Erasmus
School of
Economics

Public Perception of Nudges in Public Policy, The Effect of Transparency

MSc thesis Behavioural Economics (Strategy Track)

Vivianne Courte-Rathwell (615164)

Supervisor: Han Bleichrodt
Second assessor: Sophie van der Zee

April 30th, 2023

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

TABLE OF CONTENTS

ABSTRACT

Nudges have gained popularity among policymakers as a tool to influence behaviour. However, critics draw similarities between nudges and manipulation, especially with regards to transparency. This study aims to explore the impact of transparency on public opinion towards nudges implemented by public policymakers. Specifically, on the four types of nudges identified by Hansen and Jespersen (2013). A survey was distributed to participants in a control and a treatment group to test four hypotheses. They propose that citizens who receive complete information on nudges from public officials will be more in agreement with nudges overall (H1), citizens are least in agreement with Non-Transparent System 2 nudges (H2), are more in agreement with transparent nudges than non-transparent nudges (H3) and are more in agreement with System 2 nudges than System 1 nudges (H4).

The study found that citizens are generally in favour of nudges, with the highest agreement rate found for Transparent System 1 nudges. Providing information on the intent and method behind nudges had a positive effect on the acceptance of nudges. The results suggest that citizens value autonomy and trust in decision-making, actions, and choices.

The study's contribution to the field of behavioural economics can inform policy decisions about the use of nudges in the future. The study has some limitations, such as the use of a survey to conduct an experiment and the lack of examination of potential confounding variables. Despite these limitations, the results of this study have important implications for policymakers interested in implementing nudges in public policy.

ACKNOWLEDGEMENTS

I express my heartfelt gratitude to everyone who extended their support and dedicated time to thoroughly read my thesis and provide valuable advice. My wish was to create a work that was accessible, interesting and impactful. I am grateful to my supervisor Han for his unwavering support and guidance throughout this process. Lennard, I appreciate you for keeping me on track and focused. Shoutout to Parker for being an example of hard work. I am also grateful for Pisita and Concha, their emotional support and friendship helped me stay positive. Special thanks to Dale and Johanne for providing constructive feedback, which greatly improved the quality of my thesis.

INTRODUCTION

Nudges have become a popular tool used by public officials to influence behaviour. However, concerns about their potential for manipulation have raised questions about their transparency. This study will explore the impact of nudge transparency on public opinion, with a focus on the four types of nudges identified by Hansen and Jespersen (2013). The research question is whether the transparency of a nudge, implemented by public officials, impacts the public's opinion on nudges.

To answer this question, a survey was distributed to participants split into control and treatment groups. Four hypotheses were tested.

H1. Citizens who receive complete information on nudges from their public officials will be more in agreement with nudges overall.

H2. Citizens are least in agreement with the implementation of Non-Transparent System 2 nudges.

H3. Citizens are more in agreement with the implementation of transparent nudges than non-transparent nudges.

H4. Citizens are more in agreement with the implementation of System 2 nudges than System 1 nudges.

The significance of this research lies in its potential to inform future policy decisions and improve the effectiveness and support of nudging in the public sphere. By understanding the impact of transparency on public opinion, policymakers can establish a framework for the use of nudges ethically and democratically.

The thesis begins with a literature review, divided into two sections. The first section will cover the theory behind nudges, the mechanisms employed to create them, and the normative consequences of introducing nudges in public policy. The second section is an empirical

review that will cover the scope and methods used in the survey that was conducted to collect the results for this study. The next chapters describe the methodology and results of the survey. Finally, a discussion covering the implications of the findings and how they may contribute to nudging in policy design guidelines.

In conclusion, this study aims to contribute to the ongoing debate about nudging and its potential for manipulation. The study's contribution to the field of behavioural economics will provide insights into the impact of transparency on public opinion, which can inform policy decisions about the use of nudges in the future.

LITERATURE REVIEW

The objective of the following literature review is to understand the main concerns and questions that exist on the topic of nudging. Previous research will be reviewed in two sections, based on their contribution to this paper's study. First, the theoretical review encompasses the description of a nudge, as well as the normative elements that public policymakers face when they implement nudges. This section will introduce and justify the relevance of my research question. Second, the empirical review will limit the scope of the research question by targeting a sample of previous studies on public sentiment toward nudges. This section will also discern the framework of my empirical research.

Theoretical Review

Dual Process Theory

Dual process theory is a concept that serves as a foundation for this paper and will be recurring throughout. The critically acclaimed "Thinking Fast and Slow" by David Kahneman (Kahneman, 2011) propelled dual process theory into mainstream channels, disseminating the paradigm that our minds encompass two distinct thought processing systems.

System 1 hosts intuitive thought processes, which generate the decisions that are made automatically and swiftly. System 2 hosts deliberative thought processes, which generate deliberate, controlled, and ultimately slower judgments. Whilst System 1 is governed by habit, rendering it difficult to control or modify, System 2 is governed by rules and is relatively flexible. Where System 1 operations are effortless and emotional, System 2 operations are effortful and self-controlled. In theory, an outside observer can assess which system is at the helm by administering concurrent cognitive tasks; intuitive processes will effortlessly coexist whereas simultaneous deliberative processes experience disruption (Kahneman, 2003).

Kahneman & Tversky's work (2003) focused on the errors of intuitive thinking. One reason for this was their astonishment at the palpable discrepancy between statistical knowledge and statistical intuition discernable in themselves and their fellow doctorate colleagues. Another reason for focusing on System 1 errors is ". . . [T]heir value as diagnostic indicators of cognitive mechanisms" (Kahneman, 2003). In other words, these two authors found that empirical evidence of inaccuracy in intuition shed light on how fast thinking works.

Biases, Heuristics & Optimization

In "Thinking Fast and Slow" Kahneman (Kahneman, 2011) presents various examples of each system's operational fallacies. The purpose of listing them here is to give the unacquainted reader more perspective on what the errors of intuitive thinking may be.

System 1 displays features such as a bias to believe and confirm; a tendency to exaggerate emotional consistency; computing more than anticipated; replacing a difficult question with easier ones; being more sensitive to losses than gains; overestimating low probabilities; showing diminishing sensitivity to quantity. In the case of System 2 decisions, biases occur when we must act or commit to those decisions. Furthermore, in making those System 2

decisions, frequently, there is uncertainty- a factor for which intuitive thinking tends to take over (Kahneman, 2003).

Essentially, these errors occur because System 1, to flex its effortless speed, employs rules of thumb called heuristics. These heuristics are embedded in our System 1 processes, which is the reason why we can observe, and predict, biases systematically. In other words, because we know about heuristics, we can predict when they will lead to a bias and when they won't.

With this knowledge, it becomes possible to create environments that utilize the effects of biases to lead to a desired outcome. There is a misconception that biases are laws of human behaviour that are continuously generating poor choices and suboptimal behaviour (Smets, 2018). In reality, biases are isolated labels that describe a tendency in behaviour or choices that defy the assumptions of traditional economics (Smets, 2018).

Public policy and traditional economics have long been intertwined, with the latter serving as a normative guide for the former. This is why in behavioural economics the focus is on biases that may exist in the decision-making processes of individuals and organizations. It is widely recognized that biases can prevent us from making optimal decisions and achieving the ideal outcomes as defined by traditional economics (Kahneman, 2011). As a result, correcting these biases has become a natural part of the process of optimizing our systems and achieving the desired outcomes.

Nudge Design

Nudges are intentional designs in decision-making contexts that attempt to alter people's behaviour and choices in a predictable way (Thaler & Sunstein, 2008). This alteration is motivated by the existence of biases in individual decision making which prevent people from making rational decisions (Hansen, 2016).

