Negativity Bias and Framing Effect in Charity Donations: a Field Experiment

Antoine Soetewey

r0443256

Thesis submitted to obtain the degree of

MASTER OF SCIENCE IN ECONOMICS

General Economics Major

Promotor: Prof. Dr. Jo Swinnen
Assistants: Andrea Guariso and Mara Squicciarini

Academic year: 2013-2014

Abstract

Negativity bias is commonly known as the psychological phenomenon whereby people tend to be more sensitive to negative rather than neutral or comparably positive instances (e.g., feelings, experiences, images, objects, personal traits, etc.). Another distinct and well established phenomenon in the literature is the framing effect, which states that people tend to be inconsistent in their behavior when facing similar but differently expressed information.

The present thesis addresses these two widespread phenomena and intends to apply them to a practical framework. Consistent with the related literature suggesting that the two concepts hold in theory, we decided to implement a field experiment that should have allowed to measure whether these theoretical concepts hold in reality, and more specifically in charity donations. Concretely, the field experiment should have helped charity institutions to improve the effectiveness of their fundraising messages by illustrating the way they should frame their key message or slogan.

The results of the pilot experiment were, however, mixed and quite unexpectedly not in line with our hypotheses. Consequently, after having considered every options, we decided to collect the results of the pilot and to not implement the full-scale experiment. Two possible explanations can be given for the source of our mixed results: Either (i) our assumptions and hypotheses are wrong and/or do not hold in the domain of charity donations, or (ii) the structure and the design of the experiment are inappropriate to

measure the effects of the negativity bias and the framing effect. We concluded that the structure of the experiment is more likely to be the cause of our mixed results. Nonetheless, it might be that for some reasons that still need to be discovered, fields involving real monetary choices and real consequences on individuals revolves around different underlying concepts and theories than fields requiring hypothetical (i.e., *as if*) choices. This seems a promising avenue for future research.

Whether the experimental results are consistent or not with our hypotheses pose no problem. This thesis can be considered as a prior research that would be advisable to take into account if a similar experiment had to be performed in future research. Moreover, the issues we encountered during the pilot can also be seen as intriguing behaviors that could lead to other studies, or reveal the interest from other domains than charity donations.

Acknowledgments

I owe this Master's thesis to the help and support of many kind people around me, to only some of whom it is possible to give particular mention here.

Above all, this thesis would not have been possible without the support of my principal supervisor, Prof. Jo Swinnen. I thank him for his trust and giving me the opportunity to write about this topic. Financial support from the Katholieke Universiteit Leuven is also gratefully acknowledged.

I am grateful to my teaching assistants; Andrea Guariso and Mara Squicciarini. Firstly, I thank them for providing me with the related literature at the beginning of November 2013. This effort considerably helped me to acquire knowledge in this domain and significantly reduced the length of the prior research. Secondly, I thank them for having given me their time and support throughout the writing of this thesis and numerous meetings. Furthermore, I thank Andrea Guariso for his thoughtful comments on earlier versions of this thesis. He considerably improved the quality of the final version.

I express my gratitude to Prof. Sabrina Bruyneel, professor in Behavioral Economics at Katholieke Universiteit Leuven who allowed me to write my term paper on this topic and provided me with helpful advice, which greatly helped me to start the thesis early enough and on solid grounds. I would also like thank her for a meeting just before the decision of not going to the full-scale experiment was taken, a last meeting to make sure that we considered every options before taking the final decision.

Contents

Introduction

While most countries see their economies growing such that the Gross Domestic Product (GDP) per capita is increasing, the market for charitable donations seems to have decreased over the last years. Indeed, according to the Charities Aid Foundation [2012], the global average participation in donating money[1] has globally decreased from 30% in 2007 to 28% in 2011.

One could argue that people are not less altruistic but they have had to adapt their behaviors to a higher level of income instability caused by the economic crisis of 2008-2012. However, Pharoah and Tanner [1997] showed that the proportion of UK households giving to charity fell by 5 percentage points between 1974 and 1993−94, so that the economic conjuncture would only partially explains a decrease in charity donations.

Moreover, a more disturbing fact is that there is no correlation between the stage of the country's development and the level of charitable donations. For example, the Charities Aid Foundation [2012] ranked 148 countries based on their participation level in donating money with Belgium ranked as 34th, with a participation level of 38%. This score is way below the participation level of Indonesia or Thailand which are both developing countries but still have a surprisingly high participation level of 71%. One may therefore wonder what influences an individual's choice to donate money to charities together with the amount of his donation.

---

[1]The participation level in donating money is the percentage of people that donated money during one month prior to the survey among all the people aged 15 and older.

In 2012 in the US, total giving to charitable organizations was about 2 percent of GDP, with a total giving of $316.23 billion [The Giving Institute, 2013]. The market for charitable giving includes three main players [List, 2011]: (i) Donors supplying resources to charitable organizations, (ii) charitable organizations developing strategies to solicit resources from donors, and (iii) government establishing the tax treatment, the level of government grants, and what public goods to contribute itself. The present thesis focuses on the donors and the charitable institutions, as well as the interactions between the two.

Previous studies have shown that, for instance, attracting attention using personalized messages, setting defaults, matching donations, rewarding with small gifts, exercising peer pressure, finding the most effective period, etc. are all techniques that encourage people to donate and thus imply a higher level of participation [Behavioural Insights Team, 2013]. Relatively to the main topic of this study, a lot of work has been done on the negativity bias and the framing effect but substantially less work has been undertaken in applying these two phenomena in monetary charity donations. The aim of this thesis is to address this question from both a theoretical and a practical point of view. For this reason, the thesis is divided into two chapters.

Chapter I theoretically reviews the negativity bias and the framing effect and gives examples relevant to charity donations. This first part will lay sound foundations for Chapter II, which practically analyzes and discusses the field experiment that will enable to empirically measure the impacts of the two phenomena in the case of monetary charity donations. The objective is that the findings can then be used by managers and charity institutions to determine which type of message or slogan they should provide to people depending on their main objective.

Chapter II is structured as follows: The first section will introduce the purposes and the

benefits of this field experiment. The second section will consist of the intuitive assumptions and hypotheses constructed on the related literature and the theoretical part. The third section will explain the most appropriate methods and procedures to follow in order to measure the relevance of our assumptions. The fourth section will address the hypotheses. The fifth section will analyze our results. The sixth section will describe some potential improvements and their limitations. Finally, a discussion and some implications in the real world will conclude Chapter II.

Chapter 1

Negativity Bias and Framing Effect

1.1 Negativity Bias

Although the concept of negativity bias is quite recent, it has been introduced more than three decades ago by another closely related concept which is the loss aversion [Kahneman and Tversky, 1979]. This concept comes from the work on prospect theory which features different models of decision making under risk.[2]

A central conclusion in this domain is that preferences are significantly driven by the expected utility from the changes relative to a neutral reference point, rather than by the final state of welfare or wealth.

Another conclusion describes the fact that losses (i.e., relative changes that make things worse) loom larger than gains (i.e., relative changes that make things better, or improve them), such that people tend to be loss averse. In other words, loss aversion implies that individuals will be more affected by losses than gains of similar values in absolute terms. Much work has been undertaken in this domain and the existing evidence suggest that, at

---

[2]Some suggest that this work is probably the starting point of behavioral economics and that it is still the most influential in decision making.

least with regard to the domain of monetary valuations,[3] losses are psychologically twice as powerful as gains [e.g., Kahneman et al., 1991].

The concept of negativity bias is derived from the concept of loss aversion but remains slightly different. Like loss aversion, the negativity bias refers to the fact that people will weight more heavily negative events rather than positive ones. In other words, more than one positive instance will be necessary in order to cancel out the effects of one negative instance. An event is considered as negative if it leads to or threatens an individual with adverse outcomes, while an event is considered as positive if it leads to favorable outcomes or an opportunity to lead to favorable outcomes [Taylor, 1991].

However, the negativity bias is not restricted to the tendency for individuals to be influenced in a larger extent by negative instances rather than by relatively equal positive instances. Indeed, the negativity bias also refers to the general tendency for undesirable stimuli, unpleasant information or harmful events to generate both more persisting effects and more cognitive activity than desirable outcomes [Baumeister et al., 2001]. First, Thomas and Diener [1990] found that people underestimate the frequency of positive affect, while this is not the case with negative affect. They argue that this is because of the relative weakness of positive emotional experiences which makes them more forgettable. Second, Finkenauer and Rimé [1998] also found that people tend to report more easily and more quickly negative memories compared to neutral or positive memories. Third, Gilbert et al. [1998] found that people overestimate the extent to which events in their lives will affect them and that this overestimation is larger for negative instances than positive ones. These three findings are consistent with the view that bad things produce larger and more lasting impacts than good things [Baumeister et al., 2001].

Moreover, while negative events seem to mobilize affective and cognitive resources to

---

[3]The demonstration of loss aversion in other domains than monetary ones is unfortunately limited because a metric scale is required to measure the objective equivalence of positive and negative events.

a larger extent, Taylor [1991] claims that there appears to be an even more interesting asymmetry in individuals' responses resulting from negative and positive events, that is, people will try to minimize or even erase the undesirable effects of a negative threat by a response that elicits opposite feelings (i.e., good feelings) while they will not go through the same process (or to a significantly lower extent) in case of a positive event.

