagined when Sunstein and Thaler's "condition of convergence" should be satisfied by both LP and LB: the condition requires that for a nudge to be justifiable, individuals' fully informed and rational preferences must converge on the same choice. But situations where nudges are truly justified are actually quite rare, since people's preferences often remain diverse if they are rational: some prioritize long-term welfare, while others rationally favor short-term enjoyment.

2. According to Thaler and Sunstein (2008), when a default option is present—that is, an option that takes effect if the decision-maker remains passive—a substantial proportion of individuals adhere to this default, regardless of whether it serves their best interests. By strategically setting defaults, choice architects can influence behavior in a non-coercive manner, guiding individuals toward outcomes deemed beneficial without restricting their freedom to choose otherwise.

3. In the previous research (Loewenstein et al., 2015; Steffel et al., 2016; Paunov et al..2019b; Michaelsen et al., 2021), for example, some disclosures may take the form of stating that the *nudging agents* are "ourselves (researchers)," and that the purpose of the default is to make it easier for "you (participants)" to select the preselected option. They do not, however, explicitly present *nudging agent's identity and motivations* within political and social contexts.

4. However, this is not the only plausible pathway. In some cases, disclosure may enhance the effectiveness of a nudge—particularly when it signals that the default option is endorsed by experts, policymakers, or trusted institutions. Moreover, disclosure may sometimes have no measurable impact, especially if recipients are inattentive, already suspect default influence, or find the disclosure irrelevant to their goals.

5. Michaelsen et al. (2020) also shows that the transparent opt-out (pre-disclosure) was perceived as less fair than the undisclosed opt-out default when the nudge was experienced first-hand in an actual choice.

6. According to Weijers et al. (2024), experienced autonomy refers to the level of autonomy actually perceived by individuals after exposure to a nudge intervention, whereas expected autonomy denotes the degree of autonomy anticipated before encountering a nudge intervention, particularly regarding concerns about potential autonomy reduction.

7. Only a minority of studies have disclosed both elements (e.g., Bruns, 2018; Michaelsen et al. , 2021; Steffel et al., 2016).

8. In democratic societies, when the nudge is implemented as a public or private policy, the identity of the agent and his or her intentions and motivations are ethical and critical pieces of information, because citizens have an expectation that such information will be available to them. It is plausible to assume that even if the contents of nudges are the same, the way the nudge will be perceived by nudgees may change depending on the nudging agent.

9. To highlight a key feature of our post-disclosure experiment, it is helpful to use two terms introduced by Bovens (2008): type-level knowledge (general awareness that nudging occurs in policy contexts) and token-level knowledge (awareness that this particular choice was subject to specific manipulation). Our experimental design targets the latter: it tests whether explicitly informing individuals of the fact that they were nudged—and by whom and for what purpose—changes their decision. This is distinct from the broader question of whether people suspect they are being nudged during decision-making.

10. To illustrate, Loewenstein et al. (2015) prompt individuals to consider either the "prolong option" or the "comfort option" in relation to end-of-life preferences. However, it remains uncertain whether these choices ultimately benefit the respondents or society. In this sense, these nudges cannot be classified as policies based on LP or LB. Similarly, Weijers et al. (2024) implemented a default nudge to increase food donations in supermarkets. Although this nudge is pro-social, it does not align with LB *in a strict sense*, as participants incur costs when making food donations.

11. Conditions in which a public entity nudges for private benefit or a private entity nudges for public benefit were not implemented as experimental conditions, because these scenarios are either unnatural or rare in real-world nudging contexts. An illustration of the latter scenario would be the implementation of nudges by NGOs or NPOs with the objective of enhancing the well-being of particular disadvantaged groups.