Nudges do not forbid any existing options. In other words, nudges are designed to promote a behaviour or a choice rather than removing its alternative. The promoted behaviour or choice is selected according to the objectives and values of the designer, in the context of this paper the hypothetical designers are public policymakers. Nudges do not significantly change economic incentives, meaning that a nudge does not significantly change prices, costs, penalties, or rewards to alter behaviour. This is the charter provided by Sunstein and Thaler (2008). Each aspect must be complied with for an intervention to be considered a nudge, as opposed to traditional regulation. Other advantages provided by the authors are cost-effectiveness and efficiency (Thaler & Sunstein, 2008).

Nudging methods are geared to leverage heuristics to the advantage of the designer and their objectives. A nudge will alter the environment in such a way that when intuitive thinking is triggered, the ensuing heuristic will beget the desired behaviour or choice. It's the predictability of the heuristic that leads to the inference of "manipulation" by critics. As a result, the decisions of the affected individuals remain the object of their own System 1, but now the outcome is prearranged by the organization of the environment. Note, that nudges do not eliminate bias. Alternatively, nudges harness the existence of biases to push people into making better decisions, as judged by the designer (Hansen, 2016; Hansen & Jespersen, 2013). In fact, in response to critics, Sunstein deplored that the description of nudges as an exploitation of biases is negatively misleading. Yet, he did not explicitly disprove the description. Instead, he simply provided "take account of" [biases] as a more accurate term than "exploitation" (Sunstein, 2017).

Ideological Implications

Sunstein and Thaler named their book "Nudge: Improving Decisions About Health, Wealth, and Happiness" (2008) and they remain true to that title in their work, proposing various

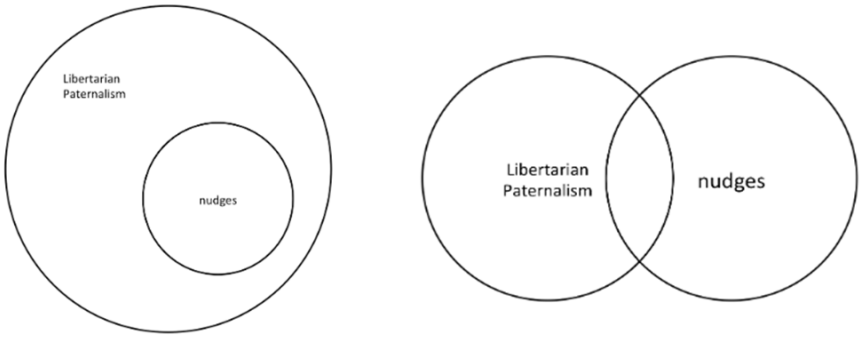
policies that use nudges to make people better off, as judged by themselves. The authors introduce a political doctrine called Libertarian Paternalism. This oxymoron expresses the combination of designing environments to promote behaviour while respecting the conditions of freedom of choice. Sunstein, a renowned legal scholar, and Thaler, a Nobel-winning economist, explain that their doctrine is paternalist because nudges are meant to make people better off. Yet it is also libertarian because nudges are not mandates nor do they eliminate options, thus preserving freedom of choice (Thaler & Sunstein, 2008).

The recent integration of nudging policies in various public spheres has prompted a multitude of inquiries and elicited substantial controversy and criticism. One of these issues is the lack of clarity surrounding the relationship between nudging and Libertarian Paternalism. Since the research question of this paper is focused on policy design in the public sphere, the political implications of nudging cannot be overlooked. Thus, a clear distinction must be made between nudging and political ideologies.

According to Hansen, the definition of nudges can be detached from Libertarian Paternalism (Hansen, 2016), which is significant because it enables a more extensive discussion beyond political doctrine. Keeping the definition within the bounds of behavioural sciences allows for a normative discussion focused on the mechanisms of nudges. The doctrine does not define a scientific tool, but instead, it defines how and if the tool should be used. Additionally, this approach allows for the understanding of nudges to be integrated with other types and parameters of policy design, which may enhance its efficiency, and applicability, and make its usage complementary to existing policy tools (see Figure 1). Hansen's definition permits us to delve into the normative applications of nudges without being caught up in the ethics of nudges (Hansen, 2016).

Figure 1

Libertarian Paternalism vs. Nudges (Hansen, 2016)



Note. Left is Sunstein's original definition. Right is Hansen's adapted definition.

Categorizing Nudges

Hansen and Jespersen (2013) propose a framework for nudging that makes use of dual process theory. The objective of doing so was to demonstrate that nudging can be categorized according to degrees of transparency and thus is not always a blatant manipulation of choice, as described by anti-nudgers. By eliciting dual process theory, the authors suggest that nudging always influences behaviour, but it only sometimes affects choices (Hansen & Jespersen, 2013). They categorize nudges into a framework of four groups. They did so to clarify the normative discussion of nudge implementation in the political sphere. This framework provides better information on how nudges function and the implications of their implementation. The central distinction between transparent and non-transparent nudges was used to identify manipulation. Manipulation is defined as the intentional decrease in the possibility of making an independent choice. Transparency is defined as the ease with which a nudged individual can discern the intention and methods behind the nudge at hand. In other words, a nudged citizen may recognize the nudge and freely choose to act in favour of or against the desired behaviour. Finally, in addition to transparency Hansen and Jespersen

(2013) distinguish System 1 to System 2 nudges as those that trigger automatic thinking versus those that trigger reflective thinking.

Transparent type 2 nudges are designed to Prompt a Reflective Choice. An example of that is the usage of an opt-out design, instead of an opt-in design for organ donations on one's driver's license registration. Nudges of this sort are easy to understand and reconstruct. They work by making the desired behavioural outcome salient. Thus, a citizen will be nudged predictably yet is still free to choose otherwise (Hansen & Jespersen, 2013).

Transparent type 1 nudges are designed to Influence Behavior, such as playing relaxing music on a plane or adding flashing lights to a sign, or painting illusions of speed bumps on streets. Here the presence and the purpose of the nudge are obvious. System 1 triggers are "*difficult, if not impossible to avoid because it activates instinctive or learned responses"* and the policymaker who executes these nudges has full responsibility for how they are influencing behaviour (Hansen & Jespersen, 2013).

Non-transparent type 1 nudges are a Manipulation of Behavior. Like Transparent type 1 nudges, these evoke instinctive, learned automatic responses. However, the subtlety in the non-transparent design makes it so that the individual on the receiving end may not recognize that they are being nudged. Consequently, citizens cannot avoid their effects. A well-documented example of this is the influence of plate size on servings and thus calorie intake (Hansen & Jespersen, 2013).

Non-transparent type 2 nudges trigger reflective thinking in such a way that the intentions and methods remain veiled. Hansen and Jespersen call this Manipulation of Choice. For example, the framing of risks when given a choice between two medical treatments. Framing means designing the presentation of information in a manner that is likely to sway a response in the direction preferred by the choice architect. Another example is anchoring people's

willingness to pay for an item, such as a chocolate bar, on a much higher number. Anchoring means setting a reference point, the anchor, to then influence following decisions in the direction favoured by the choice architect (Hansen & Jespersen, 2013).

Without knowledge of intention or method, nor even the ability to recognize the presence of a nudge, non-transparent type 2 nudges manipulate the citizen into complying with a choice, while creating the illusion of freedom and independence.

Public Opinion in Policy Design

Public opinion is an essential factor in policy design, and policymakers need to consider it to ensure the effectiveness, ethics, and legitimacy of their policies. In this regard, the following four articles explore the relationship between public opinion and policy.

