



INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Nudging the nudges

A probe into transparency effects

Tiago Alexandre Paçô Galiza Monteiro

Master's in Business Economics and Competition

Supervisors:

Nelson Campos Ramalho, Associate Professor, ISCTE - University Institute of Lisbon

Teresa Féria Almeida, Invited Assistant Professor, ISCTE- University Institute of Lisbon

Department of Human Resources and Organizational Behavior | ISCTE Business School

September, 2023

iscte

BUSINESS
SCHOOL

Nudging the nudges

A probe into transparency effects

Tiago Alexandre Paçô Galiza Monteiro

Master's in Business Economics and Competition

Supervisors:

Nelson Campos Ramalho, Associate Professor, ISCTE - University Institute of Lisbon

Teresa Féria Almeida, Invited Assistant Professor, ISCTE- University Institute of Lisbon

Department of Human Resources and Organizational Behavior | ISCTE Business School

September, 2023

Acknowledgements

The completion of this study marks one of the most important phases of my life, both academically and personally. During my master's degree, I was lucky enough to meet people who helped me enrich my knowledge and grow as an individual, for which I am deeply grateful.

First and foremost, I want to express my heartfelt gratitude to my supervisors, Professor Nelson Ramalho and Professor Teresa Almeida, for their unwavering accessibility, dedication and critical guidance.

To my parents, a big thank you for the unconditional support they have given me up to this point, which has proven to be fundamental.

To my family, friends and girlfriend, thank you for your encouragement, valuable advice and good humor.

Abstract

Traditionally, economic models relied on rationality, assuming that people were utility maximizers with perfect information capabilities. These assumptions were abandoned as several psychologists and economists have demonstrated that individuals are subject to biases and cognitive failures. Within this realm, nudge theory, which states that our behavior can be effectively influenced through interventions, has shed light on how our judgment is systematically biased and altered by the decision making architecture.

Among the ongoing discussion on the effectiveness and empirical application of nudges, a central question remains unanswered: are nudges effective only as far as they are not realized consciously by the target? This question naturally precedes a second question: Aren't overt nudges more ethical than covert ones? To address this research gap, this study aims to determine how ascribed benevolence affects nudge effectiveness through transparency.

Based on a sample of two groups (a pre-test on 86 participants and a test group of 141 participants), a 2x2 experimental study was conducted. Participants were randomly assigned to four distinct scenarios, characterized by variations in benevolence and transparency factors. These scenarios shared a common stimulus which had an anchoring nudge effect in a task involving a donation decision.

The results indicate that disclosing nudges did not have a detrimental effect on donation behavior and that organizations perceived as more benevolent tend to receive greater donations. Furthermore, results emphasize the importance of considering the ascribed benevolence of the source.

Keywords: Behavioral Economics, Nudge Theory, Anchoring, Benevolence, Transparency

JEL Codes: D91, M12

Resumo

Tradicionalmente, os modelos económicos baseavam-se na racionalidade, assumindo que os indivíduos eram maximizadoras de utilidade com informação perfeita. Estes pressupostos já não refletem a realidade atual, visto que psicólogos e economistas comprovaram que os indivíduos têm enviesamentos e falhas cognitivas. Neste contexto, a teoria *nudge*, que afirma que o comportamento pode ser influenciado, destacou como a arquitetura da tomada de decisão distorce a nossa escolha.

Na discussão sobre a eficácia e a aplicação empírica de *nudges*, uma questão permanece sem resposta: *nudges* são apenas eficazes quando não são percebidas conscientemente pelo indivíduo? Esta naturalmente antecede outra: as *nudges* explícitas são mais éticas que as ocultas? Para aferir esta falha empírica, este estudo tem como objetivo medir como a benevolência afeta o efeito transparência na eficácia de uma *nudge*.

Com base numa amostra de dois grupos (um de pre-teste com 86 participantes e outro de teste com 141 participantes), um estudo experimental 2x2 foi conduzido. Este foi feito através da alocação aleatória dos participantes a quatro cenários caracterizados por variações de benevolência e transparência. Os cenários partilhavam um estímulo comum que gerava o efeito *nudge* de ancoragem na decisão de fazer uma doação.

Os resultados indicam que a divulgação da *nudge* não influencia a doação de um indivíduo e que as organizações consideradas mais benevolentes tendem a receber mais doações. Não obstante, os resultados refletem a importância da benevolência atribuída à organização que espoleta a *nudge*.

Palavras-chave: Economia Comportamental, Teoria de Nudge, Ancoragem, Benevolência, Transparência

Códigos JEL: D91, M12

Table of content

<i>Acknowledgements</i>	<i>i</i>
<i>Abstract</i>	<i>iii</i>
<i>Resumo</i>	<i>v</i>
<i>Table of contents</i>	<i>vii</i>
<i>Index of figures</i>	<i>ix</i>
<i>Index of tables</i>	<i>x</i>
<i>Chapter I - Introduction</i>	<i>1</i>
1.1. Research problem	1
1.2. Research goals.....	1
1.3. Study's relevance	1
1.4. Structure.....	2
<i>Chapter II - Literature Review</i>	<i>3</i>
2.1. Psychology and Economics.....	3
2.2. Bounded rationality.....	3
2.3. Nudge theory	4
2.4. Two thinking systems	6
2.5. Nudgeability and transparency	7
2.6. Nudgeability and benevolence.....	8
2.7. Individual's donation behavior.....	9
2.8. Hypotheses.....	10
<i>Chapter III - Method</i>	<i>11</i>
3.1. Design.....	11
3.2. Stimuli	12
3.3. Procedure.....	14
3.3. Sample.....	15
3.4. Data analysis strategy.....	16
<i>Chapter IV - Results</i>	<i>17</i>
4.1. Manipulation check.....	17
4.2. Descriptive and bivariate statistics.....	18
4.3. Hypotheses testing.....	19

<i>Chapter V - Discussion and conclusion</i>	24
5.1. Limitations.....	26
5.2. Future research	27
<i>References</i>	28
<i>Appendix A</i>	33

Index of figures

Figure 1 - Stimuli	13
Figure 2 - QR code to survey.....	15
Figure 3 - Means plot.....	17

Index of tables

Table 1 - Experimental conditions	12
Table 2 - Descriptive and bivariate statistics	18
Table 3 - Paired samples test.....	19
Table 4 - Levene's test of equality of error variances.....	20
Table 5 - Test of between-subjects effects for first donation.....	21
Table 6 - Test of between-subjects effects for final donation.....	22
Table 7 - Descriptive statistics for Benevolence * Disclosure.....	23

Chapter I - Introduction

1.1. Research problem

Nudging grew as a topic of interest in behavioural economics ever since Herbert Simon (1956, 1982) made the pioneering research on bounded rationality observable in decision making, which earned him a Nobel prize in Economics. Its fruitful work left a legacy that gave rise to decision making theory concepts such as heuristics (Tversky & Kahneman, 1974; 1989) that also deserved another Nobel prize in Economics (Daniel Kahneman in 2002) and ultimately nudge theory (Thaler & Sunstein, 2008) and the consecration of behavioural economics (Mullainathan & Thaler, 2000; Pesendorfer, 2006).

Among the ongoing discussion on the ethical use of nudges to influence people's decision (e.g. Bovens, 2009; Selinger & Whyte, 2011) a central question on the covert use of nudges takes shape: are nudges effective only as far as they are not realized consciously by the target? This question naturally precedes a second question: Aren't overt nudges more ethical than covert ones?

The research on this topic (over vs. covert nudging) is yet in its beginning having been recently equated with the new construct of "nudgeability" (de Ridder et al., 2022) since most of the research on nudge theory has been yet exploring its use, effectiveness and empirical application in many domains (e.g. Arno & Thomas, 2016; Carlsson et al., 2021; Hansen & Jespersen, 2013; Mertens et al., 2022). On the one hand, the fundamental idea that nudge operates at the subconscious level favors the thesis that nudges efficacy depends on being covert, i.e. on the target not been aware of its existence (Bovens, 2009). On the other hand, nudges have been found to be effective even when disclosure occurs (e.g. Bruns et al., 2015; Loewenstein et al., 2015) but such findings leave some open questions, concerning the role of inferred intentions that lead the source to enact the nudge, namely its benevolence (Bang et al., 2018).

1.2. Research goals

Considering the mentioned research line on nudgeability, this study is designed to ascertain to which extent ascribed benevolence modulates the effect transparency has upon nudge effectiveness.

1.3. Study's relevance

Although nudging has attracted the attention of many researchers from all scientific domains linked to social and economic sciences, there is still room for theoretical development as regards to nudgeability (de Ridder et al., 2022). Likewise, one cannot overstate the importance of nudging for all situations involving executives, managers, policy makers, or any sort of decision maker, that are

intending to maximize their persuasive effectiveness. Therefore, a study that focuses on nudging is intrinsically of high applied value for organizations and society.