In his study of parachutists, Epstein [1967] observed that, once parachutists safely landed on the ground, a positive emotional reaction of relief or profound relaxation appeared and offset the emotions resulting from the negative emotions (e.g., the threat of a technical problem that would cause injuries or an immediate death). However, no comparable research addresses the opposite process, that is, people experience offsetting negative emotions after experiencing positive emotions. The lack of this type of observation in the related literature suggests that it may not exist [Taylor, 1991].

More generally, Rozin and Royzman [2001] argue that four types of negativity bias exist: (i) Negative potency, (ii) greater steepness of negative gradients, (iii) negativity dominance and (iv) greater negative differentiation. These four aspects of negativity bias are four different manifestations of the general bias. They should be considered as complementary concepts that might work in different aspects of life. Indeed, they are not mutually exclusive such that if one works the others does not or that they are conflicting with each other. Therefore, the following paragraphs and explanations have to be interpreted as four different representations of the same broader concept of negativity bias.

1.1.1   Negative Potency

The claim of the principle of negative potency specifies that negative events are of higher salience, i.e. more potent, compared to positive events of equal objective magnitude. This is the core principle of the loss aversion phenomenon. This phenomenon is concretely

described by the endowment effect: the willingness-to-accept (WTA)[4] set by an individual for a specific good that he possesses tends to be larger than the willingness-to-pay (WTP)[5] for the same item if he does not possess it. In other words, it means that individuals ascribe more value to goods they own than to goods they do not own merely because they own them and independently of the real valuation/value they ascribe to the good. Theoretically, the endowment effect is due to the fact that humans are loss averse and forgoing something owned feels like loosing. Furthermore, the WTA reflects both the true value the individual ascribes to the good and the compensation needed for giving up the ownership, while the WTP reflects both the value the individual ascribes to the good and the value he ascribes to the gain of ownership. Therefore, the difference between the WTA and the WTP equals the loss of ownership subtracted by the gain of ownership. If there was no loss aversion or negative potency, taken in absolute terms the difference between the WTA and WTP would be equal to zero so the selling prices would easily meet the buying prices. However, and as has been highlighted on several occasions in the literature, negative potency and loss aversion exist, that is, the feeling of loss of ownership dominates the feeling of gain of ownership, which results in a larger value ascribed to owned goods than to identical non-owned goods.

A well-known illustration of the endowment effect in the literature is the experiment conducted by Kahneman et al. [1990]. In this experiment, half of the participants were randomly offered a coffee mug. After a clear establishment of the ownership, subjects were given the opportunity to sell it through four successive artificial markets. Whereas the Coase theorem[6] predicts that about half of the mugs would be traded, the observed volume of traded mugs was always significantly lower. Indeed, since the mugs were randomly

---

[4]The WTA is the "selling price", that is, the price an individual is willing to receive for giving up a good he possesses. Because it means he would have to give something, it is considered as a loss.

[5]The WTP is the "buying price", that is, the price an individual is willing to pay for acquiring a good he does not possess yet. Because it means he would obtain something, it is considered as a gain.

[6]The Coase theorem describes that if there are no transaction costs, the market will determine an efficient outcome regardless of the initial allocation [see Coase, 1960, for further details].

distributed we expect that, on average, half of the subjects that belongs to the group that value mugs the most were given one while the other half were not. Likewise, on average half of participants in the group that value mugs the least received one and the other did not. Therefore, it implies that half of the mugs should have been traded, with the mug-owners from the group that like mugs the least selling to the non-owners from the group that like mugs the most. However, in practice less than half of the expected volume was traded, far from the 50 percent predicted volume of trade.[7]

These findings are explained by the fact that the WTA were substantially higher than the WTP, reducing the number of buyers' prices that met sellers' prices.[8] This robust instantiation of loss aversion demonstrates the underlying principle, that is, people will demand much more to undergo a loss than to benefit a gain of identical objective magnitude.

1.1.2   Greater Steepness of Negative Gradients

This type of negativity bias specifies that "negative events grow more rapidly in negativity as they are approached in space or time than do positive events" [Rozin and Royzman, 2001, p. 298]. This principle of gradient effect differs from the negative potency described above because at low levels negative events are not systematically more potent than comparable positive events. Indeed, Cacioppo and colleagues [Cacioppo and Berntson, 1994, Cacioppo et al., 1997, 1999] suggest that the effects of negative events dominate the effects of positive events only when both are strong. In other words, whereas negative potency may arise from higher subjective value of negative events at all levels, such that the two functions (with respect to negative and positive events) may have the same slope but the negative

---

[7]The number of mugs traded in the four markets was 4, 1, 2 and 2 respectively; out of 22 mugs distributed at the beginning of the experiment.

[8]Kahneman et al. [1990] found that the WTA were approximately twice as high as the WTP. This figure can be much larger; for instance, Carmon and Ariely [2000] found that participants' hypothetical WTA for tickets for National Collegiate Athletic Association (NCAA) basketball games were 14 times higher than their WTP.

function may have a higher intercept, the findings on negative gradients suggest in fact that the negative functions have a lower intercept but a higher slope [Rozin and Royzman, 2001].

Brown [1948] and Miller [1944] demonstrate this greater steepness of negative events respectively in the animal learning literature and in terms of the source of motivation. First, Brown [1948] measured the rat's tendency to pull away from a negative event (e.g., shock), at the end of an alley, at different points in the alley. He also measured the rat's tendency to pull toward a reward (e.g., food) at different points in the alley. After a comparison between the two measures, he found "steeper negative than positive gradients in terms of distance from site of feeding or shock" [Rozin and Royzman, 2001, p. 304]. Second, Miller [1944] pointed out that the intensity of response increases as the closeness to the source of external negative event increases. Conversely, the intensity of response does not increase much as the closeness to the source of positive event increases. As an example, he argues that closeness to the source of an electric shock (i.e., negative event) should increase intensity of response. Nonetheless, the underlying motivation of hunger does not vary as much when one approaches food (i.e., positive event).

1.1.3 Negativity Dominance

Although the negativity dominance is closely related to negative potency, the negativity dominance is stronger in as far as the principle is more constraining. Besides the fact that negativity dominance must meet the condition of negativity potency – negative occurrences are more potent than comparable positive ones – it must also meet the condition that "the holistic perception and appraisal of integrated negative and positive events [...] is more negative than the algebraic sum of the subjective values of those individual entities" [Rozin and Royzman, 2001, p. 298-299].

The following situation helps to understand the difference between these two principles.[9]

---
[9]In order to hold, the impacts must be taken in absolute terms.

If losing 10€ has a greater (negative) impact than the (positive) impact when winning 10€, that is, losing 10€ is worse than winning 10€ is good, then we have negative potency because the negative outcome is subjectively more potent than its positive counterpart although both are of equal objective magnitude. Nonetheless, if on the one hand the (negative) impact of losing 10€ is equal to the (positive) impact of winning 15€, that is, losing 10€ is as bad as winning 15€ is good, and if on the other hand the perception of the integrated values of losing 10€ and winning 15€ is more negative than the sum of the subjective values of these two occurrences taken as two distinct entities, then we have an instance of negativity dominance.

Similarly, Kanouse and Hanson Jr. [1987] framed the negativity dominance such that it occurs if a negative aspect has the ability to interfere with the pleasure derived from a positive aspect when both are combined. For example, "irresponsible father" is often judged more negatively than "irresponsible" [Rokeach and Rothman, 1965]. Furthermore, it illustrates negativity dominance because the evaluation of the integrated values of both aspects combined (i.e., "irresponsible father") is more negative than the algebraic sum of the subjective values of those two aspects when judged individually (i.e., "irresponsible" and "father").

Moreover, it is useful to distinguish the synchronic (simultaneous) and the diachronic (successive) manifestations of negativity dominance.

On the one hand, the synchronic type addresses the evaluation of positive and negative inputs as components of a whole. For example, under these conditions the impression of a person will be disproportionately more influenced by the negative traits than the positive traits that characterize this person.

On the other hand, the diachronic type concerns the compensation of positive events by negative ones and inversely. For instance, under these conditions the number of lives that a murderer has to save in order to compensate a single act of murder will be dramatically

large. Likewise, a tremendous number of good behaviors are usually needed to be forgiven for one bad behavior.

In the domain of close relationships, Gottman [1995] proposed that positive and good interactions between two individuals must be at least, on average, five times as frequent as negative and bad interactions for a relationship to last. These findings show that negativity bias is even stronger than loss aversion with respect to close relationships.

1.1.4   Greater Negative Differentiation

The last facet of the negativity bias is that responses to negative stimuli are more differentiated and more complex than reactions to correspondent positive stimuli. This phenomenon of greater negative elaboration is illustrated by the more varied and more diversified vocabulary that is employed to describe the characteristics of negative stimuli compared with the vocabulary that is used to describe the quality of positive phenomena [Peeters, 1971]. Another example of greater negative differentiation is the wider range of potential emotions to negative than positive outcomes, as argued by many authors [Wundt and Judd, 1907, Titchener, 1907, Carlson, 1966, Averill, 1980, Izard, 1971].