Wlezien's article provides an overview of the various ways in which public opinion influences policymaking (Wlezien & Soroka, 2016). According to him, public opinion affects the political process, leading to policy changes. Similarly, Page and Shapiro's articles examine long-term trends in American public opinion and their effect on policy, suggesting that the public's views are a strong determinant of policy (Page & Shapiro, 1993). They further emphasize the need to prioritize transparency and ensure that the public is well-informed on policy design to increase public acceptance.

Public opinion can impact policy through the political process, with politicians responsive to changes in public opinion, as found in Page and Shapiro's previous article (Page & Shapiro, 1983). Additionally, public opinion can affect the media's coverage of policy issues, leading to further public discussion and debate.

Finally, Soroka and Wlezien's article provides a comprehensive analysis of the relationship between public opinion and policy, highlighting the various factors that impact it, such as the strength and stability of public opinion, the responsiveness of politicians to public opinion, and the political institutions in place. They emphasize the importance of policies aligning with the public's perception of fairness, justice, legitimacy, and democracy to maintain public trust in the policymaking process (Soroka et al., 2011).

In conclusion, policymakers must consider public opinion in designing and implementing policies to ensure their effectiveness, ethics, and legitimacy. The articles provide arguments in favour of measuring public opinion on policy design, even if it is less informed, to optimize representative democracy. Policymakers must prioritize transparency and ensure that the public is well-informed on policy design to increase public acceptance. Furthermore, policies should align with the public's perception of fairness, justice, legitimacy, and democracy to maintain public trust in the policymaking process.

Empirical Framework/Review

Attitudes towards nudges have been tested in relation to an assorted set of explanatory variables, such as transparency, targeted system, nationality, and political inclination. This section will go over a select few previous studies to gather existing conclusions on the effect of transparency on the public’s opinion towards nudges.

General Public Sentiment

Numerous studies have shown that public opinion is generally in favour of nudges, such as Jung & Mellers, 2016; Reisch & Sunstein, 2016; and Gold et al., 2020.

A study conducted study by Sunstein and Reisch (2016) compare their findings on Europeans’ sentiment on nudges, to a similar study conducted with Americans. The general conclusion of both studies is that these populations generally support nudges. Europeans were

largely in favour of nudges if it seemed that they were designed to benefit greater welfare. Americans seemed more bipartisan on certain topics. Hence, a salient difference is that while political affiliations did not influence in Europe, they did in the United States. Nonetheless, most results from the United States are on par with those from France, Germany, Italy, and the United Kingdom (4 of the 6 countries included in the European study) (Reisch & Sunstein, 2016; Sunstein et al., 2019).

It is important to note that nudges in these two studies were designed for welfare causes, under health, safety, and clean energy. Additionally, the participants did not get information on benefits nor costs.

Public Sentiment towards targeted system

It seems that people prefer System 2 interventions to System 1 interventions (Sunstein, 2016.). Sunstein's 2016 study, had participants choose between sets of two policies that were identical in intention and differed in methods regarding the targeted system. System 2 targets usually prevailed. He cleverly reiterated his findings in People Prefer System 2 (Sort of), referring to how this preference sometimes changes (Sunstein, 2016). Notably polarized topics, such as abortion, seemed to sway attitudes. People who are of the political inclination in favour of a polarized issue prefer System 1 interventions, whereas people who are against strongly prefer System 2 interventions.

This finding alludes to the notion that System 1 nudges give less room to personal agency than their reflective counterpart, giving rise to Sunstein's proposition there exists a tradeoff between the sacrifice of personal agency and an increase in effectiveness.

Another study by Davidai and Shafir (2020) iterated similar findings, in addition to testing people's opinions in a different way. They tested people's preferences in a joint evaluation of nudges against a separate evaluation. They had one group evaluate nudges in comparison to

each other, whilst the other group evaluated lone-standing nudges. Their results show that attitudes towards System 2 nudges are more positive when evaluated alone than when compared to others. Additionally, when evaluated separately, participants were about as equally in favour of both types of nudges. They discuss that the predominantly negative rhetoric around System 1 nudges may be more influenced by how nudges have been presented, than by people's actual preferences. It is delightful to observe that framing, which is a well-known nudge, will nudge attitudes towards nudges!

The purpose of recalling those two studies is to contemplate a methodology procedure that is not only accurate but an auxiliary to previous work.

Sunstein (2016) put together a speculation, that is best formulated by himself: "*System 1 prefers System 2 nudges*. That is, people may have an automatic response to say System 2 nudges are better because, offhand they may seem to be the most respectful of individual agency.

This idea matters because, when eliciting public opinion to improve public policy design, it would be best to ensure that the results are accurate. To test this idea, further research on preference for targeting System 1 versus System 2 must stipulate if their study's methodology is eliciting System 1 behaviour or System 2 thinking.

Finally, Davidai and Shafir (2020) also discuss the most relevant way to design further studies. While joint evaluations seemed to make clear the preference between two objects, the public can only like or dislike a policy, they do not get to compare options before implementation. In this situation, only policymakers get to compare different interventions at the same point in time. On this account, it seems that further research focused on applied normative hypotheses, such as one to improve guidelines, should take on separate evaluations.

Public Sentiment Towards Transparency

A major building block of the hypothesis and methodology of this paper is based on Gold et al.'s (2020) study on how people evaluate nudges. They presented participants with five scenarios for which there was both a transparent and non-transparent nudge in use. In every scenario, participants were informed of the intention and method as well as the definition of transparency. Participants then responded to four questions, one of which measured their evaluation of a nudge's acceptability: "*To what extent do you think it is acceptable to use the psychological method described in this context to change your behaviour? (Scale 1 = Unacceptable 9 = Acceptable)"*. The results demonstrated that transparent nudges were judged as more acceptable than non-transparent nudges (Gold et al., 2020)

Additionally, this study was aimed at participants' attitudes towards nudges concerning their own behaviour. This specificity in design is due to the authors' observation that previous studies had not distinguished whether people support nudges because they want to change their own behavior or other people's behavior (Gold et al., 2020). The outcome revealed that people support nudges but believe that they will be more effective on others. In social psychology, this is called the illusory superiority bias. Accordingly, the authors of the study reiterate that preceding papers have demonstrated this unilateral realism, which is seen as skepticism towards "others". (Gold et al., 2020).

This result suggests that further research should consider the bias and design a methodology that makes it clear to the participant if they are regarding nudging themselves or nudging others. Granted, the more pragmatic answers seem to come from judging others and thus may be more applicable to normative research.

METHODOLOGY

Overview & Hypotheses

Does the transparency of a nudge, implemented by public officials, impact the public's opinion on nudges? To answer this research question a survey of public opinion was constructed to measure preference amongst the four types of nudges, as conveyed by Hansen and Jespersen (2013).

In the present study, participants were separated into two groups, a control group and a treatment group, to provide their opinion on the implementation of certain nudge policies. Each policy was designed to satisfy one of the four types of nudges. The control group was somewhat like consultants for the policymakers. They had been given all the information on the intention and method behind the nudge. They could agree or disagree with the implementation of the nudge based on their complete understanding. On the other hand, the treatment group experienced the nudge without any explanation from the policymaker. Note that the data for this experimental study was compiled through a survey. Hence, the participant experienced the nudge theoretically. We can acknowledge that a survey may yield different results than an experiment conducted in a laboratory or the real world.

Ultimately, this study tested for a difference in agreement with nudge policies between those who have been intentionally made aware of a premise and those who were only provided with limited information.

The proposed methodology will test the following four hypotheses:

H1: Citizens who receive complete information on intention and methodology will be more in agreement with nudges overall, compared to those who do not.