In the specific case of this design, both transparency and benevolence have ethical relevance, which is a dimension with ever increasing importance for corporate governance. Therefore, this topic and research goal may not only contribute by uncovering interactions between transparency and benevolence (which are still understudied) as such interactions may help decision makers to fine tune their approach to nudging, thus leveraging its effectiveness but also its ethical grounds.

All in all, the effective use of nudges can serve as a less intrusive approach to promoting the greater good compared to more harsh interferences. However, the research on nudges predominantly focuses on psychological phenomena rather than quantifying the effects of nudges (Goswami & Urminsky, 2016). That being said, this study aims to measure the extent to which nudges can effectively influence or persuade an individual's willpower.

1.4. Structure

The structure of this dissertation comprises five chapters (Introduction, Literature review, Method, Results, Discussion and Conclusion) preceded by the Abstract, Acknowledgments, Table of contents, List of figures and tables, and finishing with References and appendices.

The first chapter provides an overview of the research problem, goals, relevance and structure that serve as the foundation for the study. The second chapter is about the literature review that sustains this thesis, namely key concepts and theories within the field of behavioral economics and nudge theory, as well as the interplay and articulation of transparency and benevolence in relation with nudgeability. The third chapter delves into the methodology employed in the study and discusses the empirical research. Subsequently, chapter four presents the data collection procedures and a comprehensive analysis of the results. Finally, chapter five engages in a discussion of the findings, offering a thoughtful analysis into the limitations and potential areas of future research, and concludes the dissertation by summarizing the main points achieved throughout the study.

Chapter II - Literature Review

2.1. Psychology and Economics

Matthew Rabin (1998, p.11) made the following statement in his article in the Journal of Economic Literature: "Because psychology systematically explores human judgement, behaviour, and well-being, it can teach us important facts about how humans differ from the way they are traditionally described by economics".

As economics can be defined by the field that studies how resources are allocated by individuals and enterprises, the psychology of individual behavior should underlie and be connected to economics, much as neuroscience is connected to cognitive psychology, physics is connected to chemistry or archaeology is connected to anthropology (Camerer, 1999). That being said, ever since Rabin's contribution, behavioural economics and the interrelation between psychology and economics has been getting a lot more attention and recognition as a way of fully understanding not only human behaviour but also market outcomes.

Traditionally, economic models used to rely on rationality, assuming that people are selfish actors and utility maximisers with perfect information capabilities. Their preferences were assumed to be independent of the framing of the decision. These assumptions are long abandoned as several psychologists and economists have proven that individuals are subject to biases and cognitive failures, and because of that they fail to make the most correct choices (Lehner & Mont, 2016).

The most primordial and fundamental example, dates to 1979 when Daniel Kahneman and Amos Tversky published a paper where they demonstrate that individuals systematically contradict the premise of expected utility theory as they suffer from cognitive failures and biases.

Significant efforts have been done since then to tackle the challenge of attaching economic theory with psychological foundations. Portrayed as the reunification of these two fields, behavioral economics aims to comprise psychology insights to enrich economic conclusions, while maintaining the data use and mathematical structure that distinguish economics from other social sciences (Camerer, 1999).

2.2. Bounded rationality

Behavioral economics, despite its recent popularity, draws on ideas that harken back to the early roots of classic and neoclassical economics. In its nascent stages, economics and psychology were intertwined, with economists often taking on the role of psychologists (Camerer & Loewenstein,

2004). In this regard, Adam Smith, often referred as the father of Economics, emphasized psychological principles influencing human behavior in his book "The Wealth of Nations", foreshadowing in a sense many concepts of modern behavioral economics like loss aversion.

Subsequent ideas and theoretical frameworks have put bounded rationality – viewed as a significant yet concise terminological component within the field of economics – as one of the main challengers to the neoclassical era. This is because, as Langlois (1990) explained, economic agents are not automated machines, capable of instantaneously and effortlessly solving the intricate optimization problems that neoclassical theorists often attributed to them.

In economic theory, the shift from working with utility functions to addressing real-world economic questions often requires additional assumptions beyond optimization about the characteristics of an individual. These additional assumptions commonly pertain to the limitations on knowledge and computational abilities, reflecting the concept of bounded rationality (Egidi et al., 1992). In this context, Simon (1982) and his contributions for behavioural economics, helped the decision-making process to take shape and to place it as fundamental piece in the interactions and relations that we all have.

Bounded rationality is in that sense a term used to name rational choice while bearing in mind the cognitive limitations of the decision maker (Simon, 1990). In other words, this term asserts that decision makers are goal-oriented and adaptive, but because of human cognitive and emotional architecture, they do not choose the most optimal option available (Jones, 1999). In this line of thought, Gigerenzer and Selten (2002) emphasized that bounded rationality models not only describe how a judgment or behavior is reached but also describe the environment structure in which heuristics will succeed or not. The decisions and actions that we all take are therefore dependent on choosing between several alternatives and considering the risk and uncertainty that comes with it.

2.3. Nudge theory

Nudges, a term emphasised by Thaler and Sunstein (2008), can be described as a technique used to alter or steer people's behaviour in a predictable way, without limiting any options or significantly changing their incentives.

In a sense, Hansen and Jespersen (2013) noted that the use of nudges has revealed how our decision making and behaviour are systematically biased and altered by the interplay of psychological thinking. And because of that, lead us to fail in acting on our well-informed intentions or achieve our preferred ends. However, Kosters and Heijden (2015) outlined the act of nudging as having two beneficial sides. First, the implementation of nudges can be viewed to assist people in making the

most optimal decision in their best individual interest. Secondly, these interventions can be interpreted to steer people's choices to achieve social and collective wellbeing.

Ever since the term was brought up by Thaler and Sunstein (2008), nudging became immensely popular as a policy instrument. This rise in popularity, has been most prominently perceived among policymakers as their eagerness to alter people's behavior increased. Furthermore, countless cases of pension schemes, health incentives and sustainable actions structured in the past decade, have supported the legitimacy and success of nudge interventions among individual's willpower (Benartzi et al., 2017)

One well-known example that Thaler and Sunstein (2008) gave concerns retirement plans. In this case, an organization may require employees to opt out of, rather than opt into a retirement plan, taking advantage of the employee's tendency to select default options (i.e., decisional inertia), and thus promoting the act of retirement savings. This concept has been named libertarian paternalism and departs from the essential premise that individuals are free to choose. To put it differently, although the retirement plan example sets a default option, individuals remain free to choose to save any amount or not (Felsen et al., 2013). In this regard, this phenomenon has the purpose to overcome people's cognitive failures by manipulating their decisions towards behaviors they would have made if they were totally rational (Rebonato, 2012).

In the same line of thought, Rebonato (2012) reasoned about the legitimacy of the libertarian aspect of libertarian paternalists even when compared to other more paternalistic versions. His reasoning sided with the fact that the intentions and goals of individuals that adopt libertarian paternalism are not made aware, just like their assumptions about choice architectures that may always preserve the best interests of the individual that is being nudged, i.e. defenders of libertarian paternalism assume too much about the freedom to choose and intrinsic good of a nudging action. In this respect, nudges can be used to subconsciously influence our decision making and to improve our well-being, or not. For this to happen, Thaler and Sunstein (2008) highlighted the concept of "choice architecture". This latter designation represents the context in which individuals make decisions. It can be described as the surrounding environment and the interferences that impact the decision-making processes of an individual (Cooper, 2017).

One other well-known example in this setting, can be the school cafeteria illustration. In this scenario, the food arrangement placed in a school cafeteria can greatly impact the students' meal choices. For example, the person in charge (i.e., choice architect) of the school canteen can opt to display the dishes considering different objectives and intentions. This arrangement can follow the lucrative purposes (maximize profits), the most nutritious way (healthy food) or randomly choose the

food order. In other words, the choice architect can act as libertarian paternalist by arranging the food in a way that makes the students best off (Hausman & Welch, 2010). That is to say that due to an individual's inherent incentive to maximize their own benefit, the development of mechanisms to promote pro-social behaviours is of great importance for the well being of society (Capraro et al., 2019).

These concrete scenarios serve as illustrative examples of contexts where nudges can be effectively used to influence people decisions and, in optimal circumstances, even generate social benefits. Nonetheless, a new form of nudging has particularly gained prominence in recent years, specifically within the digital realm. Digital nudging, as the term suggests, is when individual's perception and actions within digital decision-making contexts, are directly guided through the use of user-interface design elements (Weinmann, 2016). This can be concerning, as organizations might exploit individual's diminishing attention spans and their limited ability to process information in digital contexts, in order to negatively influence them into making judging errors (Lembcke et al., 2019). Although this strategy could result in potential short-term gains for organizations, the use of digital nudging might lead to long-term repercussions such as loss of goodwill, negative publicity or even legal actions (Weinmann, 2016). This, therefore, underscores the importance of having choice architects who possess an awareness of ethical implications.