12. It should be noted that studies employing pre- and post-disclosure designs are not fully comparable without caveats (a point raised by an anonymous reviewer). In particular, to ensure comparability, an experimental design would need to hold all other conditions constant and manipulate only the timing of disclosure (pre-disclosure versus post-disclosure). Absent such a design, differences observed across studies may reflect factors other than the timing of disclosure itself. Since no such comparisons are intended here, the present research should be regarded as exploratory. This characterization is warranted because prior evidence on the effects of post-disclosure on autonomy remains limited and mixed, making a systematic evaluation challenging. Our study therefore focuses on identifying potential patterns and generating insights, providing a necessary foundation for future research that may formulate and test more definitive hypotheses.

13. Of the 3,138 individuals who agreed to participate in the survey, 236 discontinued participations before completing all items, and among those who completed all responses, 111 failed to correctly answer a simple attention-check question ("What year and month is it today?"). Thus, the final analytical sample consisted of 2,791 respondents after excluding these cases.

14. By verifying completion times, we identified a few rapid completions that may indicate inattentive responses. To address this concern, we conducted supplementary analyses excluding responses with durations below 60 seconds (seven participants) as a robustness check. The results were broadly consistent with those reported in the main text, and the substantive conclusions remained unchanged.

15. A substantial body of psychological research, conducted within the framework of Self-Determination Theory, has demonstrated that the satisfaction of the need for autonomy is a significant predictor of high self-esteem. For further details, see our supplementary material.

16. Power analysis was conducted in G*Power (version 3.1) with the following settings: test family = Z tests; statistical test = Proportions: Difference between two independent proportions (two groups); type = A priori; tail(s) = One; expected proportions $p_1 = 0.60$ and $p_2 = 0.40$; $\alpha = 0.05$; power $(1-\beta) = 0.80$; allocation ratio $N_2/N_1 = 1$.

17. Power analysis was conducted in G*Power (version 3.1) with the following settings: test family = Z tests; statistical test = Logistic regression; type = A priori; tail(s) = One; odds ratio under H_1 = 1.8; baseline event rate p_0 = 0.50; R_2 (other X with X) = 0.10; predictor distribution = Binomial; prevalence π= 0.50.

18. The discrepancy between the number of participants who selected option A in Table 2-4 and the number of participants in Table 2-3 who did not change their default choice arises because a small number of participants initially did not choose the default option but switched their choice after disclosure. In total, 40 out of 817 participants exhibited this behavior. If we exclude these participants, the essential results of the statistical analysis are not changed. But further investigation is warranted to understand why this phenomenon occurred.

19. It is worth noting that, contrary to findings in the other policy areas, we observed no statistically significant effects of default nudges in the organ donation condition. This result stands in contrast to substantial empirical evidence showing that opt-out policies and presumed consent significantly increase organ donation consent rates (e.g., Johnson & Goldstein, 2003; Abadie & Gay, 2006). Several factors may account for this divergence. Notably, cultural and institutional context may matter: Japan has historically had lower rates of organ donation and more cautious attitudes toward brain death and posthumous bodily intervention (Kato, 2013). In Japan, informed consent is legislated but presumed consent is not. Such Japanese contexts may make the nudge less effective or socially acceptable.

20. Recent evidence suggests distrust and skepticism toward experts in the post-COVID period. Sato et al. (2024) show that the estimated prevalence of holding at least one COVID-19 conspiracy belief was 24.4% in Japan. Cologna et al. (2025) report that Japan's mean trust in scientists (3.37 on a 1–5 scale) was significantly below the global mean of 3.62.

21. In the context of organ donation, the observed patterns may reflect not only cultural differences but also differences in the underlying legal frameworks. In Japan, the Act on Organ Transplantation requires not only the donor's stated intention to donate, but also explicit consent from the donor's family members. In short, organ donation is treated as a *family decision* rather than a purely *personal* one. By contrast, in the United States organ donation is generally regarded as an individual choice, with family consent not legally required once the donor's intention has been recorded. Consequently, default nudges may be less effective in Japan because the legally mandated role of family consent can moderate or counteract individual-level shifts induced by the nudge.