One may therefore wonder what causes people to be more sensitive to negative events. Currently, there are different hypotheses in the literature which have been put forward and that may explain the reasons of this bias. First, since positive events are more likely to occur than negatives ones, it is natural for human beings and animals to assume the effects of the positive while being cautious to the effects of the negative [Peeters, 1971]. Second, natural selection, which occurred through out the evolution played a large role in the development of the differences in reactions to positive and negative occurrences [Darwin, 1859]. Indeed, organisms that survived were the ones who were able to deal as quickly as possible with dangerous negative instances caused either by predators or by the

environment. Nowadays, although this mechanism may be less crucial in terms of survival for humans, it certainly remains decisive in terms of survival for businesses. This leads to the tendency for people to treat and to adapt to positive instances with lesser urgency and less importance.

From this first comprehensive definition, we can now undoubtedly suppose that negative occurrences should have a greater impact on individuals versus positive occurrences. If we apply this assumption to charity donations, we can expect that people will be more likely to respond if we provide them with negative fundraising messages than positive ones. Negative fundraising messages refer to messages that lead individuals to think of undesirable events. For instance, the message "20,000 people will die of tuberculosis if we do not support them" is a negatively framed message. On the contrary, positive messages result in people thinking of desirable outcomes. For example, the message "20,000 people will be saved from tuberculosis if we support them" is a positively framed message. The definition of the framing effect will be helpful in determining how the way pieces of information are displayed plays a role in decisions.

1.2   Framing Effect

The framing effect has been introduced by Tversky and Kahneman [1981] with the Asian disease problem. In this experiment, all respondents are provided with the following information [Tversky and Kahneman, 1981, p. 453]:

> "Imagine that the US is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows."

The first group of participants ($N = 152$) is faced with the following problem. The percentages in brackets indicate the percentage of subjects that actually chose the option

and the total number of respondents for each problem is denoted by N.

"Program A: If Program A is adopted, 200 people will be saved." [72%]

"Program B: If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved." [28%]

"Which of the two programs would you favor?"

The second group ($N = 155$) is provided with the following problem:

"Program C: If Program C is adopted 400 people will die." [22%]

"Program D: If Program D is adopted there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die." [78%]

"Which of the two programs would you favor?"

This experiment illustrates the fact that people tend to change their behavior, that is, they shift their preferences from a certain outcome in the positively framed problem to a gamble in the negatively framed problem, even though the options remain identical both in terms of probabilities and in terms of outcomes.[10] Indeed, one can easily see that both problems are similar in their effectiveness and that the only difference between the two is that the first is positively framed, such that the outcomes of the two options are quantified by the number of lives saved, while the second one is negatively framed, such that the outcomes are quantified by the number of lives lost. According to the assumption of rationality in economic theory, an agent that chooses program A should choose program C. Likewise, an agent that chooses program B should choose program D. The proportion of people choosing each program illustrates the inconsistency and the irrationality of some agents, which violates one of the main assumptions in economic theory.

---

[10]This particular pattern in behavior is often referred as *risk aversion* in the domain of gains and *risk taking* in the domain of losses.

More generally, the framing effect which describes the reversal of preferences has been defined by three major underlying theories: (i) The framing of acts, (ii) the framing of contingencies[11] and (iii) the framing of outcomes. Again, the following three subsections describe reversals of preferences caused by three different types of variation in the framing. They should also be considered as complementary concepts that might work in different aspects of life or different areas. They are presumably not mutually exclusive or conflicting with each other, although less bounded that the four manifestations of the negativity bias detailed above. Therefore, the following paragraphs and explanations have to be interpreted as three different possible representations of the same broader concept of framing effect.

1.2.1   The Framing of Acts

Roughly speaking, the framing of acts specifies that people usually fail to combine options when facing concurrent decisions that would otherwise result in greater outcomes. In other words, they often independently choose their preferred option for each problem without taking into account the possible conjunction of different options, which would lead to better outcomes.

For example, in an experiment conducted by Tversky and Kahneman [1981] participants ($N = 150$) have to make two decisions, where the first decision (i) concerns the choice between:

>"A. a sure gain of $240." [84%]
>
>"B. 25% chance to gain $1000, and 75% chance to gain nothing." [16%]

And the second decision (ii) concerns the choice between:

>"C. a sure loss of $750." [13%]
>
>"D. 75% chance to lose $1000, and 25% chance to lose nothing." [87%]

As mentioned *supra*, the percentage in brackets indicates the percentage of subjects

---

[11]Based on my read of the literature, this is also referred as the isolation effect.

that actually chose the option and the total number of respondents for the problem is denoted by *N*.

Because the two decisions (i and ii) were presented together, the participants actually had the possibility to choose one pair of options between (i) A and C, (ii) B and C, (iii) A and D, and (iv) B and D. If the problem is now no longer presented as a pair of two separate choices (decisions (i) and (ii)) but rather presented as a combination of two decisions together with the percentage of subjects that choose each of the combinations resulting from the pair of the two separate decisions, it would give us the following output:

A & C. a sure loss of \$510. [10.92%][12]

B & C. 25% chance to win \$250, and 75% chance to lose \$750. [2.08%][13]

A & D. 25% chance to win \$240, and 75% chance to lose \$760. [73.08%][14]

B & D. 6.25%[15] chance to win \$1000, 56.25%[16] chance to lose \$1000, and 37.5%[17] chance to win/lose nothing. [13.92%],[18]

From this perspective, we observe that most respondents (73.08%) chose the combination A & D, while the least popular combination (B & C) was chosen by approximately 2% of the participants. Since the combination B & C yields strictly better outcomes than the combination A & D,[19] this experiment shows that people apparently fail to combine options and rather take decisions independently although the outcomes would be greater if possibilities were jointly considered.

---

[12]84%*13%.

[13]16%*13%.

[14]84%*87%.

[15]25%*25%.

[16]75%*75%.

[17]25%*75% + 75%*25%.

[18]16%*87%.

[19]Indeed, *ceteris paribus* the win is larger and the loss is smaller. Moreover, the expected outcomes for B & C and A & D are respectively −\$500 and −\$510.

1.2.2   The Framing of Contingencies

The framing of contingencies illustrates two phenomena of choice: the certainty effect and the pseudocertainty effect.

The certainty effect corresponds to the fact that people, in the negatively framed domain, tend to overestimate the aversiveness of losses that are certain compared to losses that are merely probable. For instance, in the domain of number of lives lost which is undoubtedly relevant in the sphere of charity donations, the certainty effect is illustrated by the fact that most people will prefer the gamble "80% chance to lose 100 lives" over "a sure loss of 75 lives" while they will prefer "10% chance to lose 75 lives" over "8% chance to lose 100 lives" [Tversky and Kahneman, 1981, p. 455].

The pseudocertainty effect occurs when people dismiss a causal contingency or a sequential formulation when evaluating options. For instance, still in the negatively framed domain and as illustrated in Figure 1.1, people tend to consider options A ("Sure loss of 75 lives") and B ("80% chance to lose 100 lives") without considering the previous conditional probability of 0.1 (*Nature* at the first node of Figure 1.1).
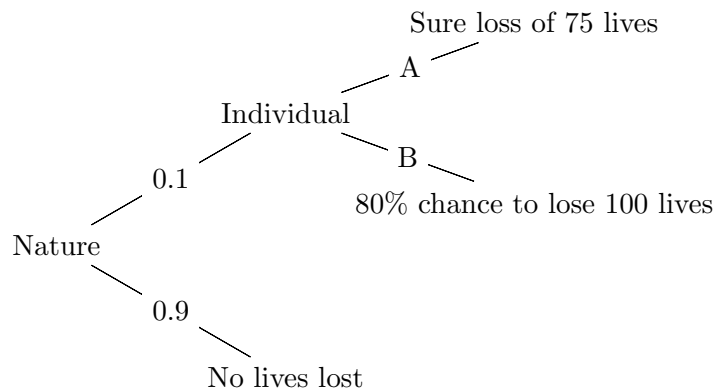

Figure 1.1: Framing of contingencies

Indeed, findings show that most people tend to have the same preference as if the first conditional probability of 0.1 did not exist, such that they prefer option B ("a 80% chance

of losing 100 lives") over option A ("a sure loss of 75 lives"). Nonetheless, as mentioned *supra* in the certainty effect, people would in contrast, prefer "10% chance to lose 75 lives" over "8% chance to lose 100 lives", which are identical outcomes if the whole tree of Figure 1.1 is taken into account [Tversky and Kahneman, 1981].[20]

Another interesting behavior pattern – with respect to the framing of contingencies and which is useful in the domain of charity donations – is that people tend to see policies that completely eliminate a specific and more restrictive risk as more attractive than policies that only reduce (and thus do not eliminate) the overall probability related to a risk at a broader level.

Slovic et al. [1982] found that people evaluating a vaccine that would reduce the probability of contracting a disease from 20% to 10% more attractive if it is described as being able to eliminate one of the two virus strains that generate similar symptoms than if it is presented as effective in 50% of the occurrences. In accord with the pseudocertainty effect theory, people tend to prefer full protection against half of the virus strains over a comparable reduction in probability of contracting the disease.