In retrospect, there is concern that some nudge policies might capitalize on individual's understated irrationality. However, it is also acknowledged that some policies, through good intentions and ethical considerations, could even guide individual's decision-making towards the most optimal path, strengthening therefore rational agency (Schmidt & Engelen, 2020). In spite of that, an unanswered question still remains: Does nudging always manipulate choice?

2.4. Two thinking systems

The underlying idea that individuals have two thinking systems became popular after Kahneman's contribution regarding prospect theory. In his study, he reasoned that people have, broadly speaking, a system 1 and a system 2, which respectively produce fast and slow thinking (Kahneman, 2011). In this regard, he alluded to the fact that system 1 operates in automatic mode, through intuitive, affective, and involuntary thoughts, whereas system 2 functions through more deliberate, reflective, and mental exhaustive thoughts. Simply put, system 1 runs automatically and only calls for system 2 to be engaged when complex computations or specific processing procedures are needed to solve a certain task in the moment (Kahneman, 2011).

Nonetheless, Kahneman (2011) pointed out that even if the division of labor between the two systems might be highly efficient, system 1 can have some limitations which can adversely affect

system 2 and thereby influence our decisions. This is because, as result of system 1 running on feelings, impressions, and short-term predictions, it may be subject to biases which can disrupt our rational behavior. Subsequently, system 1 cannot handle multiple distinct topics at once nor can comprehend data and statistics information. The reasoning for this is because, since individuals are indolent by nature, this sort of limitations are proven to be very influential to our decision making as most individuals will probably not trigger system 2 into action.

2.5. Nudgeability and transparency

The inconspicuous and discreet nature of nudges has become a significant topic of debate in recent literature, primarily due to ethical concerns (Rodrigues, 2021). By inherently exploiting human flaws in decision making processes, these concerns may be labeled as manipulative towards the individual's free will (Bovens, 2009; Hansen & Jespersen, 2013; Hausman & Welch, 2010).

When assessing nudges' transparency and effectiveness, it is generally thought that nudges work best "in the dark" as they most of the time target individual's system 1. Through nonconscious processes and function, many people do not spontaneously notice that they are being nudged or influenced (de Ridder et al., 2021).

On this subject, Bovens (2009) observed during his research on ethical questions surrounding nudge theory, that default options normally work best when they are not disclosed. At the time, he also noted that most people are less prompt to be fooled by overt nudges as they understand that they may be getting manipulated. This phenomenon can be called psychological reactance, a term brought up by Loewenstein et al. (2015) which means that reactance occurs when people feel that their freedom to behave may be threatened or loss. In other words, by alerting people about the imminent interference, it gives the individuals the idea that they are being pushed to a specific choice, which leads to a reactance on their part (Wortman & Brehm, 1975).

In this regard, transparency has been looked at from several perspectives regarding choice architecture. Thaler and Sunstein (2008) sided it with John Rawls (1971) and his contribution named the "publicity principle". The publicity principle bans governments from implementing or selecting a policy that they would not be able or willing to defend publicly to its own citizens. Thaler and Sunstein (2008) endorsed this policy because in their view it was both practical and gave the position of respect to the people. Other perspectives have focused on whether the individuals being nudged actually prefer being disclosed that they are being influenced or not.

Whether they have the cognitive capabilities to detect the nudge or not, Bang et al. (2018) pointed out that, just like when people get affected by visual illusions, even when they already know

how they work, individuals who are made aware of the choice architecture design may still get influenced by the nudge. So, awareness might not lower the effectiveness of the nudge.

In line with this, and conversely to all the previous reports and concerns, there are empirical findings that show disclosing nudges do not necessarily create reactance, or reduce their influential effect (Bruns et al., 2015; Loewenstein et al., 2015). This creates doubt as regards the existence of boundary conditions that foster or hamper the transparency of nudging in exerting its effects.

2.6. Nudgeability and benevolence

One of the main points that Thaler and Sunstein (2008) made, is that choice architects have the responsibility for organizing and orchestrating the context in which people make their respective decisions. However, if we follow the benevolent perspective implied in the Rawlsian “publicity principle”, nudging people into what is best for them requires choice architects who ultimately know what is best for others (Gigerenzer, 2015). On the other hand, Nives and Sareen (2020) noted that non-benevolent choice architects always want to leverage people’s behavior in a way that is beneficial for themselves, discarding any consideration about what is best for the individuals and their welfare.

If we look at financial markets, in which people trade for different financial arrangements, we can observe that nudges have been used to influence people’s behaviors for some decades now, and also in a malevolent way (Cai, 2020). Bucher-Koenen et al. (2011) pointed out that financial literacy and cognitive abilities play a large role on the individual’s success to return gains on financial markets and therefore could make them influenceable to external forces.

These considerations should be though prior to making a decision and in conjunction with the real nature and purpose of organizations. This is because an organization’s true intentions can differ from what is perceived by the public. If we look at a Edelman (2020) study, which compares levels of trust in different types on institutions, we can see that NGOs are perceived to be more ethically responsible than businesses. However, NGOs are also viewed as less competent in delivering on their promises compared to businesses.

In this context, Riley (2017) emphasized that the permissibility of nudges should be discussed and reflected, given the risk of epistemic injustice that nudges normally pose, because neither the moral acceptability of benevolent nudges, nor its individual vindication, appears enough to be morally accepted in the current world we live in. Sunstein (2021) strengthened his position when he reiterated that nudge theory should be in the interest of the people being nudged and aligned with their respective values. Additionally, he referred its strictly unethical character when nudges are not

compatible with human dignity, as well as when they are not subjected to political safeguards and defy constitutional understandings in a certain nation.

2.7. Individual's donation behavior

Individual's behavior concerning donations is often influenced by the aforementioned factors making it a subject of great significance to this topic. It is an area in which fundraisers and decision makers are interested to gain a deeper understanding of the characteristics that may prompt individuals to participate in acts of donation.

Mainardes et al. (2016) conducted an empirical investigation about the motivating factors and inherent characteristics that justify individual's reasoning underlying monetary donations. The study revealed that the most prominent variables justifying these donations were altruistic inclinations, as individuals with a stronger sense of altruism were more inclined to make charitable contributions, as well as the perception of the recipient organization's needs, as individuals were more motivated to donate when they recognized the organization's need to receive help.

These reports strengthened the position of Bennett (2003) around the significant impact and influence of individual's personal values on their selection of an organization to make donations to. In other words, the inclination of individuals to make donations is often shaped by their personal attributes, underscoring the importance of comprehending the main variables that precede the act of donation.

Behavioral interferences can be established to indirectly influence an individual's decision to donate. This can be done by fundraisers and institutions through the development of mechanisms rooted in nudge theory (e.g. pre-selecting a default option from a list) and thus taking advantage of individual's inertia. However, accurately anticipating the impact of setting one particular option, especially in the context of fundraisers, requires a comprehensive understanding of how these interferences influence choices. This includes an examination of circumstances that can enhance, diminish, or potentially reverse the effects of defaults (Goswami & Urmitsky, 2016).

2.8. Hypotheses

Considering the literature reviewed, the leading questions and respective hypotheses are the following:

First research question: To which extent does nudge disclosure harm the amount donated?

We posit that: Nudge disclosure has a harmful effect in the amount donated (Hypothesis 1).

Second research question: To which extent does benevolence favor the amount donated?

We posit that: Benevolence has a positive effect on the amount donated (Hypothesis 2).

Third research question: To which extent does nudge disclosure harm the amount donated depending on perceived benevolence?

We posit that: Benevolence mitigates the harmful effects from nudge disclosure (Hypothesis 3).

Chapter III - Method

3.1. Design

The method builds upon the experimental design used by Caroço (2022) to integrate “anchoring” as a nudging strategy. The factorial design is a preferred approach as it enables us to measure one or more factors at the same time (i.e., independent variables), while estimating each factor’s effects at levels of the dependent variable. All possible combinations between factors will lead to several experimental outcomes (Kalaian & Kasim, 2008) and this is a commonly accepted method in behavioral economics to infer causal nexus (Oxoby, 2006).

This design operates with scenarios that bring the participant into a situation that requires him/her to make a decision on a donation within the context of a fundraising campaign picture. We created four scenarios corresponding to a 2x2 factorial between-subject design.