References

Abadiea, A., & Gayb, S. (2006). The impact of presumed consent legislation on cadaveric organ donation: A cross-country study, *Journal of Health Economics*, 25 (4), 599–620.

Bovens, L. (2008). The ethics of nudge. In M. J. Hansson & T. Grüne-Yanoff (eds.), *Preference Change: Approaches from Philosophy, Economics and Psychology*. Springer, pp. 207–220.

Bruns, H., Kantorowicz-Reznichenko, E., Klement, K., Jonsson, M. L., & Rahali, B. (2018). Can nudges be transparent and yet effective? *Journal of Economic Psychology*, 65 (C), 41–59.

Cialdini R. B. & Goldstein, N. J. (2004). Social influence: compliance and conformity. *Annual Review of Psychology,* 55 (1), 591–621.

Chirkov, V., Ryan, R. M., Kim, Y., & Kaplan, U. (2003). Differentiating autonomy from individualism and independence: a self-determination theory perspective on internalization of cultural orientations and well-being. *Journal of Personality and Social Psychology*, 84 (1), 97–110.

Cologna, V., ... , Rolf A. Zwaan. (2025). Trust in scientists and their role in society across 68 countries. *Nature Human Behaviour,* 9, 713–730.

Congiu, L., & Moscati, I. (2022). A review of nudges: Definitions, justifications, effectiveness. *Journal of Economic Surveys*, 36 (1), 188–213.

Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. *Handbook of Theories of Social Psychology*, 1(20), 416–436.

Götz, F.J., Mitschke, V. & Eder, A.B. (2023). Conflict experience and resolution underlying obedience to authority. *Scientific Report*, 13 (11161), 1–13.

Grüne-Yanoff, T. (2012) Old wine in new casks: libertarian paternalism still violates liberal principles. *Social Choice and Welfare*, 38 (4), 635–645.

Hagman, W., Andersson, D., Västfjäll, D. & Tinghög, G. (2015). Public views on policies involving nudges. *Review of Philosophy and Psychology,* 6 (3), 439–453.

Hansen, P. G., & Jespersen, A. M. (2013). Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy. *European Journal of Risk Regulation*, 4 (1), 3–28.

Inoue, A., Shimizu, K., Udagawa, D., & Wakamatsu, Y. (2016). How broad is the scope of Sunstein's and Thaler's theory? *Australian Journal of Legal Philosophy* , 41, 1–14.

Johnson, E. J., & Goldstein, D. G. (2003). Do defaults save lives? *Science*, 302 (5649), 1338–1339.

Kato, Y. (2013). Conscience in health care and the definitions of death. *Croatian Medical Journal*, 54 (1), 75–77.

Kroese, F. M., Marchiori, D. R., & De Ridder, D. T. (2016). Nudging healthy food choices: a field experiment at the train station. *Journal of Public Health*, 38 (2), e133–e137.

Loewenstein, G., Bryce, C., Hagmann, D., & Rajpal, S. (2015). Warning: You are about to be nudged. *Behavioral Science & Policy*, 1 (1), 35–42.

Michaelsen P., (2024). Transparency and nudging: An overview and methodological critique of empirical investigations. *Behavioural Public Policy*, 8 (4), 807–817.

Michaelsen, P., Nyström, L., Luke, T. J., Johansson, L. O., & Hedesström, M. (2020). Are default nudges deemed fairer when they are more transparent? People's judgments depend on the circumstances of the evaluation, PsyArXiv Preprints.

Michaelsen, P., Nyström, L., Luke, T. J., & Hedesström, M. (2021). Downstream consequences of disclosing defaults: Influences on perceptions of choice architects and subsequent behavior. *Comprehensive Results in Social Psychology*, 5 (1–3), 25–48.

Michaelsen, P., Johansson, L. O., & Hedesström, M. (2024). Experiencing default nudges: autonomy, manipulation, and choice-satisfaction as judged by people themselves. *Behavioural Public Policy*, 8 (1), 85–106.