H2: Citizens are least in agreement with the implementation of Non-Transparent System 2 nudges.

H3: Citizens are more in agreement with the implementation of transparent nudges than non-transparent nudges.

H4: Citizens are more in agreement with the implementation of System 2 nudges (reflective thinking and choice change) than System 1 nudges (automatic thinking and behaviour change).

Participants

Respondents were solicited through means of social media (LinkedIn and Facebook), survey exchange websites (Surveyswap.io and SurveyCircle.com) and through networking for broader distribution. These sourcing channels allowed for a large range of nationalities from North America, primarily Canadian, and Europe. It is vital to be aware that despite the efforts for a randomized sample, I suspect that the majority of the participants are within a third degree of myself. Additionally, the information provided in the survey required an intermediate English reading level. Thus, all participants are native to or educated in English.

In the span of three weeks, the survey saw 200 participants. From this sample, 15% were ignored due to incomplete input and for another 3%, their nationality did not correspond to a democracy, according to the Economist Intelligence Unit's index of democracy (EIU, 2023). Excluding participants from non-democratic countries from this survey on public policy design is done to ensure that the results of the survey accurately represent the views of individuals who live in democratic societies. The inclusion of participants from non-democratic countries may skew the results of the survey, as individuals living in non-democratic societies may have different perspectives and experiences related to policy design compared to those living in democratic societies. The Economist Intelligence Unit's index of democracy is a widely used index that is based on ratings for five categories of indicators

such as electoral process and pluralism; civil liberties; the functioning of government; political participation; and political culture.

Given the above exclusions, the final sample size included 164 participants, of which 52% completed the control survey.

Variables & Design

The primary dependent variable in this study is the public sentiment scores for each nudge category and nudges in general. Throughout the survey, sentiment is tallied as the sum of points, measured with a 5-point Likert scale. The independent variable is the presence or absence of supplemental information on the motivation and mechanisms behind the theme of the nudge. The control group experienced the presence of supplemental information, whereas the treatment group experienced an absence of supplemental information.

The participants began the survey with four randomly issued, demographic questions – about their age, gender, occupation, and nationality.

In the second section of the survey participants were provided with an accessible (user-friendly) definition of "Nudge". Next, the control group and the treatment group were introduced to the corresponding transparency obligations that their public policymakers must comply with when implementing a nudge. The control group received supplemental information on the motivation and mechanisms behind the theme of the nudge, making their vignettes more transparent than those of the control group. The control group and the treatment group were then presented with eight identical nudges, in random order. Every nudge corresponded to one of the four categories described by Hassen and Jensen (2013) as well as covering two potential public policy themes, meat consumption (diet) and individual savings (finance). These topics were picked because they include personal choice and socio-economic consequences. The intentions for each nudge were the same for each respective

theme. The policymakers in the meat consumption theme believe that the health of the climate and of individuals are negatively affected by meat consumption and would like to make it decrease. The policymakers in the individual savings consumption theme believe that the economic health of society and individuals is positively affected by increased income savings and would like to make it increase. Appendix 1 illustrates all eight nudges and methods presented to the participants. Those methods and nudges were selected based on studies done by Hansen & Jespersen (2013) and Caraban et al. (2019) where they categorize different nudges using the former's nudge matrix framework. Figures 2 and 3 illustrate the latter.

In the final section of the survey, all participants were presented with eight identical opinion questions – about their beliefs concerning the motivation behind the overarching public policy themes, about their beliefs in the bias that nudges are designed to deviate, and about their perception of public policymakers and nudging in general.

Figure 2

Nudges organized by category (Hansen & Jespersen, 2013)

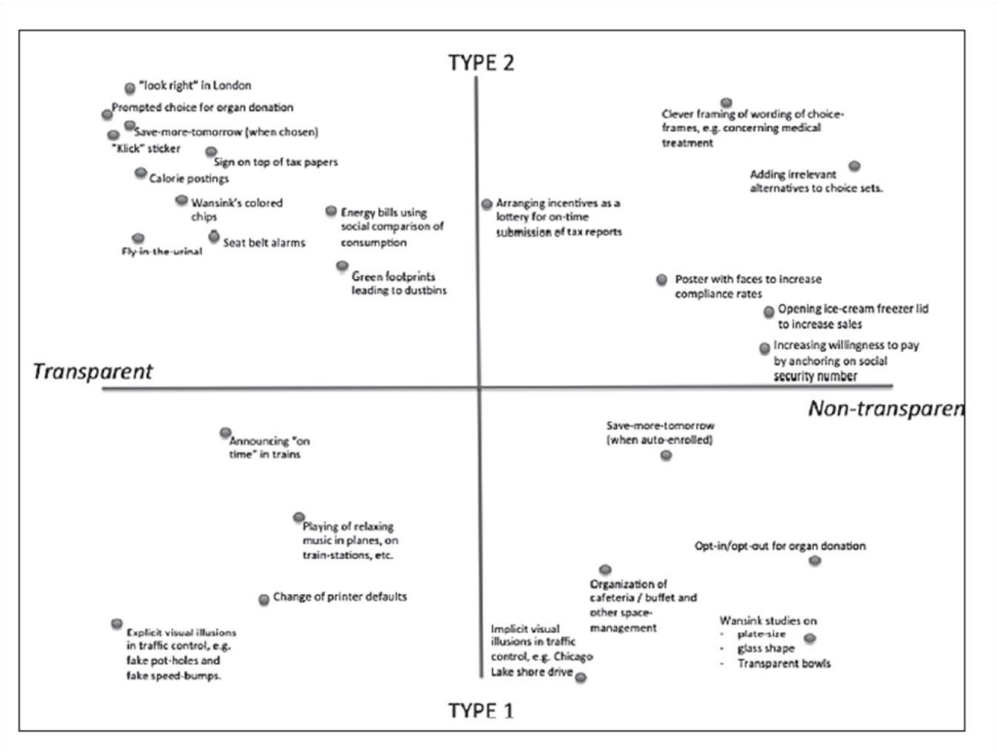
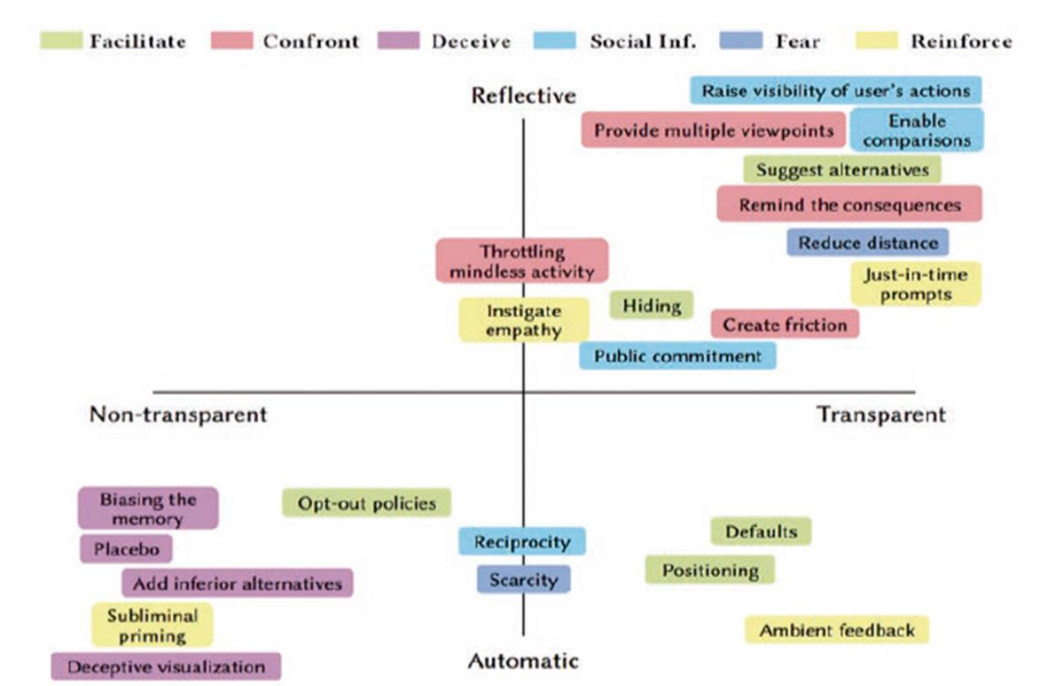


Figure 3

Nudges organized by category (Caraban et al., 2019)



Descriptive Statistics

The data from the survey is a sample population with a gender split of 58% male and 41% female. The average age of the sample is 32.