The first factor concerns “covert vs. overt” nudging which refers to making the participant aware or not of the anchoring, while the second factor concerns “ascribed benevolence vs. ascribed self-interest” which is manipulated via the nature of the organization that is promoting the fundraising campaign. To depict a highly benevolent fundraiser, we have chosen an unnamed non-profit organization represented by a NGO (which we deemed to be taken as a benevolent charitable-like organization) while to depict a highly self-interested fundraiser, we have chosen a for-profit unnamed organization represented by a commercial bank (which we deemed to be taken as a self-interest organization). These choices were based on organizational trust studies that have found non-profit organizations such as NGO to be given more public trust and perceived as more ethical than for-profit businesses such as commercial banks (Edelman, 2020).

In total, the experiment comprehends four treatment groups with each participant randomly assigned (condition 1: benevolence*covert nudge; condition 2: benevolence*overt nudge; condition 3: self-interest*covert nudge; condition 4: self-interest*overt nudge). Table 1 shows the four experimental conditions.

Table 1 - Experimental conditions

	Overt nudging (after-decision nudge is disclosed)	Covert nudging (nudge never disclosed)
Benevolent (non-profit NGO, with altruistic social representations)	Webpage asking for donations to help children + with nudging disclosure	Webpage asking for donations to help children + without nudging disclosure
Self-interested (for-profit entity, with non-altruistic social representation)	Webpage asking for donations to help children + with nudging disclosure	Webpage intended to ask for donations to help children + without nudging disclosure

3.2. Stimuli

Firstly, to ascertain how much the ascribed benevolence is rightfully placed upon the NGO and Commercial bank, we pre-tested the participants with the following questions:

When the NGO is presented:

“In the present, the NGO Volunteer Doctors started to carry out humanitarian actions/campaigns with the aim of raising donations. In your understanding, please indicate in the spectrum below the real motivation of the NGO¹.”

When the Commercial Bank is presented:

“In the present, a Commercial Bank started to carry out humanitarian actions/campaigns with the aim of raising donations. In your understanding, please indicate in the spectrum below the real motivation of the Bank².”

The participant is requested to answer in a sliding rule from “100% self-interested” to “100% altruistic”, as shown below:

¹ In the original study, a portuguese version was used as follows: No presente ano, a ONG Médicos Sem Fronteiras, começou a realizar ações/campanhas humanitárias, com o objetivo de angariar doações. No seu entendimento, indique no espectro abaixo a real motivação da ONG.

² In the original study, a portuguese version was used as follows: No presente ano, um Banco Comercial com fins lucrativos, começou a realizar ações/campanhas humanitárias, com o objetivo de angariar doações. No seu entendimento, indique no espectro abaixo a real motivação do Banco.

100% self-interested _____ 100% altruistic

Then, according to the four conditions, we showed the stimulus that was built to enact the donation decision.

When conditions 1 and 2 (benevolence) are presented:

“This social campaign, created in line with its mission, aims to support children in Africa who lack clean drinking water³.”

When conditions 3 and 4 (self-interest) are presented:

“This social campaign, created due to the need to integrate ESG (Environmental, Social and Corporate Governance) policies into its agenda, is intended to support children in Africa, who lack clean drinking water⁴.”

After presenting both conditions, participants were exposed and asked to hypothetically disclose how much they would be willing to donate. This experiment was done through the following webpage:

Figure 1 - Stimulus



The image shows a screenshot of a donation webpage. The background features a close-up of a smiling young child. Overlaid on the right side is a white form with a blue header and footer. The header text reads: "Junte-se à nossa comunidade de doadores mensais, para apoiar crianças em África a ter água potável". Below this is a grey box with the text "Selecione uma quantia para doar uma única vez". The main content area contains a grid of buttons for donation amounts: 1€, 2€, 5€, 10€, 20€, and "Outra quantia". The 5€ button is highlighted in blue. Below the grid is a large blue button labeled "DOAR QUANTIA" and a checkbox labeled "Prefiro não doar". At the bottom, a grey box contains the text: "A contribuição de 5€ vai permitir que uma criança tenha água potável para uma semana".

³ In the original study, a portuguese version was used as follows: Esta campanha social, criada em linha com a sua missão, tem a intenção de apoiar crianças em África, com falta de água potável.

⁴ In the original study, a portuguese version was used as follows: Esta campanha social, criada em virtude da necessidade de integração de políticas ESG (Environmental, Social e Corporate Governance) na sua agenda, tem a intenção de apoiar crianças em África, com falta de água potável.

After showing the stimulus, in the condition “covert” nothing else is stated to the participant. However, in the “overt” condition, the anchoring nudge is made explicit so that the participant realizes the donation stimulus was not entirely transparent.

The disclosure was made with the following sentence:

“It may have gone unnoticed, but the previous donation request utilised a persuasion strategy of highlighting the value of the donation so that readers are willing to donate more. This type of nudge is called anchoring, because it will anchor the reader's attention to a value that might otherwise seem a bit excessive for such a donation.

Now that you've realised, would you be willing to keep the same donation or would you prefer to change it⁵?”

After explicitly disclosing to the participant that the stimulus was not entirely transparent, the following options were made available whether they want to keep the donation or change it.

1. I will keep my donation.
2. I will change it to: any amount between 0 and 100
3. I will not donate any quantity.

The anchoring nudge effect is produced by highlighting some information that anchors the participant to a specific value for donation.

3.3. Procedure

The experiment was designed and conducted in Qualtrics to expose participants to the conditions and stimuli mentioned above. This online survey tool allows random allocation of the conditions, with each participant being evenly allocated to one of the four conditions.

Participation was kindly and voluntarily requested through invitation sent either by private message, e-mail or social platforms like LinkedIn and WhatsApp. In this request, a message was written containing details about the context of the survey (not disclosing any information about the conditions), the anonymous participation, the confidentiality of data, as well as the expected time it

⁵ In the original study, a portuguese version was used as follows: Pode ter passado despercebido, mas o pedido de doação anterior utilizou uma estratégia de persuasão que consiste em destacar o valor da doação, de modo a que os leitores estejam dispostos a doar mais. Este tipo de nudge é chama-se anchoring, porque irá ancorar a atenção do leitor a um valor que de outra forma poderia parecer um pouco excessivo para uma tal doação. Agora que se apercebeu, estaria disposto a manter o mesmo donativo ou preferiria alterá-lo?”

would take to complete the task. This message was accompanied by either a link or the following image containing the QR Code of the survey page:

Figure 2 - QR code to survey



Lastly, a debriefing message was showed containing the researcher's contact information (in case of any doubt arising) as well a thanking message for the consideration and possible contribution to the study.

3.3. Sample

The sample comprises two groups: one with 86 individuals exclusively to test the manipulations as described in section 4.1. The second group is the conceptual model test group that comprises 141 participants, aged between 18 and 87 with average age of 38.5 years-old ($SD=16.3$), mostly females (56%) and educated (41.6 holds a Bachelor degree; 29.2% a Master degree). Most participants (86.8%) report to work in organizations unrelated to either the non-governmental sector neither finance industry and also to be sufficiently well off to donate money (64.7%), and actually having done so in the past year (49.6%).

The random allocation of participants to each of the four conditions originated similar sized groups as follows: Bank*covert comprised 36 participants, ONG*covert 36 participants, Bank*overt 35 participants, and ONG*overt 34 participants.

The groups are equivalent as regards age (ANOVA $F(3, 137)=1.679, p=.174$), gender ($\chi^2(3)=1.955, p=.582$), education (ANOVA $F(3, 133)=.206, p=.892$), available income for donations ($\chi^2(3)=1.480, p=.687$), donated last year ($\chi^2(3)=4.301, p=.231$) or working in either NGO or finance industries ($\chi^2(3)=2.606, p=.456$)

3.4. Data analysis strategy

The data analysis for this study was performed utilizing the IBM SPSS Statistic software. Firstly, tasks involving data screening were conducted in order to identify and exclude incomplete responses that lacked utility.

As commonly reported in experimental studies, we will start by computing the manipulation check which is tested with an Analysis of Variance (ANOVA) comparing the means for the conditions tested against its absence.

For descriptive purposes we will use frequencies and bivariate correlations with control variables (e.g. income, prosocial behaviors, donation habits, gender, age, education).

To test the hypotheses we will use Analysis of Covariance (ANCOVA) as the data analysis technique which is commonly used in this sort of research design to ascertain the main effects and interaction effects that indicate the existence of absence of the hypothesized relations.