Milgram, Stanley (1963). Behavioral Study of Obedience. *Journal of Abnormal and Social Psychology*, 67 (4), 371-378

Paunov, Y., Wänke, M., & Vogel, T. (2019a). Transparency effects on policy compliance: Disclosing: how defaults work can enhance their effectiveness. *Behavioural Public Policy*, 3 (2), 187–208.

Paunov, Y., Wänke, M., & Vogel, T. (2019b). Ethical defaults: Which transparency components can increase the effectiveness of default nudges? *Social Influence*, 14 (3–4), 104–116.

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology*, 34 (2), 243–281.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55 (1), 68–78.

Sato, Y., Kawachi, I., Saijo, Y., Yoshioka, E., Osaka, K., & Tabuchi, T. (2024). Correlates of COVID-19 conspiracy theory beliefs in Japan: A cross-sectional study of 28,175 residents. *PLoS ONE*, 19 (12), 1–17.

Steffel, M., Williams, E. F., & Pogacar, R. (2016). Ethically deployed defaults: Transparency and consumer protection through disclosure and preference articulation. *Journal of Marketing Research*, 53 (5), 865–880.

Sunstein, C. R., & Thaler, R. H. (2003). Libertarian paternalism is not an oxymoron. *University of Chicago Law Review,* 70 (4), 1159–1162.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press.

Van Rookhuijzen, M., De Vet, E., & Adriaanse, M. A. (2023). The effect of transparency on the temporal spillover effect of default nudges. *British Journal of Social Psychology*, 62 (3), 1363–1375.

Wachner, J., Adriaanse M., & De Ridder, D. (2020). The influence of nudge transparency on the experience of autonomy. *Comprehensive Results in Social Psychology*, 5 (1–3), 49–63.

Wachner, J., Adriaanse M., & De Ridder, D. (2021). The effect of nudges on autonomy in hypothetical and real life settings. *PLoS ONE*, 16(8), 1–16.

Weijers, R. J., Wachner, J., & de Koning, B. B. (2024). The effect of a default nudge on experienced and expected Autonomy: A field study on food donation. *Journal of Behavioral Decision Making*, 37 (4),1–14.

Appendix A: Complete script of questionnaire

After agreeing to the informed consent procedures for participation in the experiment, participants provide a response to only one of the sixteen experimental routes as follows:

Participants in Group 0 were assigned to one of A-0, B-0, C-0, or D-0;
after replying to A-1, participants in Group 1 to one of A-1-1, A-1-2, or A-1-3;
after replying to B-1, participants in Group 2 to one of B-1-1, B-1-2, or B-1-3;
after replying to C-1, participants in Group 3 to one of C-1-1, C-1-2, or C-1-3; and
after replying to D-1, participants in Group 4 to one of D-1-1, D-1-2, or D-1-3.

After addressing the aforementioned questions, respondents proceed to answer inquiries related to numeracy, political trust, party allegiance, self-esteem, and typical demographic characteristics (such as age, gender, and income), along with control-related questions.

Group 0

A-0: Please answer the question based on the following situation.
You are about to ask for a set meal in a casual dining. There are two types of set meals, low-salt and regular. You have been asked the following question by the restaurant staff: "Is a low-salt set meal okay with you?"
Which one will you order?
□I will order a low-salt set meal.
□I will order a regular set meal.
*Note: Participants assigned to A-0 are required to respond solely to this question.

B-0: Please answer the question based on the following situation.
You are asked to fill out an organ donor card.
Which one will you choose?
□I will donate my organs after my death.
□I will not donate my organs after my death.
*Note: Participants assigned to B-0 are required to respond solely to this question.