1.2.3   The Framing of Outcomes

As we have seen earlier with loss aversion, people tend to evaluate negative variations from the reference point[21] more heavily than positive comparable variations, that is, the convex value function for losses is usually steeper than the concave value function for gains.[22] This assumption explains why losses loom larger than gains as well as why people are more willing to forego a discount than accept a surcharge [Thaler, 1980].

---

[20] 10% chance to lose 75 lives is similar to 0.10*(Sure loss of 75 lives). Likewise, 8% chance to lose 100 lives is similar to 0.10*(80% chance to lose 100 lives).

[21] An individual's reference point is its initial level of outcome before any (positive or negative) variation and that is therefore judged neutral.

[22] The convexity and concavity of the value functions are explained by the decreasing marginal rates.

More generally, Tversky and Kahneman [1986] claim that people will evaluate the difference between two options to a lesser extent when it is framed as an advantage of one option rather than a disadvantage of the other option. For instance, a difference that favors option A over option B can be seen as an advantage of A or a disadvantage of B, depending on the neutral reference point. If A is the reference point, the difference between the two outcomes will be seen as a loss, conversely if B is the reference point the difference will be seen as a gain. Due to loss aversion, the difference will loom larger when A is the reference point and the difference considered as a loss [Tversky and Kahneman, 1981]. Consistent with this finding, this would therefore suggest that a salesman could for instance make the client buy the high-quality product among two when this product is presented as the norm and when the emphasis is on the potential loss of quality attached to the low-quality product.

Those interesting results demonstrate that framing a message in different ways sometimes causes individuals to be inconsistent, such that preferences reversal arises. More generally, this section showed that both the negativity bias and the framing effect have a significant impact on individuals' choices. The beauty of this thesis is the fact that we tried to put them together by looking at the impact of negative framing. Hence, we present them as two different phenomena that intersect with each other.

If wisely used, a sound and deep understanding of both concepts should allow all fundraisers to improve their campaign's efficiency, that is, levy more funds through a higher participation rate and a higher amount of money collected from donors. Although, I believe that applying this understanding will yield higher return with respect to the participation rate than regarding the amount of donation.

This specific prediction is based on the fact that I assume that it is easier for non-donors to start donating a small amount than for donors to increase their donations by the same

amount. Indeed, this seems to hold if we assume convexity of preferences,[23] diminishing marginal utility and the fact that agents usually derive utility mainly from their revenues but from altruistic behaviors as well, as authored by Dunn et al. [2008]. Moreover, we need to assume a similar willingness-to-donate across both those who do not donate and those who already donate. Otherwise, the following problem arises: people that are not donating and those that are donating are by definition different people. Those that do not donate might not do so because for instance they really do not want to let part of their revenues go and they do not really care at all about charity. While those that are already donating reveal that they care at least slightly about donating. So with different willingness-to-donate or different "type" of people, there is no reason to believe that it should be easier to move the former and not the latter.

Having said this, from the point of view of non-donors that start donating a given amount of money we have on the one hand, a decrease in utility induced by losing a part of their revenues and on the other hand, an increase of utility induced by the pleasure of giving. Because of the convexity of preferences and the diminishing marginal utility hypothesis, it may be that the gain in utility is larger than its loss. Whereas, from the point of view of donors that increase their donations by the same given amount, we have on the one hand an increase in utility induced by the pleasure of giving more than what they are used to, and on the other hand a decrease of utility induced by a second reduction of their income. In this case, we cannot assume whether the gain in utility will be larger than the loss. Nonetheless, we can assume that the loss in utility for non-donors will be smaller than the one for the donors. Likewise, we can assume that the gain in utility for non-donors will be larger than the one for the donors. This should lead to a larger increase in participation rate than the increase in amount collected, ceteris paribus.

---

[23]Intuitively, convex preferences refer to the idea that averages are better than the extremes, that is, a bundle consisting of two different goods will be preferred over a bundle consisting of two similar goods, i.e., $\forall \mathbf{x}^1, \mathbf{x}^2$ and $\mathbf{x}^3 \in X, \mathbf{x}^1 = (x_1, x_2) \succ \mathbf{x}^2 = (x_1, x_1) \sim \mathbf{x}^3 = (x_2, x_2)$.

Chapter 2

Field Experiment

2.1 Purposes and Benefits

The main purpose of this type of experiment is to help charitable institutions to find what are the incentives for people to donate. Furthermore, it aims to help them in the design of campaigns that generate the highest level of charitable donations. It could also attempt to shed light on observed trends in donations following noticeable and dramatic events. To illustrate this we could find the root cause of the recent increase in NGO's, funding for agriculture despite the financial crisis and following the big attention media paid to the world food crisis. Likewise, it could be used to better understand the behavior of some NGOs and International Organizations that seem to switch the content of their messages depending on punctual specific circumstances [Swinnen and Squicciarini, 2012].

The findings should also be advantageously used to manage social welfare through donations and beneficiaries satisfaction management. Indeed, beside the fact that charities increase the welfare of the beneficiaries thanks to countless beneficial actions that are undertaken with the money that is collected, donations also increase the welfare of the donors. Dunn et al. [2008] have shown that individuals spending money on others experi-

enced greater happiness than those who do not.[24]

Above we mentioned a clear distinction between participation and amount donated; a difference between effects on the intensive versus extensive margin of donors (whether we attract new donors or we attract those donating already to give more). However, our experiment only refers to the participation as we expect to receive standard 1€ donations. So the goal of the experiment is to focus on the choice of the donors.

2.2   Assumptions and Hypotheses

Based on Chapter I, we suppose that negatively framed messages will have higher impacts on respondents and should therefore produce a higher level of participation. This leads us to our main hypothesis:

> HP1: The effects of negatively framed messages on potential donors will be greater than the effects of positively framed messages. Since respondents will be more sensitive to negatively framed messages we expect that they will more likely donate when faced with negatively framed information than when faced with positively framed information.

This assumption is, however, not sufficient to construct the most appropriate design for the experiment and to present very precise messages that could be compared with each other in order to measure the ones that yield the best outcomes. Fortunately, several research about the effectiveness of messages on individuals have been undertaken so we will refer to them – and more especially to the variables that influence individuals' responses – to construct our own messages for the field experiment. The reason why I refer to them at this point is that we are now focusing on the design of the experiment and those references

---

[24]We assume that welfare includes happiness. See Greve [2008] for a detailed meaning of *welfare*.

confirm that our approach should work.

First, Baumeister et al. [2001] demonstrate that loss aversion does not only hold with monetary rewards or punishment but that it can be generalized to many areas. In order to measure whether loss aversion applies to non-monetary occurrences, they review the hypothesis that negative events have a greater impact on individuals than comparably positive events. For all the areas they covered, they were unable to find evidence where good events were stronger than bad ones so they conclude that "bad is stronger than good at a pervasive, general level" [Baumeister et al., 2001, p. 355]. The work of Rozin and Royzman [2001] offers a complementary range of examples that supports the more general principle that negativity bias is also true for non-monetary facets featured in the human existence.

Second, Hilbig [2009] reported that negatively framed information is generally deemed more valid. In order to measure whether negative information was perceived as truer than positive information, he asked respondents to judge the validity of some statements. The two groups were provided with the same intrinsic information but one group received the positive versions (e.g., proportion of marriages lasting 10 years or longer) while the other group received the negative versions (e.g., proportion of marriages divorcing within 10 years). As one could expect, the negatively framed statements were judged significantly more valid and those results were replicated for three different statements. Furthermore, the results were not altered when individual disparities in pessimism and optimism were controlled.

However, Burt and Strongman [2005] argue that the effectiveness of message framing is not only determined by the negative–positive context of the message but rather by the type of evidence with which the message is given. According to their results, negatively framed fundraising messages are more effective when they are combined with abstract informa-tion, whereas positively framed messages are more effective when combined with anecdotal

information. Abstract information generally refers to statistical evidence or the number of patients who suffer from particular symptoms, whereas anecdotal information generally tells the story of one patient (most of the time a young patient because it is considered as more appealing and thus more persuasive) who suffers from the symptoms. In addition, they found that for a message to be highly effective, it should highlight that others have already contributed so to make donors feel that the goal of the charity will be reached and that their donation will indeed be used to fund a cause. It should also stress the usefulness and the impact of each individual donation so to make potential donors feel that their donations will not be worthless among all other donations.

It is clear that many studies have been conducted on the effectiveness of fundraising messages or more generally, on the persuasive effects of messages on individuals. However, I have summarized only the findings of the studies which I am aware of and that are linked with the negativity bias and the framing effect since it is the subject of the present thesis.[25]

These conclusions together with the definitions of the negativity bias and the framing effect enabled us to draw the methods and procedures for the field experiment.

2.3 Methods and Procedures

This study manipulated the negativity bias and the framing effect in a fundraising project. It was hypothesized that people would donate to a larger extent to projects that were displaying a negative message. Specifically, message framing (i.e., negative versus positive) were expected to have interactive influences on donors' choices. We thus took into account the vast literature we described in designing this intervention. The design of the experiment was informed from existing findings in the literature and in particular we considered the

---

[25]See List [2008], Allen and Preiss [1997] among others for an overview of the findings on messages' persuasiveness in field experiments.

existing body of evidence to design fundraising activity that could differently solicit donors.