Chapter IV - Results

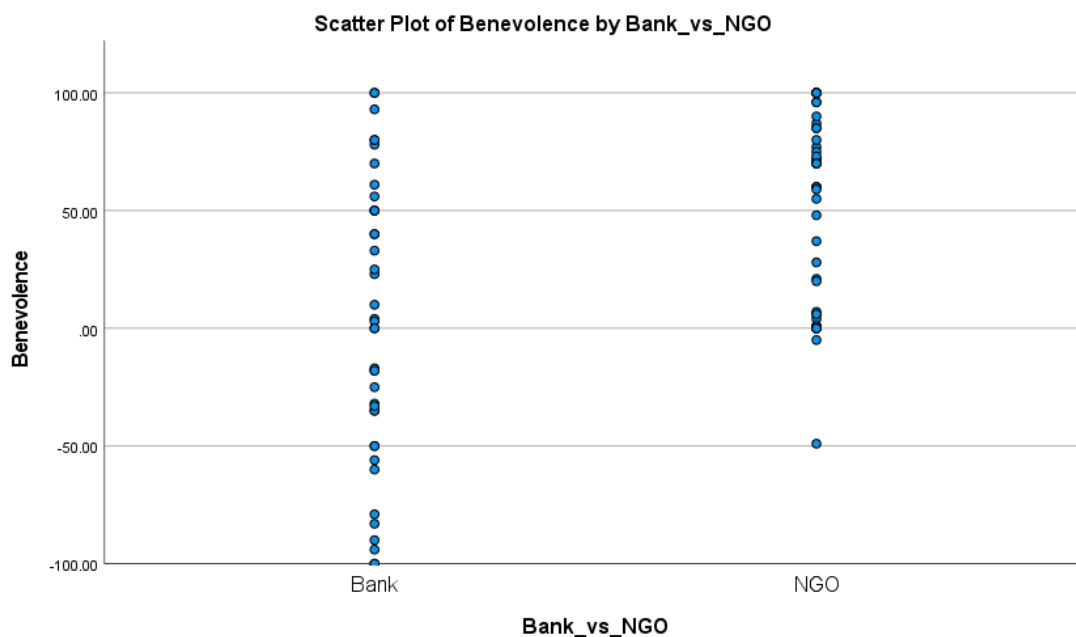
The results section starts by showing the manipulation tests which are intended to gauge to which extent the stimuli was successful in inducing the intended perception (ie. to what degree the stimuli induced a sense of low benevolence attached to Banks and high benevolence attached to NGO). After this, the descriptive and bivariate statistics are presented to proceed to findings pertaining to hypotheses testing.

4.1. Manipulation check

The manipulation test of the stimuli, which focused on benevolence, was conducted with a sample of 86 individuals. The conditions were presented in the exact same way as planned in the experiment with two scenario for Bank condition and two for NGO with the expectation of finding a higher mean for benevolence in the NGO condition, should be manipulation be successful.

The individuals were allocated in a random way to one of the conditions (Bank or NGO). Both groups have equivalent age ($F(1, 84)=0.916, p=.341$) and gender proportion ($\chi^2(1)=3.668, p=0.055$) scenario. As Levene test indicate there is no homogeneity of covariance matrices between both groups (Levene statistic (1, 84)=8.269, $p=0.005$), Robust tests of equality of means were conducted showing a clear mean difference ($M_{\text{bank}}=1.77, SD_{\text{bank}}=59.95, M_{\text{NGO}}=54.04, SD_{\text{NGO}}=40.13$; Brown-Forsythe (1, 66.559)=21.881, $p<.001$) depicted in Figure 3.

Figure 3 - Means plot



4.2. Descriptive and bivariate statistics

Descriptives show that the first donation mean ($m=5.07$, $SD=5.34$) is close to the nudging point (5 euros) which is supported by a one-sample t-test ($t(140)=0.157$, $p=.875$, 95% CI [-0.82; 0.96]), which is also observed in the final donation mean ($m=4.71$, $SD=5.31$, $t(140)=-0.634$, $p=.527$, 95% CI [-1.17; 0.60]).

The bivariate statistics show firstly that no significant correlation occurs between the sociodemographic variables (age, gender, education, industry, capability, and donated) with neither benevolence (BankNGO) not with transparency (covert_overt) which is the expected result for a randomly assigned situation as the one deployed in the study.

From these variables, capability is the only one that is associated with donation amount both as regards covert ($r=.214$, $p<.05$) and overt ($r=.208$, $p<.05$) which means this is an expectable outcome from more financial availability. Finally, benevolence is positively correlated with donation thus encouraging our hypothesis, although transparency is not associated, which discourages it.

Table 2 - Descriptive and bivariate statistics

	Mean	SD	1	2	3	4	5	6	7	8	9
1. Age	38.53	16.30	-								
2. Gender	56% F.	-	.095	-							
3. Educ	3.07	.85	-.165	.136	-						
4. Industry	87% no	.34	.201*	-.128	-.297**	-					
5. Capability	1.35	.48	.143	.151	.262**	.246**	-				
6. Donate	1.50	.50	-.083	.081	-.265**	.087	-.246**	-			
7. BankNGO	0.49	.50	.115	-.108	-.057	.119	-.002	.110	-		
8. Transparency	0.48	.50	.034	.047	.027	-.055	.101	-.139	-.007	-	
9. Donat_covert	5.07	5.34	.034	-.017	.050	-.030	.214*	-.129	.216*	-.040	-
10. Donat_final	4.71	5.31	.048	-.012	.030	.010	.208*	-.134	.217**	-.108	.889**

* $p<.05$; ** $p<.01$

4.3. Hypotheses testing

To which extent does nudge disclosure harm the amount donated?

Although literature has claims that support the harmful effect of nudge disclosure as well as a no-effect, the hypothetical answer to this guiding question is: Nudge disclosure has a harmful effect in the amount donated (Hypothesis 1).

To test this, we conducted a paired t-test on the comparison of first-donation amount versus after-disclosure donation amount. The test suggested a statistically significant difference for a one-tailed 95% confidence interval ($t(68)=-1.691$, $p=.048$) but this p-value does not hold under a 5000 repetition bootstrapping (95% CI [-1.579; 0.130] although this is a two-tailed test (Table 3). This casts doubt on the conclusion, but from a conservative viewpoint, doubts must lead us to rejecting the first hypothesis.

Table 3 - Paired samples test

	Paired Differences					Significance			
	Std. Error		95% CI Difference			t	df	One-Sided	Two-Sided
	Mean	Std.Dev.	Mean	Lower	Upper			p	p
2 nd Donat vs. 1 st Donat	-.724	3.552	.428	-1.579	.130	-1.691	68	.048	.095

As per the literature review, ascribed benevolence to the agent should be taken into account when judging on donation decisions.

To which extent does benevolence favor the amount donated?

The hypothetical answer to this guiding question is: Benevolence has a positive effect on the amount donated (Hypothesis 2).

To test this hypothesis, an ANCOVA is deployed taking the first donation as the dependent variable, the Bank vs. NGO condition as a independent variable representing low versus high benevolence respectively, and age, gender, education, industry, having donated, capability to donate as covariates. The p-value was set to 95% confidence interval and the bootstrapping confidence is 95% with 5000 repetitions (Hayes, 201

According to Rutherford (2011), ANCOVA requires that the conditions have equivalent variances (homogeneity of variances) as well as being homoskedastic (meaning, the errors must not be dependent on the independent variable). To test homogeneity of variances between the conditions, we ran Levene’s test that is designed to test the null hypothesis that the error variance of the dependent variable is equal across groups.

When a first test was conducted, Levene’s statistic indicated this assumption was not met. In fact, the descriptives showed the first donation variance was much larger in the NGO condition ($s=38.7\%$) than in the Bank conditions ($s=16.4\%$). This was corrected by means of computing the inverse of the variance (Kutner et al., 2005) and create a dummy variable that was subsequently used as a weight in the Weight Least Squares function associated with the ANCOVA.

This corrective procedure was successful and the resulting Levene statistic is $F(1, 133)=.123$, $p=.726$ (Table 4) allows the ANCOVA analysis to proceed.

Table 4 - Levene’s test of equality of error variances

F	df1	df2	p-value
.123	1	133	.726

Dependent Variable: First_donation

The second requirement, was checked with the White test for Heteroskedasticity that tests the null hypothesis that the variance of the errors does not depend on the values of the independent variables (White, 1980). The analysis showed a X^2 statistic of -13.5 for 30 degrees of freedom which has an exact p-value of 1, thus ruling out issues emerging from homoskedasticity.

The resulting between-subjects effects showed that when an organization is perceived as being more benevolent (represented NGO in the stimuli as compared to Bank) the individuals tend to donate more (Table 5) thus supporting the second hypothesis.