C-0:
Coronavirus, which is highly contagious and has a high mortality rate, is reappearing throughout Japan. However, the risk of infection can be greatly reduced by vaccination, the safety of which has been medically confirmed. Yet, it can be accompanied by severe side effects (e.g. high fever and headache that lasts for one or two days after the vaccination). This vaccination is free of charge.
Which one will you choose?
□I will be vaccinated.
□I will not be vaccinated.
*Note: Participants assigned to C-0 are required to respond solely to this question.

D-0:
A viral infectious disease, which is highly contagious and has a high mortality rate, is prevalent throughout Japan. However, the risk of infection can be greatly reduced by vaccination, the safety of which has been medically confirmed. Yet, it can be accompanied by severe side effects (e.g. high fever and headache that lasts for one or two days after the vaccination). This vaccination is free of charge.
Which one will you choose?
□I will be vaccinated.
□I will not be vaccinated.
*Note: Participants assigned to D-0 are required to respond solely to this question.

Group 1

A-1: Please answer the question based on the following situation.
You are about to ask for a set meal in a casual dining. There are two types of set meals, low-salt and regular. You have been asked the following question by the restaurant staff: "Is a low-salt set meal okay with you?"

☑I will order a low-salt set meal.
(If you do not want to order a low-salt set meal, please uncheck□.)
*Note: Respondents who answered A-1 are required to respond to one of the following questions.

A-1-1(Neutral):
In the previous question, "I will order a low-salt meal" was initially set as the selected option. In fact, from the viewpoint of psychology and cognitive science, it is well known that this way of asking the question leads many people to choose this option. Now we would like to ask you a question. Given this, which option would you choose?
□I will order a low-salt set meal.
□I will order a regular set meal.

A-1-2 (Public):
In the previous question, "I will order a low-salt meal" was initially set as the selected option. In fact, from the viewpoint of psychology and cognitive science, it is well known that this way of asking the question leads many people to choose this option.
And this way of asking the question is for the following reason: in order to promote the maintenance of your health conditions, the government induced your answer to choose a low-salt set meal, for the policy of increasing the number of people who choose this option. Now we would like to ask you a question. Given this, which one would you choose?
□I will order a low-salt set meal.
□I will order a regular set meal.

A-1-3 (Private):
In the previous question, "I will order a low-salt meal" was initially set as the selected option. In fact, from the viewpoint of psychology and cognitive science, it is well known that this way of asking the question leads many people to choose this option.
And this way of asking the question is for the following reason: in order to make more profit, the restaurant induced your answer to choose a low-salt set meal, for the policy of increasing the number of people who choose this option because it is cheaper to make the low-salt set meal. Now we would like to ask you a question. Given this, which one would you choose?
□I will order a low-salt set meal.
□I will order a regular set meal.

Group 2

B-1: Please answer the question based on the following situation.
You are asked to fill out an organ donor card.

☑I will donate my organs after my death.
(If you do not want to donate your organs, please uncheck□.)
*Note: Respondents who answered B-1 are required to respond to one of the following questions.
B-1-1 (Neutral):
In the previous question, "I will donate my organs after my death" was initially set as the selected option. In fact, from the viewpoint of psychology and cognitive science, it is well known that this way of asking the question leads many people to choose this option. Now we would like to ask you a question. Given this, which option would you choose?
□I will donate my organs after my death.
□I will not donate my organs after my death.

B-1-2(Public):
In the previous question, "I will donate my organs after my death" was initially set as the selected option. In fact, from a psychological or cognitive science point of view, it is well known that this way of asking the question makes many people tend to choose this option. And the reason the question was asked this way is as follows. In order to promote the overall social benefits, the government wanted you to choose this option, to increase the number of people who choose it, since everyone could one day need an organ donation. Now we would like to ask you a question. Given this, which option would you choose?
□I will donate my organs after my death.
□I will not donate my organs after my death.