2.3.1   Design and Participants

In order to measure whether negativity bias and framing effect have significant impacts on monetary donations, we proceeded as follows:

> A group of two people erected a stand in the city center of Leuven, consisting of a table with a flag displaying the logo of the Katholieke Universiteit Leuven. Since the experiment was conducted in Leuven; a Dutch-speaking city, one of the experimenter was fluent in Dutch, and the other experimenter was fluent in English and French, in case some respondents were not Dutch-speakers.

It is important to note that we obtained the authorization from the university to use its affiliation for this experiment and to expose the official banners. The stand was located in a commercial street close to typical places where people pass by, so as to attract as many people as possible, but still in a quite confined area otherwise people could easily avoid us. Moreover, we also took the idea from the existing literature of providing a small gift (i.e., a pen in our case) to solicit donations.

> In addition, two posters were held in front of the stand (approximately two meters away) displaying a message that informed people that we were collecting money for charities. The message was: "Buy a KU Leuven pen for 1€ and send 2€ to charity". One poster displayed the message in English and another displayed the message in Dutch (See Appendix A).
>
> Beforehand, we had prepared four boxes on which we had written the following messages:
>
> 1. A-: "In 2010, about half of the HIV-affected people in low- and middle-income countries did not have access to the lifesaving antiretroviral ther-

apy.[26] This project aims to prevent more people from dying of HIV."

2. A+: "In 2010, about half of the HIV-affected people in low- and middle-income countries did have access to the lifesaving antiretroviral therapy. This project aims to help more people survive HIV."

3. B-: "Without the vitamin A supplements, the chance of dying for a child in low- and middle-income countries is 24% higher.[27] This project aims to prevent more people from dying of malnutrition."

4. B+: "With the vitamin A supplements, the chance of survival for a child in low- and middle-income countries is 24% higher. This project aims to help more people survive malnutrition."

The four messages were written in black for the experiment, color was just added here for explanatory purposes and to emphasize the positive versus the negative framing. The letter and the sign in front of the messages (A-, A+, B- and B+) were also not visible for the donors. This was added here so we can easily refer to them later. Furthermore, these four messages vary in terms of framing but length was held constant in order to ensure both that the messages are as similar as possible and that the only variable that varies is the framing (i.e., the variable we want to measure).

Moreover, pre-printed vouchers were necessary in order to record basic information of the donors we attracted (gender, amount donated, etc.) and to have a proof of the amount collected (See Appendix B for the voucher).

The pre-printed vouchers enabled us to observe preferences of individuals depending on easy-to-observe characteristics. With basic econometric models, this information could

---

[26]According to UNAIDS and WHO estimates, 47% (6.6 million) of the estimated 14.2 million people eligible for treatment in low- and middle-income countries were accessing lifesaving antiretroviral therapy in 2010, an increase of 1.35 million since 2009 [UNAIDS, 2011].

[27]An academic study involving more than 200,000 children between the ages of 6 months to 5 years found

be used in further studies to predict the choices of individuals depending on their characteristics. It also helped us in increasing our trustworthiness, by showing that we have something pre-arranged to carefully record donations.

Furthermore, we gave the possibility of leaving the emails on the vouchers for three main reasons:

  (i) to keep participants informed about the whole amount raised and about the details of the final donation and also to provide proof that the donation was actually made, if they wanted to remain informed,

 (ii) to increase our credibility, and

(iii) to refer to further communications in case people had other questions.

Once everything was set up, we waited for people to come. We kept on the table the KU Leuven pens, the vouchers that donors had to fill in after the donation and two of the four boxes that were alternated. Indeed, we always kept two boxes on the table, but every 30 minutes we interchanged the boxes to have the following combinations: (i) A+B-, (ii) A+B+, (iii) A-B+ and (iv) A-B-. We displayed each combination more than once in order to increase the precision of the experiment. This would increase its precision because, due to the changing environment (weather, time of the day, number of people present in the street, closeness to lunch time, etc.),[28] changing them every 30 minutes leads to more comparable results than if we do two hours shifts for example. Indeed, more changes happen in two hours than in 30 minutes.

These four combinations were indeed required to avoid any bias and thus flawed results. Combinations A+B+ and A-B- were needed in order to observe how people perceive

---

that administering vitamin A supplements to children reduced the mortality by 24% in low- and middle-income countries [Mayo-Wilson et al., 2011].

[28]It might be that for unknown reasons, these variables have a significant impact on donations. Further

differently messages A and B. This is equivalent to a reference group. Ideally, we wished to create two messages in such a way that people are indifferent choosing between A- and B-, and between A+ and B+. If we could be sure that this was the case, we would not need to test for A-B- or A+B+. However, as the messages are intrinsically different (i.e., in the sense that they refer to two different projects; one is about HIV and the other is about malnutrition), we cannot really be sure that two messages were truly "similar". It is conceivable that for whatever reason, people have intrinsic preferences for one message or the other, so that for instance, in both A-B- and A+B+ they prefer donating to A. Therefore, we need to include this comparison to understand how things actually work in this "baseline case".

In order to illustrate this, suppose that when combining A+B- we observe that 60% of people donate to A (that is, the opposite of what we would expect). One might conclude that the negativity bias does not actually hold. However, it may simply be that people like message A more than message B. Now, assuming that we check also the combination A+B+ and we see that in this case 80% of the people donate to A. In that case we could conclude that the negativity bias worked. Indeed, people preferred message A over message B to begin with (i.e., when A+ is combined with B+), but then if message B is framed negatively, some of them shift their preference towards message B. Message A is still preferred over message B on average, but there has been a significant shift in preference due to the negativity bias. Of course, the messages were written so that we believe that whenever faced with the combinations A+B+ and A-B-, people are indifferent between choosing A or B, so that people donations are close to 50%-50% but we would still need to verify this, otherwise our conclusions could be distorted.

With more time and resources at disposal, we could also pretest whether both types of messages are seen as equally appealing. For this, we could for instance implement a laboratory experiment or an online survey prior to the field experiment where we would

---

research is nevertheless needed to verify it.

ask respondents to evaluate the messages.[29] The questions would be structured as to see whether one message is preferred, or seen as more attractive when presented to the subjects. Nonetheless, for the ease of the implementation, we measured whether our messages are equally appealing directly via the field experiment.

2.3.2   Procedure

The pilot experiment went as follows. When people approached the table and appeared to read the posters, we told them *"We are students from the university collecting charity donations to be sent to support an aid project"*. We explained them that if they donate 1€ we would give them one KU Leuven pen in return. Moreover, we told them that for each donation we would match their donation with another euro coming from the university, so that 2€ would actually be sent to support the project. This allowed to increase total contributions since matching grants is the policy that has the largest positive impact on donors' behaviors, that is, the largest return on total donations to charities [Eckel and Grossman, 2008].

Once they decided to donate, we told them that they have to choose to which project they want the money to go. In all cases we avoided telling them beforehand that they were going to have to make a choice, to avoid a self-selection problem.

Furthermore, when someone decided to make a donation, we asked to have a look at the two messages concerning the two projects and to put the amount in his preferred project. It was always very important to pay attention to not indicate one box before the other one so to not influence their choice. Then we gave the person a pen. We decided to give the pen after he had chosen so as to keep his attention. Indeed, I believed that by doing so, the donor would pay more attention to the messages and take the pen as a reward, so that he would do his best to read the messages thoroughly.

---

[29]The group of respondents for the pretest would, of course, be different than the group of respondents for the field experiment in order to prevent any awareness of the second group when actually donating money.

From time to time, one experimenter went a bit further in the street to distribute flyers to passersby. The flyers were displaying the exact same message than on the posters, except that they were also displaying the date and the venue of the experiment. They were two-sided; one side in English the other in Dutch (See Appendix C). We did this to increase the curiosity of the people and to give them the opportunity to think before they were confronted to our stand. This approach considerably increased the number of people reading our posters and to some extent also increased the number of donors.

Finally, if a donor wondered why we made him choose between two projects rather than choosing ourselves, we avoided telling the real aim of the experiment, to avoid distortion of the donor's choice. For instance, we told him that we would like to support both projects but thought that leaving the choice to the donors was a more appropriate way of splitting the money.

Sampling

Ideally, on the one hand we wanted a representative sample of the population that allows us inferring some patterns that are common to the whole population. So the point is to have a representative sample that is large enough to make inference. On the other hand, there were some budget and time constraints that limited our sample size. The question therefore becomes "How large is enough?". Based on the existing literature and other examples we thus approximated the minimum number of subjects that would be needed to have a statistically representative sample and avoid the problems generated by the law of low numbers.[30]

A rule of thumb in psychological experimentation is that 20 participants in each condition (i.e., in each combination) is usually the lower bound limit. Nonetheless, as it is not a laboratory experiment but a field one, the number of respondents should be larger.

---

[30]The law of low numbers states that extreme or contradictory outcomes are more frequent in small samples than in large samples, because a large sample is less likely to stray from the expected outcome [Kahneman et al., 1982].