Table 5 - Test of between-subjects effects for first donation

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Correct Model	1363.356 ^b	7	194.765	1.989	.061	.099	13.924	.758
Intercept	557.516	1	557.516	5.694	.018	.043	5.694	.658
Age	29.700	1	29.700	.303	.583	.002	.303	.085
Gender	86.820	1	86.820	.887	.348	.007	.887	.154
Education	19.399	1	19.399	.198	.657	.002	.198	.073
Industry	.011	1	.011	.000	.992	.000	.000	.050
Donated	24.232	1	24.232	.247	.620	.002	.247	.078
Capacity / Slack	497.097	1	497.097	5.077	.026	.038	5.077	.609
BankNGO	532.491	1	532.491	5.439	.021	.041	5.439	.639
Error	12434.751	127	97.911					
Total	26140.190	135						
Corrected Total	13798.107	134						

a. Weighted Least Squares Regression - Weighted by Weight

b. R Squared = .099 (Adjusted R Squared = .049)

Because donation decisions under a nudge effect that can or cannot be disclosed, do not occur independently of the ascribed benevolence to the agent, an interaction effect was proposed.

To which extent does nudge disclosure harm the amount donated depending on perceived benevolence?

The hypothetical answer to this guiding question is: Benevolence mitigates the harmful effects from nudge disclosure (Hypothesis 3).

To test this we conducted an ANCOVA with the final amount donated (after disclosure of nudging in half the sample) using Bank vs. NGO and Disclosure as independent variables, and controlling for age, gender, education, industry, capability and having donated. We kept the weight least squares procedure as the same issues pertaining to homogeneity of variances between conditions were found. Subsequently, no issues remained as regards homogeneity as well as heteroskedacity in the data matrices.

Because 2*2 conditions (cover vs overt * low vs high benevolence) with the cases that have explicit overt donation values would hamper the power of the test due to small sample size, we tested for mean differences between the two conditions that answered overt as against those that were not given that chance. An ANOVA showed no mean differences for the covert donation ($F(1,139)=0.219$, $p=.641$) thus suggesting no reason to anticipate a divergent behavior with the one observed for the overt groups. Therefore, we transported the covert donation values to the overt donation both for the Bank and NGO. This procedure may lower the chances of finding differences where they may occur thus raising Type II error in hypothesis testing (which is the most conservative one).

Table 6 - Test of between-subjects effects for final donation

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Correct Model	2154.941b	9	239.438	2.468	.013	.151	22.216	.916
Intercept	729.837	1	729.837	7.524	.007	.057	7.524	.777
Age	11.604	1	11.604	.120	.730	.001	.120	.064
Gender	145.341	1	145.341	1.498	.223	.012	1.498	.229
Education	6.792	1	6.792	.070	.792	.001	.070	.058
Industry	2.808	1	2.808	.029	.865	.000	.029	.053
Donated	660.984	1	660.984	6.814	.010	.052	6.814	.736
Capacity / Slack	42.697	1	42.697	.440	.508	.004	.440	.101
BankNGO	766.072	1	766.072	7.898	.006	.059	7.898	.796
Transparency	143.507	1	143.507	1.479	.226	.012	1.479	.227
interaction	140.067	1	140.067	1.444	.232	.011	1.444	.222
Error	12125.015	125	97.000					
Total	25453.118	135						
Corrected Total	14279.956	134						

a. Weighted Least Squares Regression - Weighted by Weight

b. R Squared = .099 (Adjusted R Squared = .049)

To test the hypothesis, we conducted an ANCOVA taking the final donation as the dependent variable, the benevolence (Bank vs. NGO) condition, and the Transparency (overt vs. covert) condition as independent variables as well as their interaction term, while controlling for age, gender, education, industry, having donated, capability to donate as covariates. Findings (Table 6, previous page) show no interaction effect between transparency and benevolence ($F(1, 125)=1.444, p=.232$) thus rejecting the third hypothesis.

Still, when one considers the means for final donations, they substantially vary not only with perceived benevolence ($M_{\text{Bank}}=3.61$ vs. $M_{\text{NGO}}=6.16$) but also with disclosing vs. not-disclosing the nudge, as shown in Table 7.

Table 7 - Descriptive statistics for Benevolence * Disclosure

BankNGO	Transparency	Mean	Std. Deviation	N
Bank	Covert	4.56	11.62	36
	Overt	2.62	7.60	34
	Total	3.61	10.09	70
NGO	Covert	6.05	10.21	34
	Overt	6.29	10.02	31
	Total	6.16	10.04	65
Total	Covert	4.98	10.97	70
	Overt	3.64	9.43	65
	Total	4.33	10.32	135

a. Weighted Least Squares Regression

Within the Bank condition, the means for covert condition is 4.56 ($SD=11.6$) while for overt condition it is 2.62 ($SD=7.60$). Within the NGO condition, the means for covert condition is 6.06 ($SD=10.21$) while for overt it is 6.29 ($SD=10.02$). Although this may, at first sight, suggest a clear interaction where the decision to disclose nudging is harmful when organizations are not given high-benevolent intentions, the large data dispersion as indicated by standard deviations advises otherwise, and the ANCOVA interaction effect does support this caveat.

Chapter V - Discussion and conclusion

Nudging has attracted much attention ever since the rational limits of decision making have been highlighted by Herbert Simon. This topic is especially important in Economics and Management because decision making is at the micro foundations of phenomena studied in Economics, Organizational Science and Management. Although it has become a popular topic in research and professional education, there are many issues that are still in open debate. Namely, to which extent is any form of paternalism truly ethical? What are the conditions that maximize its effectiveness? What are the characteristics of the target that interfere with its effectiveness? And how is the ascribed intention of the source of nudging important for its effectiveness? These all entail some sort of tradeoff between the effectiveness of nudging and ethics.

Therefore, this study took on the challenge of answering a relevant question within nudge research: are nudges effective only as far as they are not realized consciously by the target? Put otherwise, are nudges effectiveness hampered by transparency?

Because transparency is an important principle in ethics, the nature of nudging brings ethical implications due to the subconscious induction of a decision violating the principle of informed consent (Kiener, 2021). The blurred ethical frontier in nudging is expressed in the entanglement with “libertarian paternalism” as used by Thaler and Sunstein (2008) because all expressions of paternalism entail a restriction in freedom enforced by third parties. The ideal situation would be to guarantee the “self-declared interest of the nudged” individuals (Hansen, 2016) by means of making the nudge overt, i.e. consciously aware of it.

Ongoing research diverges as regards the answer (cf. de Ridder et al., 2021 vs. Bruns et al., 2015) but the harmful effects of disclosure are more in line with the subconscious nature of nudging. As an addition to this discussion, it is advantageous to consider the nature of the source of nudging. As in any persuasion process, the ascription of a benevolent nature to an influencer will improve the chances of complying. Therefore, the most favorable situation to nudge could be the one that is made covert and ascribed by the target to a source that is positively evaluated, e.g. for its benevolent intention. This implies two main effects and an interaction effect. The first main effect concerns transparency, the second benevolence, and the interaction between these.

Firstly, results on the effect the stimuli had upon donation, suggest that disclosing nudges is not harmful, thus rejecting hypothesis 1. Although findings are not entirely clear (as the p-value falls liminally below the threshold of 0.05, but the bootstrapping limits do cross the zero value) it is more

prudent to not reject the null hypothesis and thus, we must interpret these results as being supportive of the disclosure thesis. Therefore, our results go in line with Bruns et al. (2015) and Loewenstein et al. (2015). The overt nudge was made after the covert decision and in this circumstance the psychological reactance would have been noticed because not only has the manipulation been made explicit as it was made explicit after its occurrence, which means the participant could have felt tricked and therefore, react negatively to it. Still, this did not happen which reinforces Bang et al. (2018) view that being aware of the manipulation might not suffice to escape the architecture of decision.

The second hypothesis, that proposed benevolent sources of nudging receive higher compliance from targets, was supported by findings. This goes in line with the importance of trust in complying to any persuasive action. Because benevolence is one of the dimensions that compose relational trust (Colquitt & Salam, 2012) it is logical to corroborate its importance for the acceptance of a donation request.

Lastly, the third hypothesis was built upon the assumption that the first hypothesis would receive support (“Benevolence mitigates the harmful effects from nudge disclosure”) but although the first hypothesis did not behave in the expected way, this does not preclude the testing of the third hypothesis, eventually revised as “there is an interaction effect of benevolence on the effectiveness of nudge transparency (overt versus covert anchoring as a nudging strategy) in such a way that benevolence increases the chances of transparency being related to the donation amount”. Findings support an interaction effect where low benevolence in the overt condition produces the lowest mean donation value ($M = \text{€}2.62$) while covert nudging produces a closer-to-target value ($M = \text{€}4.56$). Still, this effect is not observed in the high benevolence condition where the donation is independent from the nudge being shown covert or overt and substantially higher than any of the other conditions ($M_{\text{total}} = \text{€}6.16$).