B-1-3(Private):
In the previous question, "I will donate my organs after my death" was initially set as the selected option. In fact, from a psychological or cognitive science point of view, it is well known that this way of asking the question makes many people tend to choose this option. And the reason the question was asked this way is as follows. People on the waiting list for organs wanted you to choose this option, in order to increase the number of people who choose it, and they lobbied politicians to make this happen. Now we would like to ask you a question. Given this, which option would you choose?
□I will donate my organs after my death.
□I will not donate my organs after my death.

Group 3
C-1 :
Coronavirus, which is highly contagious and has a high mortality rate, is reappearing throughout Japan. However, the risk of infection can be greatly reduced by vaccination, the safety of which has been medically confirmed. Yet, it can be accompanied by severe side effects (e.g. high fever and headache that lasts for one or two days after the vaccination). This vaccination is free of charge.
Which one will you choose?
☑I will be vaccinated.
(If you do not want to be vaccinated, please uncheck□.)
*Note: Respondents who answered C-1 are required to respond to one of the following questions.

C-1-1(Neutral):
In the previous question, "I will be vaccinated" was initially set as the selected option. In fact, from the viewpoint of psychology and cognitive science, it is well known that this way of asking the question leads many people to choose this option. Now we would like to ask you a question. Given this, which one would you choose?
□I will be vaccinated.
□I will not be vaccinated.

C-1-2 (Public):
In the previous question, "I will be vaccinated" was initially set as the selected option. In fact, from the viewpoint of psychology and cognitive science, it is well known that this way of asking the question leads many people to choose this option.
And this way of asking the question is for the following reason: in order to promote the overall social benefits, the government induced your answer to choose vaccination, to increase the number of people who choose vaccination. Now we would like to ask you a question. Given this, which one would you choose?
□I will be vaccinated.
□I will not be vaccinated.

C-1-3(Private):
In the previous question, "I will be vaccinated" was initially set as the selected option. In fact, from the viewpoint of psychology and cognitive science, it is well known that this way of asking the question leads many people to choose this option.
And this way of asking the question is for the following reason: the pharmaceutical companies that produce the vaccine have induced your answer to choose vaccination, to increase the number of people who choose vaccination by lobbying politicians. Now we would like to ask you a question. Given this, which one would you choose?
□I will be vaccinated.
□I will not be vaccinated.

Group 4
D-1 :
A viral infectious disease, which is highly contagious and has a high mortality rate, is prevalent throughout Japan. However, the risk of infection can be greatly reduced by vaccination, the safety of which has been medically confirmed. Yet, it can be accompanied by severe side effects (e.g. high fever and headache that lasts for one or two days after the vaccination). This vaccination is free of charge.
Which one will you choose?
☑I will be vaccinated.
(If you do not want to be vaccinated, please uncheck□.)
*Note: Respondents who answered D-1 are required to respond to one of the following questions.

D-1-1(Neutral):
In the previous question, "I will be vaccinated" was initially set as the selected option. In fact, from the viewpoint of psychology and cognitive science, it is well known that this way of asking the question leads many people to choose this option. Now we would like to ask you a question. Given this, which one would you choose?
□I will be vaccinated.
□I will not be vaccinated.

D-1-2 (Public):
In the previous question, "I will be vaccinated" was initially set as the selected option. In fact, from the viewpoint of psychology and cognitive science, it is well known that this way of asking the question leads many people to choose this option.
And this way of asking the question is for the following reason: in order to promote the overall social benefits, the government induced your answer to choose vaccination, to increase the number of people who choose vaccination. Now we would like to ask you a question. Given this, which one would you choose?
□I will be vaccinated.
□I will not be vaccinated.

D-1-3(Private):
In the previous question, "I will be vaccinated" was initially set as the selected option. In fact, from the viewpoint of psychology and cognitive science, it is well known that this way of asking the question leads many people to choose this option.
And this way of asking the question is for the following reason: the pharmaceutical companies that produce the vaccine have induced your answer to choose vaccination, to increase the number of people who choose vaccination by lobbying politicians. Now we would like to ask you a question. Given this, which one would you choose?
□I will be vaccinated.
□I will not be vaccinated.