Indeed, since field experiments are more likely to be subject to noises[31] and biases due to the lack of control on the environment, the number of participants should be larger in order to compensate this noisiness with a more representative sample. I believe, therefore, that the literature offering similar field experiments (e.g., experiments on framing effect or negativity bias or, alternatively, any other experiments involving people in the street) would give a good approximation of the minimum number of observations required.

I am aware of four similar experiments which we can refer to determine the required number of participants.

First, Tversky and Kahneman [1981] collected data on the effects of variations in framing from groups ranging from 77 to 200 students who answered a brief questionnaire in a classroom setting.

Second, Das et al. [2008] obtained data on the effectiveness of eight different fundraising messages from passersby in the vicinity of a university campus of a large city. 160 of them agreed to participate.

Third, Hilbig [2009] conducted three experiments to demonstrate that negative instances are deemed more veridical. His first two experiments were conducted as online-surveys, while his third experiment was run using simple 1-page questionnaires dispersed within a community. With respect to the first experiment, 110 participants were recruited and randomly assigned to one condition amongst two. Concerning the second experiment, 38 participants were recruited and randomly assigned to one of two groups. With regard to the last experiment, 33 participants were recruited and randomly assigned to one condition among two.

Fourth, Moriarty [1975] conducted two field experiments on the responsiveness of 72 bystanders witnessing a theft on the basis of the level of prior commitment with the victim.

Based on these four illustrations and the rule of thumb, I personally set the number of

---

[31]Noise in this context is defined by the "**random** patterns that might easily be mistaken for signal", which is itself defined by "an indication of the underlying truth behind a statistical or predictive problem" [Silver, 2013, p. 416].

respondents that was needed for our field experiment to, on average, 25 donors in each of the four combinations, with strictly not less than 20 donors in any of them. This means that we expected the total number of donors to be 100.

Moreover, we conducted a one-day pilot prior to the full-scale experiment because it is a standard practice for field experiments and because it would be helpful to improve the design by detecting the small errors. By detecting them in advance we would thus be able to avoid them during the actual field experiment. More especially, we also thought that it would give us the opportunity to see how many people are, on average, willing to participate in our experiment, and we could therefore adapt it by increasing the required number of days to meet the minimum number of participants we previously set. Alternatively we could increase the minimum size of the sample if we found that many people were willing to participate in our experiment.

Experiment Type

One may wonder why we aim at observing real monetary donations involving cash transfers instead of hypothetical donations. We could have thought, for example, of a laboratory experiment where we would ask people in which project they would choose to donate if we would give them the money.

The first advantage of this method is that it would be easier to determine the exact number of respondents this being dependant on the number of subjects we would be willing to survey, instead of the number of people who would be willing to donate to charities. The number of respondents would in that case be independent of the number of people willing to give their own money. Such a number would be very hard to estimate prior to the implementation of the pilot.

The second advantage is that it would allow us to have a larger sample than if we apply the method detailed *supra* involving real monetary donations. Indeed, I believe that people

would be more willing to fill in a survey where they simply have to choose one of the two projects, rather than choose one project and actually give some of their own money. The conclusions drawn from a larger sample have the benefit of being more pertinent in as far as the data is more accurate and typically gives a better projection of the entire population.

Thirdly, laboratory experiments are usually conducted in a closed environment where all variables that are not being tested and that could influence the outcomes are controlled, leading to experimental results that are less noisy. Field experiments are in fact subject to factors that change the way respondents behave and that are much harder to keep constant across the whole process. Conducting a laboratory experiment where we would ask respondents to complete a questionnaire and choose between two projects in a classroom setting, would indeed induce a less noisy experiment.

Nonetheless, the study was designed as detailed *supra* in order to also potentially measure the magnitude of the negativity bias and the framing effect. If the experiment was limited to observing the choice of the individuals between two options, it would not be possible to collect information on the amount of money people would wish to donate. In contrast, if we give respondents the opportunity to choose between one of the two projects, and to state how much money they would be willing to donate for the chosen project, then the hypothetical amount would probably be biased upwards. This can be explained by the fact that people tend to overstate their hypothetical donations in comparison to the real amount of cash donations that is collected. Therefore, since hypothetical donations are typically between two to eight times larger than cash payments,[32] it would be literally impossible to correctly measure the magnitude of the negativity bias [Brown et al., 1996, Champ et al., 1997, Navrud and Veisten, 1996].

---

[32]This ratio increases to two to ten times if Contingent Valuation (CV) payments (valuation of non-market resources, for instance, the view of a mountain or the biodiversity) are compared with the actual cash equivalent [Seip and Strand, 1992, Duffield and Patterson, 1992, Navrud and Veisten, 1996, Champ et al., 1997].

Moreover, I believe that using real monetary donations leads to less biased results because people think much more and behave more as they would do in real life when there is real money at stake. Indeed, in reality, people are never making decisions concerning donations while sitting in a classroom, and knowing someone else is observing and testing them. This experiment structure would thus be the one which most closely reflect real life.

Furthermore, as respondents will donate their own money and that they are hence more likely to truly reveal their preferences, this incentive-compatible method would lead them to both truthfully report their donations, and choose their preferred project.

Another advantage is that a field experiment allows for a greater diversity of respondents. Indeed, if our experiment was designed as a laboratory experiment it would mainly attract students from the Katholieke Universiteit Leuven, as our main way of letting people know about the experiment would be through the university. The drawbacks of having a sample consisting mostly of students is that the findings may be flawed as students tend to be more aware of the negativity bias,[33] and that the sample is less representative of the general population. Also, the Charities Aid Foundation and National Council for Voluntary Organisations [2012] found that most of real life donations come from adults and elderly people: 67% of women and 62% of men between 45 and 64 years old gave in 2012 in the UK, while these percentages fall to 47% and 49% respectively for women and men between 16 and 24 years old [National Council for Voluntary Organisations, 2014]. So basically by selecting only students we would focus on the age category that are the least likely to donate in real life, thus compromising the relevance of our study.

By conducting a field experiment where we survey passersby in the city center of Leuven, we wish to attract men and women belonging to different socioeconomic classes and of different ages. This we hope will limit the number of individuals aware of the negativity bias. Moreover, a more diversified sample is more representative of the entire population

---

[33]This assumption is based on common sense. Most of the students in psychology and many students in economics are at least aware of this effect. Nonetheless, further research is needed to make sure that students, and more precisely, education affects the results of this kind of experiment.

and tends to capture the people that are more likely to donate. We will therefore be able
to draw stronger conclusions.

Last but by no means least, this structure of the experiment is truly simulating what
NGOs actually do. Indeed, they provide a public message and hope it will attract donations
in the real world, not in a lab environment.

2.4   Hypotheses

Three hypotheses arise:

HP1: Donations are significantly higher for the project that states the negatively framed
message rather than the project that states the positively framed message.

HP2: Donations are not significantly different when both projects display negative or
positive messages together.

HP3: Donations are not significantly different between A-B- and A+B+.

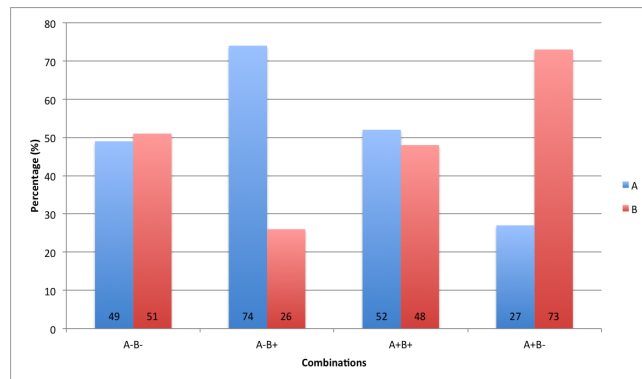The hypothetical results of the hypotheses we wanted to test are summarized in Figure
2.1.

Figure 2.1: Hypothetical percentage of donors choosing project A and project B for the
reference groups (A+B+ and A-B-) and the treatment groups (A-B+ and A+B-).

2.5 Experimental Results

As discussed above, we planned to conduct a pilot before the experiment. The pilot was intended to determine whether our methods and procedures were appropriate for the subject.

2.5.1 Summary Statistics

Tables 2.1 to 2.4 summarize the results of the donations during the pilot, by subject and by combination.

	Combination	
Subject	A-	B-
Subject 1	2	1
Subject 2	1	1
Subject 3	0	1
Subject 4	1	1
Subject 5	1	0
$\bar{x} =$	1	0.8

Table 2.1: Amount of donations by subject for the combination A-B-.

	Combination	
Subject	A-	B+
Subject 1	0	2
Subject 2	0	1
Subject 3	0	1
Subject 4	1	1
Subject 5	2	3
Subject 6	0	1
Subject 7	2	0
Subject 8	2	2
$\bar{x} =$	0.875	1.375

Table 2.2: Amount of donations by subject for the combination A-B+.