This effect brings light to the divergent findings pertaining to transparency of the nudge. On the one hand, those that claim high transparency to be harmful do find a suitable empirical support in our low-benevolence conditions. However, they will not find such support in the high-benevolence condition, and the same is valid vice-versa for those claiming transparency has no harmful effect. Thus, both theses may be right in the sense that they are context dependent and eventually it would be interesting to review findings based on the categorization of ascribed benevolence to the source of nudging (if it is visible).

These findings cannot be directly transferred to all sort of decisions that nudging targets. This is solely for money donation decisions for philanthropic purposes. In many other cases, e.g. nudging to

exercise or change eating habits, transparency and benevolence may behave differently. If a given organization has a vested interest in fostering the consumption of any service or good linked to the nudge, then it is quite likely the individual will react negatively if such low benevolence is inferred.

Conversely, if the nudging stems from a neutral institution or any organization that may not directly benefit from the nudge, then it is more likely that the high benevolence condition applies. The fact that the interaction pointed towards an effect only in the low-benevolence condition suggests that organizations that may be seen as self-interested in the nudging will be more careful in keeping it covert if they want to be effective. In such cases, ethical issues do apply as the target would behave differently should they realize the source of the nudge.

5.1. Limitations

A couple of limitations were clear in this study. Firstly, the study experience uses a small sample size, which included only 141 valid responses, considering the factorial design it corresponds to 34 to 36 responses per condition. This creates a caveat as regards the ability to comprehensively represent the range of questions that were under analysis, as it increases the likelihood of sampling errors. Still, the random assignment of conditions is a step towards mitigating it. Secondly, the absence of a control group prevented the assessment of causal relationships and excluded the chance to determine a baseline comparison to a group of people who would not face the nudge effect of the stimuli.

Another limitation stems from the laboratory-like nature of this experiment. As usual, laboratory contexts are known for its ability to offer better control over the factors at play but to detach from real-world settings where the natural decisions actually happen. In this case, the participants are aware of the research and their behavior may be truly influenced by this awareness. Also, the fact that their donation was mocked may give an optimistic view of the real donation because in real world they would truly have to give away the amount donated. For ethical and legal reasons we could not opt for a hidden strategy where individuals would be asked to donate real money without being duly informed ahead, as this is clearly deceitful and also falls within fraud class requests even if it was disclosed later and the donation returned. Still, we trust the eventual inflation that mock situations may produce are equally distributed by all groups and thus this does not harm the test of the transparency and benevolence effect. Eventually, only the true amount of estimated donation. It is also important to keep in mind that the stimuli was designed based on a real donation campaign by an international renowned institution that uses the anchoring nudge, which offers some assurance about its effect.

5.2. Future research

To overcome these limitations future research may be designed with a baseline group that works as a control group so to understand if their donations without any sort of nudging would suggest a true nudging effect. Still, future research may explore what level of benevolence is ascribed by the nudged individuals when the source of nudging is not declared. In this circumstance it might be useful to bring in individual dispositions such as personality to consider the propensity to trust as well as other pro-social variables that can influence the decision to donate and how much to donate.

Overall, despite the limitations mentioned, we think the study offers insight as regards the relative harmless nature of disclosing nudges for organizations taken as benevolent, which adds to their ethical grounds. Conversely it brings concern for those organizations that are intending to nudge for self-vested interests, e.g. commercial purposes, which is more prone to covert strategies as the overt ones are more harmful for the intended objectives. These are the cases where ethical breach is more likely. Still, we believe this study shed light into the divergent claims about transparency effects upon nudging found in literature. Future research may benefit from considering the ascribed benevolence linked to the source of the nudge, which should be taken as a boundary condition but also a desired goal as regards corporate reputation.

References

- Bang, H., Shu, S. B., & Weber, E. U. (2018). The role of perceived effectiveness on the acceptability of choice architecture. *Behavioral Public Policy*, 2, 1–21. <https://doi.org/10.1017/bpp.2018.1>
- Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., ... & Galing, S. (2017). Should governments invest more in nudging?. *Psychological science*, 28(8), 1041-1055.
- Bennett, R. (2003). Factors underlying the inclination to donate to particular types of charity. *International Journal of Nonprofit and Voluntary Sector Marketing*, 8(1), 12-29.
- Bovens, L. (2009). The Ethics of Nudge. In T. Yanoff-Grüne and S. O. Hansson (eds). *Preference Change* (pp. 207–220), Dordrecht: Springer.
- 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, 41-59.
- Bucher-Koenen, T., & Ziegelmeyer, M. (2011). *Who lost the most? Financial literacy, cognitive abilities, and the financial crisis*. ECB working paper, No. 1299, European Central Bank, Frankfurt a. M.
- Cai, C. W. (2020). Nudging the financial market? A review of the nudge theory. *Accounting & Finance*, 60(4), 3341-3365.
- Camerer, C. (1999). Behavioral economics: Reunifying psychology and economics. *Proceedings of the National Academy of Sciences*, 96(19), 10575-10577.
- Camerer, C. F., & Loewenstein, G. (2004). Behavioral economics: Past, present, future. *Advances in behavioral economics*, 1, 3-51.
- Capraro, V., Jagfeld, G., Klein, R., Mul, M., & de Pol, I. V. (2019). Increasing altruistic and cooperative behaviour with simple moral nudges. *Scientific reports*, 9(1), 11880.
- Caroço, C. B. (2022). Nudging strategies and acceptance of human resource policies: The case of framing and social proof heuristic [*Dissertação de mestrado, Iscte - Instituto Universitário de Lisboa*].
- Colquitt, J. A., & Salam, S. C. (2012). Foster trust through ability, benevolence, and integrity. In E. Locke (ed). *Handbook of principles of organizational behavior* (2nd ed.) (p. 389-404), Chichester: John Wiley & Sons.
- Cooper, E. J. (2017). To nudge or not to nudge: Promoting environmentally beneficial behaviors. *Bard Center for Environmental Policy*, 10
- de Ridder, D., Kroese, F., & van Gestel, L. (2022). Nudgeability: Mapping conditions of susceptibility to nudge influence. *Perspectives on Psychological Science*, 17(2), 346-359.
- Edelman (2020). 2019 Edelman trust barometer: Global report. Author. https://www.edelman.com/sites/g/files/aatuss191/files/2019-02/2019_Edelman_Trust_Barometer_Global_Report.pdf


- Egidi, M., Marris, R. L., & Viale, R. (Eds.). (1992). *Economics, bounded rationality and the cognitive revolution*. Edward Elgar Publishing.
- Felsen, G., Castelo, N., & Reiner, P. B. (2013). Decisional enhancement and autonomy: public attitudes towards overt and covert nudges. *Judgment & Decision Making, 8*(3), 202-213.
- Gigerenzer, G. (2015). On the supposed evidence for libertarian paternalism. *Review of Philosophy and Psychology, 6*(3), 361-383.
- Gigerenzer, G., & Selten, R. (Eds.). (2002). *Bounded rationality: The adaptive toolbox*. MIT press.
- Goswami, I., & Urminsky, O. (2016). When should the ask be a nudge? The effect of default amounts on charitable donations. *Journal of Marketing Research, 53*(5), 829-846.
- 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.
- 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.
- Hausman, D. M., & Welch, B. (2010). Debate: To nudge or not to nudge. *Journal of Political Philosophy, 18*(1), 123-136.
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications
- Jones, B. D. (1999). Bounded rationality. *Annual Review of Political Science, 2*(1), 297-321.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux
- Kalaian, S., & Kasim, R. (2008). Research Design. In P. J. Lavrakas, (Eds.). *Encyclopedia of Survey Research Methods* (pp. 724-731). Thousand Oaks: Sage
- Kiener, M. (2021). When do nudges undermine voluntary consent?. *Philosophical Studies, 178*(12), 4201-4226.
- Kosters, M., & Van der Heijden, J. (2015). From mechanism to virtue: Evaluating Nudge theory. *Evaluation, 21*(3), 276-291.
- Kutner, M.H., Nachtsheim, C.J. , Neter, J. & Li, W. (2005). *Applied Linear Statistical Models*, McGraw-Hill, New York.
- Langlois, R. N. (1990). Bounded rationality and behavioralism: A clarification and critique. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft, 146*(4), 691-695.
- Lehner, M., Mont, O., & Heiskanen, E. (2016). Nudging—A promising tool for sustainable consumption behaviour?. *Journal of Cleaner Production, 134*, 166-177.