In terms of occupation, 40% of the sample are full or part-time workers, while 55% are students and the remaining 6% fall under the category of "other."

The sample also seems to be diverse in terms of geographical location, with 40% of the participants being from Canada, 12% from Germany, 11% from the Netherlands, 5% from the USA, 23% from other countries in Europe, and 9% from elsewhere.

Finally, 70% of the sample comes from full democracies and 30% come from flawed democracies. See Appendix 2 for more details on nationality, democracy type and age.

Analytic technique

Between-group analysis

Since the data generated from the survey is ordinal and rank-based, a nonparametric method to measure the between-group results is preferable. Also, the sample size is small and not all the assumptions of parametric tests can be met. Considering that the survey was distributed in two versions to two independent groups, the Mann-Whitney U test best applies.

RESULTS

As described previously, two randomized groups were asked to provide their opinion on the implementation of certain nudge policies. Participants recorded their opinion by indicating how much they would agree or disagree with the implementation. The control group received a description of the nudge and was made aware of the motivation and mechanisms behind the nudges. Recall that the objective of a nudge is to steer behaviour and choices. The treatment group received no information on the motivation and mechanisms behind the nudges. Ultimately, this experiment tests for a difference in agreement with nudge policies between those who have been intentionally made aware of a premise and those who were only provided with limited information.

The results are presented as follows:

1- Results by category will show the participant's sentiment towards the four categories of nudges defined by Hansen & Jespersen (see literature review). The fact that both transparency and system are integral to nudge mechanism design makes these results useful for practical application purposes.

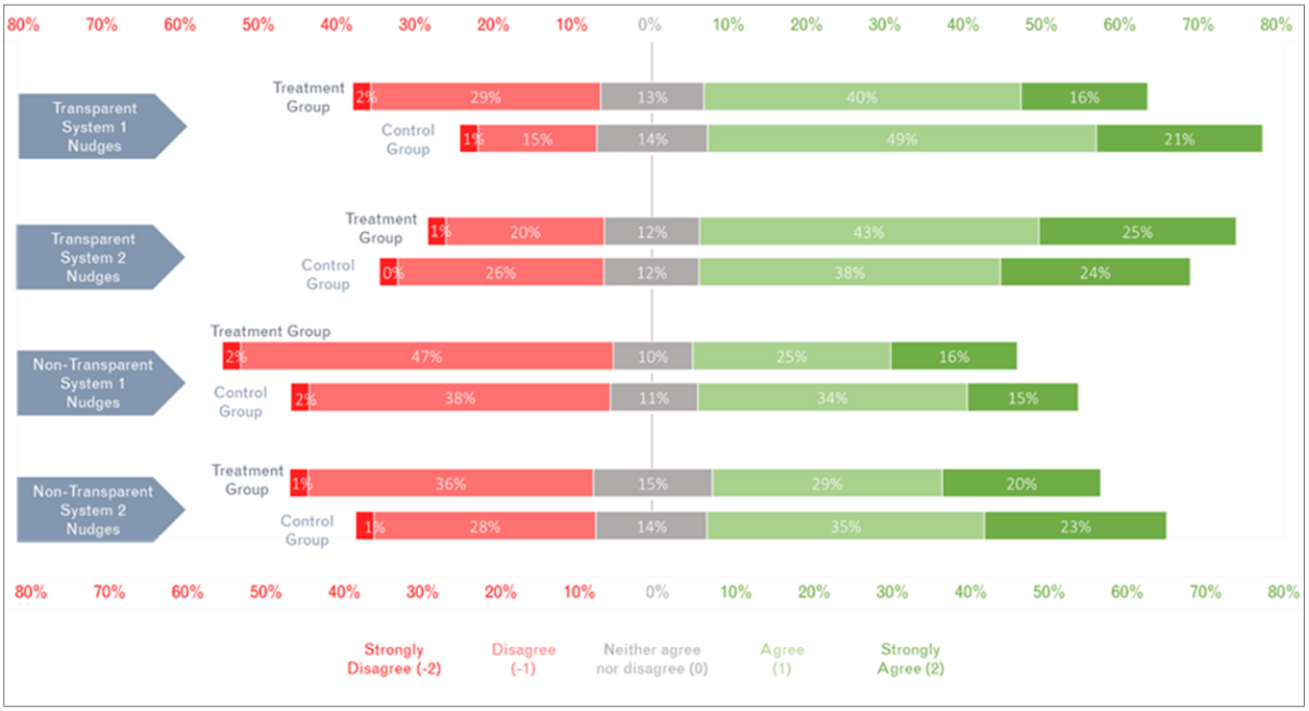
2- Results by type will show participants' overarching sentiments towards different types of nudges concerning their mechanism: whether they trigger System 1 or System 2 thinking and whether they are transparent or non-transparent. This analysis provides a theoretical understanding of which elements of a nudge design are more likely to be accepted.

3- The exploratory questions were designed to provide additional insights into participants' attitudes towards nudges, beyond the predefined categories and types. Further research should investigate further into the effects of demographics as they yielded no conclusive outcomes in the current study.

Results by Category

Figure 4

Degree of agreement with the implementation of nudges, count in % by category



Note. Results by category will show the participant's sentiment towards the four categories of nudges defined by Hansen & Jespersen

The following results are illustrated in Figure 4 and explained in detail by level of agreement below.

Strongly Disagree can be considered an extreme value due to how infrequently it was selected across all nudge categories and for both groups.

Neither agree nor disagree represents between 10% to 15% of responses for every category and changed very little from the treatment group to the control group within each category, remaining relatively consistent across all types of nudges.

Strongly Agree represents between 15% to 25% of responses for every category and varied little from control to treatment group within each category. The largest difference was 5% in the case of Transparent System 1 nudges, favoured by the control group. No group has a stronger tendency to pick *Strongly Agree*.

Disagree and *Agree* collectively represent approximately two-thirds of the responses for every category. Thus, driving the difference in results between the two groups.

- Transparent System 1 nudges saw *Disagree* decrease by 14% and *Agree* increase by 9% from the treatment group to the control group.
- Transparent System 2 nudges saw *Disagree* increase by 6% and *Agree* decrease by 5% from the treatment group to the control group.
- Non-Transparent System 1 nudges saw *Disagree* decrease by 9% and *Agree* increase by 9% from the treatment group to the control group.
- Non-Transparent System 2 nudges saw *Disagree* decrease by 8% and *Agree* increase by 6% from the treatment group to the control group.
- Transparent System 1 nudges, Non-Transparent System 1 nudges and Non-Transparent System 2 nudges, gained appreciation when intent and method were known.