For an easier visualization of the results, I now present some key figures:

- 35 subjects donated

- the average donation was 1.6€

- 15 subjects donated more than 1€

- 3 subjects obviously did not read or choose

- 8 subjects donated to both projects

| | Combination | |
Subject	A+	B-
Subject 1	0	1
Subject 2	2	0
Subject 3	0	1
Subject 4	1	0
Subject 5	0	1
Subject 6	0	1
Subject 7	1	0
Subject 8	0	1
Subject 9	0	1
Subject 10	1	0
Subject 11	2	0
Subject 12	0	1
Subject 13	1	1
Subject 14	1	0
Subject 15	0	2
Subject 16	1	0
Subject 17	0	1
$\bar{x} =$	0.59	0.65

Table 2.3: Amount of donations by subject for the combination A+B-.

| | Combination | |
Subject	A+	B+
Subject 1	0	1
Subject 2	0	2
Subject 3	1	1
Subject 4	0	2
Subject 5	0	1
$\bar{x} =$	0.2	1.4

Table 2.4: Amount of donations by subject for the combination A+B+.

2.5.2 Issues Encountered

During the implementation of the pilot we rapidly encountered some issues related to the design of the experiment. The conditions and the environment were not the source of the problem as there were enough people passing by, enough people donating, etc. However, we rapidly discovered a few issues that were critical for the experiment. I highlight eight.

First, couples tended to give to both projects. Indeed, if one partner donated to a project the other partner wanted to give to the other project independently of the messages they were facing.

Second, single elderly people tended to give to both projects in order to be altruistic or because they did not want to favor one project over the other. This was possible because people were not limited in their choices of how much to donate, so that they could make

two donations or more.

Third, if someone saw which project the previous donor supported, he was most likely to choose the other project, even though the two donors did not know each other.

Fourth, some asked us which project was the least successful.

Fifth, most respondents that gave more than 1€ did so because they did not have 1€ coins.

Sixth, some asked us to choose for them.

Seventh, many subjects were asking their friends whether they should give to the project that supports HIV or the project that supports malnutrition without their friends looking at the messages, cancelling out the potential effect of the negativity bias and framing effect. This specific problem illustrates that some people chose only because of the content and were not affected by the framing.

More importantly, we felt that most donors were not reading the messages thoroughly. Indeed, a lot of participants wanted to give 1€ and then quickly go back to their activities. These behaviors were repeatedly observed during the pilot.

Although these issues were problematic for our experiment, they can also be seen as fascinating behaviors in other studies or domains. Indeed, researchers sometimes observe unexpected results and discover something even more interesting for future research. More-over, many inventions are the result of mistakes or unexpected outcomes. This is, however, beyond the scope of this thesis.

It is true that we sometimes encountered people that seemed to have read the messages. However, we were first concerned that this fraction of people was too small to be repre-sentative. Second, we were never completely sure about this belief because the boundary - between people that read and people that do not - was blurred (i.e., not rigorously de-limited). Otherwise, we would have taken into account only the choices of these donors. But in this case, I personally believe that I would not have been always objective. Indeed, although I read many academic papers and spent a lot of time on the negativity bias, I

found almost no counter effects or counter examples in the literature. Therefore, I think that I may unconsciously overestimate the effects. So if I had to determine whether a donor has or has not read the messages carefully, I would have probably overestimated the number of times that he has read them carefully if he had chosen to donate to the negatively framed message, and inversely, I would have probably overestimated the number of times that he has not read them carefully if he had chosen the positively framed message. For this reason, we decided to collect all responses without making a difference between people that seemed to have read the messages carefully and those that did not.

Therefore, after having considered every options, we decided to stop after the pilot and not to implement the full-scale experiment.

Actually, the problems we mentioned above would not be per se a problem. Suppose many people do not read the messages and are subject to all the problems and limitations we mention above. There is then in most of the cases no reason to believe that they would be **systematically** biased in their choices and this would not therefore invalidate the experiment but simply add noise to the measure.

What I mean is that if people donate without caring about the messages, they will mostly choose randomly. This means that on average we would have many people choosing the positive option and many choosing the negative option. Nonetheless, as long as some people read the message and the negativity effect holds, with a sample that is large enough we should still see that at the end the negative messages receive relatively more. Where the actual problem lies is in the fact that this large noise in the data calls for a much larger sample than what we initially forecasted, requiring more time and more resources for performing a meaningful analysis. This is indeed the real problem.

One exception to this is one specific problem we mention above, that is that some people donate the opposite of what the previous person has donated to compensate for that. This is indeed a setting in which the donation would not be *random* (suppose the previous person reads the message and gives to the negative message, then the next person

would always give to the positive one). So this is the only real problematic case, for which an alternative would be needed. For instance, one solution could be to have two people working at the stand at the same time and using two different tables. In the first table we "welcome" donors, while in the second table they have to make the choice, so that people do not really see or focus on what the previous person is doing, but can talk with some of us, while the other makes the choice.

The remainder of this chapter is organized as follows. First I will still present the analysis and interpretation of the experimental results of the pilot. In this subsection we will focus on the choices of donors (either A or B) for each combination rather than on the amount of donation because most subjects that gave more than 1€ did so because they did not have 1€ coins. In the next subsection I will perform an ANalysis Of VAriance (ANOVA) to measure the group means.[34] Although we usually need a larger sample, I will still perform this empirical analysis in order to show the methodology and to show how I would have done, had the data set been larger and more reliable. Third, I will describe the possible explanations for those mixed results. I will then present some potential solutions and improvements for future research together with their limitations. Finally, a discussion and some implications in the real world will conclude this chapter.

2.5.3 Data Analysis and Interpretation

As described in the previous section, to remain objective we have not identified whether the respondents have carefully and thoroughly read the messages before making the donations. However, we still noted when a donor obviously did not look at the messages. This happened for example when the donor left when asked to choose or when he picked a project because the previous donor selected the other project. Moreover, we also noted

---

[34]Although the methodology compares means, it is called analysis of variance because it actually tests for statistical significance between means by analyzing variances [StatSoft, Inc., 2013].

when a donor gave to both projects at the same time.

This allows us to compare four types of results: when

 (i) all donations are taken into account (i.e., including donors that obviously did not read or choose and including donors that gave to both projects),

 (ii) including donors that obviously did not read or choose but excluding donors that gave to both projects,

(iii) excluding donors that obviously did not read or choose but including donors that gave to both projects, and

(iv) excluding donors that obviously did not read or choose and, also excluding donors that gave to both projects.

The first category (i) can be seen as the most noisy one since it includes the two main issues while the last category (iv) can be seen as the purest one since it excludes these issues. Presenting the four categories would be redundant, so we are only going to present the two most interesting ones: the first category because it is the most representative cohort, and the fourth category because it is the most unbiased cohort. It is important to note that we should be careful with the interpretation that follows because of the limited sample. The interpretation is presented to show how we would have interpreted the experimental results, had the sample size been large enough.

Figure 2.2 illustrates the results including all donations. Figure 2.3 illustrates the results excluding donors that obviously did not read or choose and excluding donors that donated to both projects. In both graphs, N denotes the global number of choices made by participants for each of the combinations.[35]

---

[35]The disparities in the sample size between the first three combinations and the fourth one can be explained by the fact that as soon as one or two potential donors were considering to donate, other passersby were in turn much more likely to donate while if no one was considering to donate, passersby were inclined
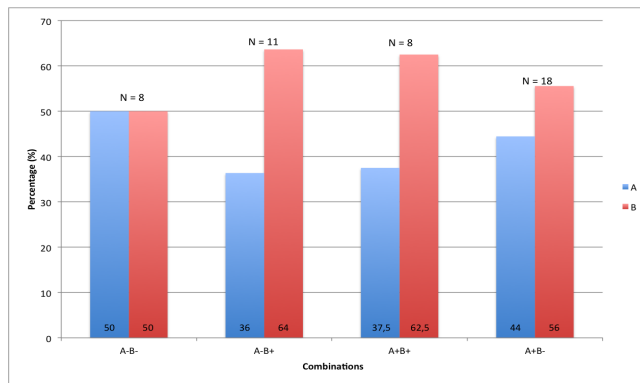
Figure 2.2: Results (i): Percentage of choices for project A and project B, all donations taken into account.
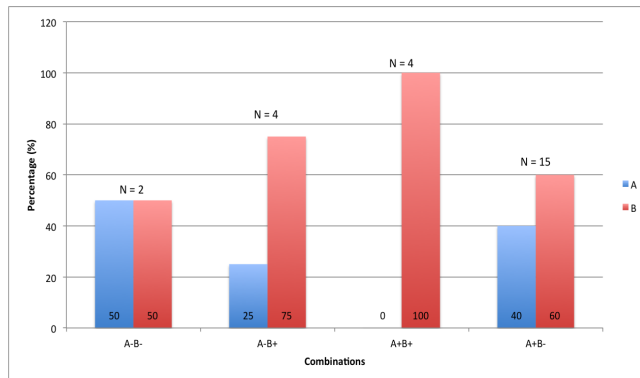

Figure 2.3: Results (iv): Percentage of choices for project A and project B, excluding donors that obviously did not read or choose and excluding subjects that donated to both projects.

As we can see on the graphs, the results of the pilot are quite inconsistent with our hypotheses mentioned in Section 2.4. Regarding figure 2.2, we see that there are two combinations that are in line with our hypotheses; A-B- and A+B-. First, people were indifferent between A and B when the combination A-B- was presented to them, which is in line with HP2. This seems to show that we managed to find two types of message that

to avoid us. This issue will also be highlighted *infra* in Section 2.6.

are equally attractive. Second, a larger proportion of people donated to the negatively framed project in the combination A+B-, which is in line with HP1. However, when we compare those two combinations with the two other combinations A-B+ and A+B+, we see that the results are not corroborating our hypotheses for three main reasons.