- Lembcke, T. B., Engelbrecht, N., Brendel, A. B., & Kolbe, L. M. (2019, June). *To Nudge or not to Nudge: Ethical Considerations of Digital nudging based on its Behavioral Economics roots*. In Proceedings of the 27th ECIS, Stockholm.
- Loewenstein, G., Bryce, C., Hagmann, D., & Rajpal, S. (2015). Warning: You are about to be nudged. *Behavioral Science & Policy*, 1(1), 35–42.
- Mainardes, E. W., Laurett, R., Degasperi, N. C. P., & Lasso, S. V. (2016). What motivates an individual to make donations of money and/or goods?. *International Review on Public and Nonprofit Marketing*, 13, 81-99.
- Oxoby, R. (2006). Experiments and behavioral economics. In M. Altman (ed). *Handbook of Contemporary Behavioral Economics* (pp. 441-456). New York: M.E. Sharpe.
- Rabin, M. (1998). Psychology and economics. *Journal of Economic Literature*, 36(1), 11-46.
- Rawls, J. (1971/1999). *A theory of justice* (2nd ed). Cambridge: Harvard University Press.
- Rebonato, R. (2012). *Taking liberties: A critical examination of libertarian paternalism*. London: Palgrave Macmillan.
- Rodrigues, J. C. (2021). *Impact of Disclosing Nudges in a Financial Context* (Doctoral dissertation, Universidade Catolica Portuguesa (Portugal)).
- Rutherford, A. (2011). *ANOVA and ANCOVA: A GLM approach* (2nd ed.). Hoboken: John Wiley & Sons.
- Schmidt, A. T., & Engelen, B. (2020). The ethics of nudging: An overview. *Philosophy compass*, 15(4), e12658.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63(2), 129-138.
- Simon, H. A. (1990). Bounded rationality. In *Utility and probability* (pp. 15-18). Palgrave Macmillan, London
- Sunstein, C. (2021). Much anew about "nudging". (R. Fusaro, & J. Sperling-Magro, Interviewers). <https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/much-anew-about-nudging>
- Thaler, R., & Sunstein, C. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven: Yale University Press.
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, 185(4157), 1124-1131.
- Tversky, A., & Kahneman, D. (1989). Rational choice and the framing of decisions. In B. Karpak and S. Zionts (eds). *Multiple criteria decision making and risk analysis using microcomputers* (pp. 81-126). Berlin: Springer.

- Weinmann, M., Schneider, C., & Brocke, J. V. (2016). Digital nudging. *Business & Information Systems Engineering, 58*, 433-436.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*, 817-838.
- Wortman, C. B., & Brehm, J. W. (1975). Responses to uncontrollable outcomes: An integration of reactance theory and the learned helplessness model. In *Advances in experimental social psychology* (Vol. 8, pp. 277-336). Academic Press.

Appendix A

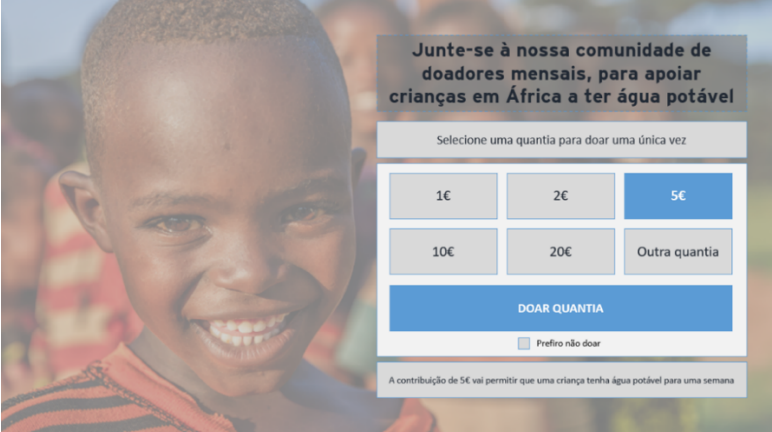
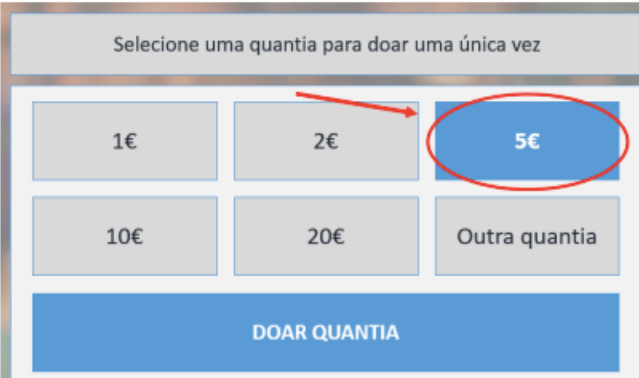
1. Benevolence conditions (Covert+Over)

Manipulation check	<p>No presente ano, a ONG Médicos Voluntários, começou a realizar ações/campanhas humanitárias, com o objetivo de angariar doações.</p> <p>No seu entendimento, indique na barra abaixo a real motivação da ONG.</p> <p>(Arraste o indicador para a posição que corresponde ao que pensa)</p>
Context	<p>Esta campanha social, criada em linha com a sua missão, tem a intenção de apoiar crianças em África, com falta de água potável.</p> <p>(Clique na seta para ver a campanha)</p>
Stimuli	<p>Hipoteticamente, ao observar esta campanha, por favor selecione a opção que optaria.</p>  <p>The image shows a donation interface for 'Médicos Voluntários'. It features a background image of a smiling African child. Overlaid on the right is a grey box with the text: 'Junte-se à nossa comunidade de doadores mensais, para apoiar crianças em África a ter água potável'. Below this is a white box with the instruction 'Selecione uma quantia para doar uma única vez'. There are six buttons for donation amounts: 1€, 2€, 5€, 10€, 20€, and 'Outra quantia'. The 5€ button is highlighted in blue. Below the buttons is a blue button labeled 'DOAR QUANTIA' and a checkbox labeled 'Prefiro não doar'. At the bottom, a small text line reads: 'A contribuição de 5€ vai permitir que uma criança tenha água potável para uma semana'.</p>

<p>Nudge Disclosure</p> <p>(Only for overt condition)</p>	<p>Pode ter passado despercebido, mas no pedido de doação anterior, a ONG Médicos Fronteiras utilizou uma estratégia de persuasão que consiste em destacar um valor de doação, de modo a que os leitores estejam dispostos a doar mais.</p> <p>Este tipo de nudge chama-se ancoragem, porque irá ancorar a atenção do leitor a um valor que de outra forma poderia parecer um pouco excessivo para uma tal doação.</p> <div data-bbox="609 622 1254 999" data-label="Image"> </div> <p>Agora que se apercebeu, estaria disposto a manter o mesmo donativo ou preferiria alterá-lo?</p> <ul style="list-style-type: none"> • Não quero alterar a minha doação • Quero alterar a minha doação
---	---

2. Self-Interest conditions (Covert+Overt)

<p>Manipulation check</p>	<p>No presente ano, um Banco Comercial com fins lucrativos, começou a realizar ações/campanhas humanitárias, com o objetivo de angariar doações.</p> <p>No seu entendimento, indique na barra abaixo a real motivação do Banco.</p> <p>(Arraste o indicador para a posição que corresponde ao que pensa)</p>
<p>Context</p>	<p>Esta campanha social, criada em virtude da necessidade de integração de políticas ESG (Environmental, Social e Corporate Governance) na sua agenda, tem a intenção de apoiar crianças em África, com falta de água potável.</p> <p>(Clique na seta para ver a campanha)</p>

<p>Stimuli</p>	<p>Hipoteticamente, ao observar esta campanha, por favor selecione a opção que optaria.</p> 
<p>Nudge Disclourse (Only for overt condition)</p>	<p>Pode ter passado despercebido, mas no pedido de doação anterior, o Banco Comercial utilizou uma estratégia de persuasão que consiste em destacar um valor de doação, de modo a que os leitores estejam dispostos a doar mais.</p> <p>Este tipo de nudge chama-se ancoragem, porque irá ancorar a atenção do leitor a um valor que de outra forma poderia parecer um pouco excessivo para uma tal doação.</p>  <p>Agora que se apercebeu, estaria disposto a manter o mesmo donativo ou preferiria alterá-lo?</p> <ul style="list-style-type: none"> • Não quero alterar a minha doação • Quero alterar a minha doação

3. Final characterization

1.	Habilitações Literárias: <ul style="list-style-type: none">• Ensino Básico (até ao 9º ano)• Ensino Secundário (12º ano)• Licenciatura• Mestrado• Doutoramento
2.	Trabalha numa organização não governamental (ONG) ou no setor financeiro? <ul style="list-style-type: none">• Sim• Não
3.	Tem atualmente rendimentos próprios e margem financeira para fazer doações? <ul style="list-style-type: none">• Sim• Não
4.	No decorrer do último ano, fez algum tipo de doação? <ul style="list-style-type: none">• Sim• Não