- Transparent System 2 Nudges, each category decreased in appreciated when intent and methods were known.
- In the treatment group, Transparent System 2 nudges had the highest agreement rate (43%).
- In the control group, Transparent System 1 nudges had the highest agreement rate (49%).
- In both groups Non-Transparent System 1 nudges had the lowest agreement rate.
- In the treatment group, the most contrasting difference between *Agree* and *Disagree* is 22%. The weight falls on the *Agree* side for Transparent System 2 and on the *Disagree* side for Non-Transparent System 1.
- In the control group, the most contrasting difference between *Agree* and *Disagree* is 33%, falling on the *Agree* side, for Transparent System 1.

Results by Type

The following results are illustrated in Figure 5 and explained in detail by level of agreement below.

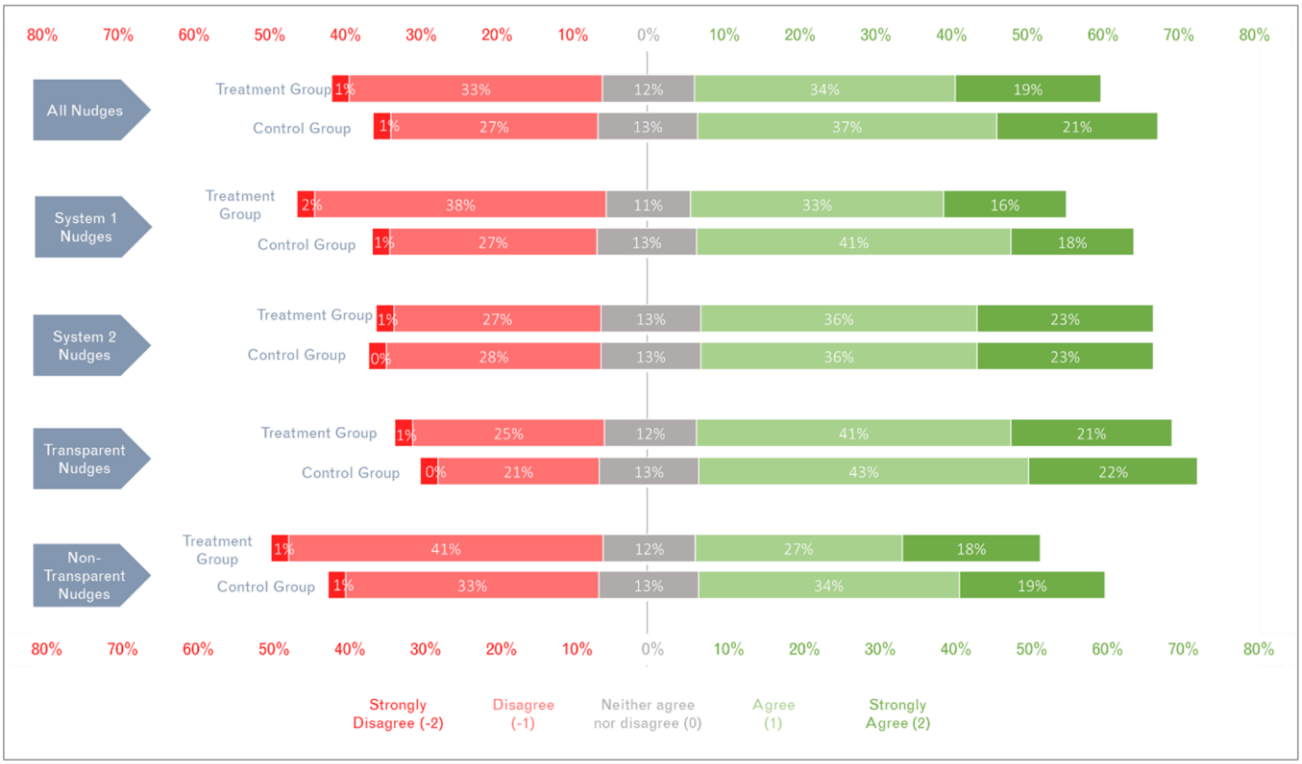
Strongly Disagree can be considered an extreme value due to how infrequently it was selected across all nudge categories and for both groups.

Neither agree nor disagree represents between 11% to 13% of responses for every category and changed very little from the treatment group to the control group within each category, remaining relatively consistent across all types of nudges.

Strongly Agree represents between 16% to 23% of responses for every category and varied no more than 2% from control to treatment group within each category.

Figure 5

Degree of agreement with the implementation of nudges, count in % by type



Note. Results by type show participants' overarching sentiments towards different types of nudges concerning their mechanism: whether they trigger System 1 or System 2 thinking and whether they are transparent or non-transparent.

Disagree and *Agree* collectively represent approximately two-thirds of the responses for every category. Thus, driving the difference in results between the two groups.

- System 1 nudges saw *Disagree* decrease by 11% and *Agree* increase by 8% from the treatment group to the control group.
- System 2 nudges saw *Disagree* increase by 1% and *Agree* remain the same from the treatment group to the control group.
- Transparent nudges saw *Disagree* decrease by 4% and *Agree* increase by 2% from the treatment group to the control group.
- Non-Transparent nudges saw a decrease of 8% and *Agree* increase of 7% from the treatment group to the control group.

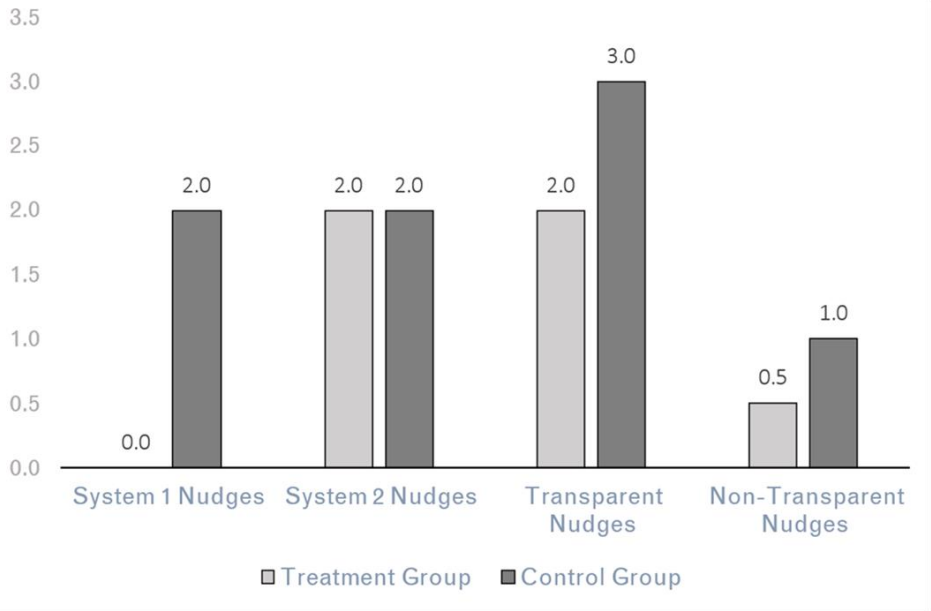
- System 1, transparent and non-transparent nudges gained appreciation when intent and method are known.
- System 2 Nudges stayed at a constant appreciation level.
- In both groups Transparent nudges had the highest agreement rate.
- In both groups Non-Transparent nudges had the lowest agreement rate.
- In the treatment group, the most contrasting difference between *Agree* and *Disagree* is 22%, falling on the *Agree* side, for Transparent System 1.
- In the control group, the most contrasting difference between *Agree* and *Disagree* is 17%, falling on the *Agree* side, for Transparent System 1.