Each participant answered all of the questions listed below.

Numeracy

- The price of one baseball bat and one ball together is 1100 yen. If the price of the bat is 1000 yen more than the price of the ball, what is the price of the ball?
- It takes 5 minutes to make 5 pieces of a certain product using 5 machines. Then, if 100 machines are used to make 100 products, how many minutes will it take?
- Part of a pond is covered with water lilies, and the area doubles every passing day. If it takes 6 days for the entire pond to be covered with water lilies, how many days will it take for half of the pond to be covered?

Political awareness

How trustworthy do you think your country's political system is?
(Options) trustworthy, somewhat trustworthy, somewhat untrustworthy, untrustworthy

Party support

Apart from which party you vote for in elections, which party do you usually support? Please select only one.

Liberal Democratic Party
Constitutional Democratic Party of Japan, Japan Innovation Party, Komeito Party, Japanese Communist Party, Kokumin Democrat Party, Reiwa Shinsei Gumi, Realizing Free Education Party, Social Democratic Party, Sansei Party (Minna de Tsukuru Party), Other political parties, and I do not support any political party.

Self-respect (autonomy)

- My pride will not allow others to decide what concerns me.
- You do not have to decide for yourself, even if it concerns you, if the result is for your own good.
- If you are a full-fledged human being, you should decide for yourself.
- I would rather grit my teeth and endure than rely on the mercy of others.
- I do not want to show weakness to others.
- It is easier and better to have an expert who knows more than you decide for you, even if it is about yourself.
- (5-point scale from "very much agree" to "not at all agree")

Demography

- What month of what year is today? (pull down answers)
- What is your occupation?
- For all those who answered "Student" to "What is your occupation?"
    - Which of the following is the number of all people who belong to the school you attend?
    - Which of the following describes the number of people who usually attend your school campus?
- For those who answered "Work" to "What is your occupation?
    - Which of the following is the number of people who belong to the company you work for?
    - Which of the following describes the number of people who usually go to your place of work?

For all of them

- What is the annual after-tax income of your entire family?
- Which of the following describes your own annual after-tax income?

Questions for control

- Have you ever been infected with coronavirus from January 2020 (when the first COVID cases were discovered in Japan) to the present? (a binary question)
- Have you ever been infected with influenza from January 2020 (when the first COVID cases were detected in Japan) to the present? (a binary question)
- Are you interested in organ transplantation? (5-point scale from "very much agree" to "not at all agree")
- How concerned are you about your health? (5-point scale from "very much agree" to "not at all agree")

Appendix B: Balance-test results

Group	Policy	Type of Disclosure	male	age		obs.
				ave.	(std.)	
0	Healthy food choice	Null	0.405	48.3	(16.6)	84
	Organ donation		0.444	48.9	(16.9)	90
	COVID-19 vaccination		0.533	50.9	(15.9)	90
	Viral infectious disease vaccination		0.482	51.3	(15.1)	83
1	Healthy food choice	Neutral	0.443	50.0	(15.9)	201
		Public	0.527	49.9	(16.9)	203
		Private	0.495	50.1	(16.1)	210
2	Organ donation	Neutral	0.480	51.6	(17.1)	200
		Public	0.434	50.5	(16.7)	205
		Private	0.480	49.8	(17.0)	200
3	COVID-19 vaccination	Neutral	0.502	49.5	(16.6)	215
		Public	0.554	51.0	(15.9)	202
		Private	0.514	50.3	(16.6)	208
4	Viral infectious disease vaccination	Neutral	0.495	49.5	(16.2)	196
		Public	0.512	49.4	(16.7)	207
		Private	0.497	51.1	(16.6)	197
						2791
	gender: Chi2=14.0, p=0.527					
	age: F-statistic=0.416, p=0.975					