The subsequent transformation of responses into a score paints a similar picture. Recall that the score is constructed by the addition of the scale numbers corresponding to each of the 8 nudge questions given to the participants. As a result, participants can score between -8 and 8 in each nudge category. Considering that the infrequency of *Strongly Disagree* (-2) causes a skewed distribution, the median, as seen in Figure 6 best illustrates the differences in agreement trends between the two groups. Despite the similar median scores across all four categories, two observations stand out. First, Non-Transparent nudges are the least agreed with and second, System 1 nudges have the largest difference between the two groups. Then, testing for the significance of these differences, or lack thereof in the case of System 2 nudges, we employ a Mann-Whitney U test.

Notwithstanding the descriptive results, only the cumulative scores of System 1 nudges are significantly different between the control group to the treatment group, at a 5% level ($p=0.011$).

Figure 6

Degree of agreement with the implementation of nudges, median of cumulative score by type



Note. Participants can score between -8 and 8 in each nudge category. The control group had access to full information on the presented nudges, whereas the treatment group had limited information.

Exploratory Questions

Questions such as "I believe that the health of the climate and of individuals are negatively affected by meat consumption" and "All policymakers that implement nudges will always design a nudge that is in the best interest of the People" were asked to better understand the participants and their agreement rates. However, the average answer for all questions hovered around 0 *(Neither agree nor disagree),* and there was little variation in results between the control and treatment groups.

DISCUSSION

The study investigates whether the transparency of nudges implemented by public officials impacts public opinion on nudges. A survey was used to measure preferences for the four types of nudges. Participants were separated into a control and treatment group. The control group was exposed to the intent and method behind every nudge they were asked to assess,

the treatment group was not. This was done to test four hypotheses. H1: Citizens who receive complete information on intention and methodology will be more in agreement with nudges overall, compared to those who do not. H2: Citizens are least in agreement with the implementation of Non-Transparent System 2 nudges. H3: Citizens are more in agreement with the implementation of transparent nudges than non-transparent nudges. H4: Citizens are more in agreement with the implementation of System 2 nudges than System 1 nudges.

Findings & Implications

The results generally align with the hypotheses presented in the methodology. Participants who received complete information were indeed more likely to agree with nudges overall, supporting H1. The finding that Non-Transparent System 2 nudges were the least likely to be agreed with thus rejecting H2, while the higher agreement rates for transparent nudges and System 2 nudges support H3 and H4, respectively.

The results show that the transparency and the System being triggered in the nudge design have an impact on participants' agreement or disagreement.

The results suggest that citizens may generally be in favour of nudges, however, the aggregated rates of agreement were not overwhelmingly higher than the rate of disagreement. Thus, it becomes important to pose more precise questions to better evaluate people’s views on nudge designs.

Hansen and Jespersen (2013) categorized nudges into four categories and provided guidelines on how to avoid the likeness and possibility of manipulation. Nevertheless, they tagged Non-Transparent System 1 and System 2 nudges as manipulations of behaviour and choice, respectively. Non-transparent System 1 nudges had the lowest agreement rate in both groups, by a noticeable amount, compared to the other categories. This finding rejects Hypothesis 2; citizens are least in agreement with the implementation of Non-Transparent System 2 nudges.

This raises the question of how accurate an individual's perception of their autonomy over their System 2 truly is. It is a question that Sunstein (2017) has brought up and is worth looking to investigate a better understanding of Kahneman's two systems.

A notable result however is the slight increase in agreement rate by those who were exposed to intent and method. This allows the conclusion that increasing transparency, by providing explanations, tends to increase approval, even in the least approved categories. However, the rate of disagreement implies that in practice those nudge categories may be avoided in policy design.

The results indicate that providing information on the intent and method behind nudges has a positive effect on the acceptance of nudges, particularly for Transparent System 1 nudges. *Disagree* decreased and *Agree* increased from the treatment group to the control group for all categories except for System 2 nudges, which remained relatively consistent. The highest agreement rate was found for Transparent nudges in both groups, while Non-Transparent nudges had the lowest agreement rate. The most contrasting difference between *Agree* and *Disagree* was found for Transparent System 1 nudges in both groups, indicating that this type of nudge was the most effective in generating agreement.

Next, the implications of participants' higher rate of agreement with Transparent and System 2 nudges compared to their respective counterparts implies that citizens value autonomy and trust in decision-making, actions and choices. Where System 2 nudges trigger reflections, System 1 nudges trigger reactive behaviours and may be perceived as limiting freedom of choice. The same can be said for transparent and non-transparent nudges. However, it is interesting to note that the rate of agreement for System 1 and non-transparent nudges were higher among the control group – implying that receiving an explanation may pardon the limitation of free choice, knowing intent and method ultimately increases overall

transparency and lets the citizens “in” on the policy. They may feel that they are better equipped to make their own choices when they know how the nudge architecture functions.

Additionally, *Strongly Disagree* is a rare response across all nudge categories and types, for both groups. This lack of strong negative opinions could be due to survey biases, the chosen scenarios, and the hypothetical approach. *Neither agree nor disagree* is a consistent response across all categories and types of nudges, while *Strongly Agree* was a relatively common and consistent response.

In summary, the results show that providing information on the intent and method behind nudges can positively affect their acceptance, particularly for certain types of nudges, and that Transparent nudges are generally more effective in generating agreement than Non-Transparent nudges.

Limitations

A limitation of the study is that it was conducted in a hypothetical scenario through a survey, and future research may benefit from examining the impact of transparency nudges in a laboratory or the real world. As is the case with most surveys, and especially in those of small scale, there are a variety of limitations. Just the nature of a survey eliminates the option of discussing causality from the analysis. Surveys tend to bear self-report flaws which impact the reliability and the validity of the results. Increasing the sample size and with it the randomization to reduce sample bias is how one could get closer to the population's real sentiments. There is also a possibility of central tendency bias and social desirability bias, caused by the wording of the options and/or complexities of the information provided in each question. Additionally, the use of a survey instead of an experiment means that the participants did not experience the nudges in a real-world setting, which may limit the generalizability of the findings.

Additionally, responses could be biased due to inattentiveness, misinterpretation and indolence. As a consequence of 30 incomplete responses, I believe that this study is also subject to survey fatigue bias. There may also be potential to improve instructions and to better communicate the length of the survey. The study did not examine potential confounding variables that may have impacted the results, such as the participants' demographics or experience with the issues related to the nudges. Increasing the sample size to a larger variety of countries would lead to changes in the sentiment data retrieved or focusing on one nation at a time could also lead to more accurate results. Consequently, the external validity of these results may be limited.

Contribution

The purpose of this study is to provide policymakers with a framework that they can use to implement nudges effectively. The results demonstrate that prioritizing transparency can increase public acceptance. Public acceptance plays a crucial role in policy design in terms of both ethics and effectiveness. Firstly, effective implementation and public acceptance are positively related. The absence of public support may result in resistance, which can undermine effectiveness. Secondly, policies that align with the public's perception of fairness and justice are more likely to be accepted. Ethical misalignments may lead to civil unrest, whereas perceptions of fairness and justness can help build trust in the policymaking body. Thirdly, policies that align with the public's perception of legitimacy and democracy are also more likely to be accepted. This public validation reinforces the definition of democracy, while illegitimate and undemocratic policies destroy public trust. Policymakers must consider the views and values of the public in designing and implementing policies to ensure their legitimacy, effectiveness, and ethical soundness.

The findings of this study support previous research that suggests transparency is a key factor in shaping public perception of nudges. This study expands upon existing literature by specifically examining the impact of transparency in the context of nudges implemented by public officials. The results demonstrate the importance of transparency and explanation in gaining public support for policy decisions and highlight the potential downside of implementing "manipulative" nudges.

Further Exploration

Continued analysis into potential demographic differences in agreement rates for nudges would be beneficial for policymakers seeing to implement nudges. Additionally, future research could use laboratory or real-world experimental methods to examine the public's opinion of nudges when in application. A longitudinal study could also assess the potential unintended consequences of nudges. Finally, it may be useful to investigate the long-term effects of nudges on public opinion and trust in government, particularly considering the increasing use of AI and machine learning in policymaking.

Conclusion

The study investigated the impact of transparency of nudges on public opinion towards nudges. The findings suggest that citizens are generally in favour of nudges, with the highest agreement rate found for Transparent System 1 nudges. The results show that providing information on the intent and method behind nudges has a positive effect on the acceptance of nudges, particularly for Transparent System 1 nudges. Additionally, the results suggest that citizens value autonomy and trust in decision-making, actions, and choices. The study has limitations, including the use of a survey to conduct an experiment and the lack of

examination of potential confounding variables. Notwithstanding the limitations of this study, the results of this study have important implications for policymakers interested in implementing nudges in public policy.

BIBLIOGRAPHY

Caraban, A., Karapanos, E., Gonçalves, D., & Campos, P. (2019, May 2). 23 Ways to Nudge: A review of technology-mediated nudging in human-computer interaction. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3290605.3300733

Davidai, S., & Shafir, E. (2020). Are 'nudges' getting a fair shot? Joint versus separate evaluation. *Behavioural Public Policy*, *4*(3), 273–291. https://doi.org/10.1017/bpp.2018.9

EIU. (2023). *Democracy Index 2022: Frontline Democracy and the Battle For Ukraine*. www.eiu.com.

Gold, N., Lin, Y., Ashcroft, R., & Osman, M. (2020). 'Better off, as judged by themselves': do people support nudges as a method to change their own behavior? *Behavioural Public Policy*, 1–30. https://doi.org/10.1017/bpp.2020.6

Hansen, P. G. (2016). The definition of nudge and libertarian paternalism: Does the hand fit the glove? *European Journal of Risk Regulation*, *7*(1), 155–174. https://doi.org/10.1017/S1867299X00005468

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. https://doi.org/10.1017/s1867299x00002762

Jung, J. Y., & Mellers, B. A. (2016). American attitudes toward nudges. In *Judgment and Decision Making* (Vol. 11, Issue 1).

Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.

Kahneman, D. (2003). Maps of Bounded Rationality. *The American Economic Review*, *93*(5), 1449–1475.

Page, B., & Shapiro, R. (1983). Effects of Public Opinion on Policy. *Source: The American Political Science Review*, *77*(1), 175–190. https://www.jstor.org/stable/1956018

Page, B., & Shapiro, R. (1993). The Rational Public: Fifty Years of Trends in Americans' Policy Preferences. *Oxford University Press*, *57*(4), 612–614. https://www.jstor.org/stable/2749489

Reisch, L. A., & Sunstein, C. R. (2016). Do Europeans like nudges? In *Judgment and Decision Making* (Vol. 11, Issue 4). https://ssrn.com/abstract=2739118

Smets, K. (2018). *There Is More to Behavioral Economics Than Biases and Fallacies*. Behavioral Scientist. https://behavioralscientist.org/there-is-more-to-behavioral-science-than-biases-and-fallacies/

Soroka, S. N., Wlezien, C., & Barabas, J. (2011). Degrees of Democracy: Politics, Public Opinion, and Policy. *Public Opinion Quarterly*, *75*(1), 192–200. https://doi.org/10.1093/poq/nfq049

Sunstein, C. R. (2016). People Prefer System 2 Nudges (Kind of). *Duke Law Journal*, *66*(1), 121–168.

Sunstein, C. R. (2017). Misconceptions About Nudges. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3033101

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

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness.* Penguin Books.

Wlezien, C., & Soroka, S. N. (2016). Public Opinion and Public Policy. In *Oxford Research Encyclopedia of Politics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190228637.013.74

APPENDIX 1

Nudges and information presented in the survey

DESIRED EFFECT	increase individual's savings	decrease meat consumption
MOTIVATION	Explain why a policymaker wants people to have more savings - Decrease default risks - Old-age poverty	Explain why a policymaker would want to decrease meat consumption - Health risks - Climate change
BEHAVIORALLY ORIENTED NUDGES (System 1)		
Non-Transparent	Default Options : Being enrolled in a savings plan is the default. Examples of methods: - One must "opt-out" in order to decrease savings or not save at all.	Size alterations : Smaller portions & packages for meat options. Examples of methods: - Limiting packaged meat to a certain weight/quantity
Transparent	Convenience enhancements : The option to increase savings is easier to find than the option to decrease savings. Examples of methods: - Larger savings option is on top of employee dashboard - Frequent informative emails about the benefits of saving more	Convenience obstructions : Meat options are less convenient to select, compared to meat-free options. Examples of methods: - no presliced, preportioned, or preserved food - meat-free food placed earlier in a cafeteria line - decrease amount available in shelves, specific cuts and quantities are accessible through the store's butcher only.
COGNITIVELY ORIETNTED NUDGES (System 2)		
Non-Transparent	Framing: Using the influence of word selection to alter people's choices in favor if larger saving plans. Examples of methods: - using positive and rewarding words for large saving plans - using cautionary words for small saving plans	Visibility alterations : Meat options are less visible. Examples of methods: - not eye-level shelf positions - less "see-though" packaging - middle of the menu
Transparent	Social Comparisons : Receiving information about the saving habits of others (anonymous) Examples of methods : - You are made aware of how many people choose which plan - how much you save compared to your peers, national average etc. - examples of what other people have achieved in terms of savings	Evaluative nutritional labeling : Standardized labeling differentiating meat-free vs meat options. Examples of methods : - Negative stickers next to meat options - Environmental impact scale

APPENDIX 2

Descriptive Statistics: Nationality & Age

CONTROL GROUP			TREATMENT GROUP			ALL PARTICIPANTS		
ıtry	Count of Nationality	Average of Age	Country	Count of Nationality	Average of Age	Country	Count of Nationality	Average of Age
	24	29	Flawed	22	27	Flawed	46	28
Belgium	4	42	Belgium	4	42	Argentina	1	27
Brazil	1	31	Brazil	1	31	Belgium	4	42
Croatia	2	22	Croatia	2	22	Brazil	1	31
Czech Republic	1	24	Czech Republic	1	24	Colombia	3	30
France	1	27	France	1	27	Croatia	2	22
Greece	2	21	Greece	2	21	Czech Republic	1	24
Italy	2	22	Italy	2	22	France	1	27
Poland	1	18	Poland	1	18	Greece	5	23
Portugal	3	24	Portugal	3	24	India	1	20
Singapore	1	21	Singapore	1	21	Indonesia	1	23
South Africa	1	28	South Africa	1	28	Italy	6	23
United States	5	38	United States	5	38	Malta	2	23
	54	32	Full	64	35	Philippines	1	21
	28	38	Canada	37	41	Poland	2	19
	10	23	Germany	9	31	Portugal	4	24
	2	26	Ireland	2	23	Singapore	1	21
ıurg	2	23	Netherlands	9	23	Slovenia	2	24
nds	8	24	Norway	2	26	South Africa	1	28
land	1	36	Switzerland	1	35	United States	7	42
ınd	1	19	Taiwan	1	26	Full	118	34
	2	51	UK	3	26	Canada	65	40
	78	31	Total	86	33	Germany	19	26
						Iceland	2	26
						Ireland	2	23
						Luxembourg	2	23
						Netherlands	17	23
						New Zealand	1	36
						Norway	2	26
						Switzerland	2	27
						Taiwan	1	26
						UK	5	36
						Grand Total	164	32