First, the combination A+B+ shows that people seemed to prefer message B, which is inconsistent with HP2. It might be that some behaviors were influenced by the fact that message B concerns children. For instance, women, mother or simply parents may have been affected by the content such that the framing was too weak to overcome the innate parenting skills of this category of subjects. As explained in Section 2.3.1, this is not a problem per se. Indeed, if the difference between A and B for the combination A+B- was larger than the difference in the combination A+B+ (e.g., 20% for A and 80% for B for the combination A+B- and the same percentages as illustrated in figure 2.2 for the combination A+B+) we could have said that the negativity bias holds. However, in our case a larger percentage of people selected the positive option when they were faced with the positive and negative messages than when both messages were positive, as if there was a *positivity* bias, which is of course inconsistent with HP1.[36]

Second, if we compare the combination A-B- with the combination A-B+, we see that 64% of respondents supported the project with the positive message when faced with a positive and a negative message, while they were indifferent between A and B when both messages were negative (i.e., 50% for A and 50% for B). This is not in line with HP1 since donations are higher for the positive option, but this is in line with HP2 because donations between the two negative options are equal.

Third, the comparison between the combinations A-B+ and A+B+ does not give better results. Indeed, approximately the same percentage of donors selected the positively framed option when they were provided with a negative and a positive option as when they were

---

[36]Sears and Whitney [1973] and Sears [1976] firstly describe the positivity bias as the "consistent favorable evaluation of specific individuals". However, in this context I use the term as the opposite of the negativity bias, that is, people are more affected by positive than negative instances.

facing two positive options. This is also inconsistent with HP1, as less people would have selected the positive option in the first combination (A-B+) compared to the second combination (A+B+) if the results were in line with the hypothesis.

The only comparison that supports these effects is the comparison between the combinations A-B- and A+B-. Indeed, as soon as one of the two messages was negatively framed, some donors shifted their preference to this option. This is in line with HP1 and HP2.

Figure 2.3 illustrates two comparisons that support our hypotheses and two that do not.

First, although the percentages are not the same the comparison between A+B+ and A+B- gives the same interpretation than for figure 2.2, that is, the results support neither HP1 nor HP2. Indeed, all donations went for a particular option when both messages were positive, so HP2 is violated. And some donors shifted their preferences to the positive option when they were faced with the positive and negative option, which is conflicting with HP1.

Second, the comparison between A-B- and A-B+ also shows opposite effects since a smaller percentage of donors chose the project with the negative message although they were indifferent when both messages were negative, which demonstrates consistency with HP2 but inconsistency with HP1.

Concerning the two comparisons that support the negativity bias hypothesis, the first one is the comparison between A-B+ and A+B+ and the second one is the comparison between A-B- and A+B-.

First, when both messages were positive all donors preferred message B but one quarter shifted their preferences to the negative message when the messages were oppositely framed, which is in line with HP1.

Second, we see that 10% more people supported the project with the negative message than when both projects were displaying a negative message, which is also consistent with

HP1.

To sum up the results of the pilot, among the four comparisons when all donations are taken into account three of them illustrate a violation of the negativity bias. Furthermore, only half of the comparisons support our hypotheses when the cleanest results are interpreted.

It is also worth to note that there is a huge difference between A-B- and A+B+, which is inconsistent with HP3. This finding is puzzling because in theory, if people perceive A-B- as similar, not much should change when faced with A+B+. This finding is quite interesting as it comes out striking from both figures. It might certainly be partly due to the noise in the data, but it might also suggest that when framed positively or negatively the salience of the messages changes significantly. For example, it may be that as long as the messages are both negative, people do not give weight to whether they concern health or hunger but they rather focus on the framing. Suggesting that the negative framing overcomes the content, people are indifferent because both messages are negative, so both have the same potential "strength". While whenever framed positively, they pay more attention to the fact that they concern health or hunger. Suggesting that the positive framing is too weak to dominate the content, people in this case choose their preferred project depending on their preferences between health and hunger and not depending on the potential "strength" implied by the framing. This might be an interesting element coming out from this pilot that we were not planning to focus on but which might deserve further exploration in the future.

2.5.4   ANOVA (ANalysis Of VAriance)

The purpose of an analysis of variance is "to test for significant differences between means" [StatSoft, Inc., 2013]. As said earlier, a larger sample is usually needed in order to perform

an ANOVA. However, we still present the results of the analysis to show its methodology.[37]

The results we obtain with our dataset look probably bad according to specialists, but their "ugliness" is justified by the problems we highlighted before.

The ANOVA command is the following:[38]

```
fit <- aov(var ~ categ, fr)
```

The order of arguments is important. The first argument `var` indicates the dependent variable Y (the output; the amount of the donation). The second argument `categ` denotes the independent variable X (the input; the choice of the donor). The final argument `fr` is the name of the data structure that is being analyzed and under which we want the analysis to store [Revelle, 2013].

The results can be seen with the summary command:

```
summary(fit)
```

This produces the following output:

```
> summary(fit)
            Df Sum Sq Mean Sq F value Pr(>F)
categ        7   7.65  1.0929   2.148 0.0514 .
Residuals   62  31.55  0.5089
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

The null hypothesis of ANOVA is that the means are all equal: $H_0 : \mu_1 = \mu_2 = ... = \mu_k$, where k is the number of conditions.[39] In our experiment, $k = 8$. The alternative hypothesis is that at least one of the means is different from at least another mean [Akey,

---

[37]For this analysis, the RStudio program was used. RStudio is a professional software for the R statistical computing environment.

[38]Another method which gives the same results is the following command: `anova(lm(var ~ categ, fr))`.

[39]This non-specific null hypothesis is sometimes called the *omnibus null hypothesis* [Lane, n.d.].

2008]. Since the p-value is larger than 0.05^{40} we are unable to reject the null hypothesis that means are all equal. Thus, if our data set had been nicer we would have concluded that there is no credible evidence that the negativity bias or the framing effect held since the means are not significantly different from each other. This implies that people were indifferent with respect to the messages. Nonetheless, we have to be careful with this conclusion because the p-value is only slightly greater than 0.05 (p-value = 0.0514). Indeed, we can not reject the null hypothesis but this non-rejection is not so obvious.

Furthermore, if the ANOVA had shown significant differences between the means, we would have performed a Tukey Honestly Significant Difference (Tukey HSD) test in order to know between which factors there is a difference.[41] Indeed, the ANOVA reveals whether the means are different or not but it does not reveal which means are different from which. The Tukey HSD test offers this specific information [Lane, n.d.]. This test is usually a follow-up of ANOVA and it can be performed only if the ANOVA shows a significant difference [Duclert, 2013]. Although the ANOVA did not show a significant difference between the means (p-value = 0.0514), we still present the Tukey HSD test because the p-value is only slightly not rejected and to show the methodology [Duclert, 2013]:[42]

```
tes <- aov(var ~ categ, fr); TukeyHSD(tes)
```

The p-values show no significant difference between the means (See Appendix D for the output of the Tukey HSD test and Appendix E for the whole set of commands in the program R together with the output). With our small sample size, it is not too surprising because results from small samples are unstable [Lane, n.d.].

---

[40]This value is the default choice for the significance level of $\alpha = 5\%$. This level gives the type I error (or error of the first kind), i.e., the probability of rejecting H_0 when it holds.

[41]This test can be performed provided that there are more or less the same amount of observations in each category. This is again not the case with our dataset but we still provide it to show the methodology.

[42]The default level of confidence for this test is 0.95. It can be changed thanks to the following command: `TukeyHSD(tes, conf.level = 0.90)`, where 0.90 corresponds to the confidence level.

2.5.5 Possible Explanations

When it comes to the source of our mixed results, two possible explanations can be considered: (i) Either our assumptions and hypotheses are wrong and/or do not hold in the domain of charity donations, or (ii) the structure and the design of the experiment are inappropriate to measure the effects of the negativity bias and the framing effect.

Despite thorough research in the literature and several discussions with specialists, there is no research I am aware of that supports the opposite of what we expected, that is, research arguing that there may be a psychological phenomenon that people tend to be more sensitive to positive than neutral or comparably negative instances. Furthermore, there is also no research or experiment I am aware of yet that shows no significant effect. Otherwise, this thesis would corroborate these findings. The fact that I was not able to find any example of research that supports the null hypothesis in the literature may be due to the *publication* bias. This bias is the tendency of editors and researchers to under-report results that are negative (i.e., supporting the null hypothesis) compared to results that are positive (i.e., showing a significant finding) mostly due to the experimenter's loss of interest in the topic since he anticipates that others will not be interested in null results [Song et al., 2010, Easterbrook et al., 1991].

Indeed, despite the fact that studies with a null result do not appear to be inferior to studies with significant results with respect to the quality of design, it has been found that papers affirming a null result are three times less likely to be published than statistically significant results [Dickersin et al., 1987]. Therefore, the negativity bias could possibly be overrated.

In any case, based on my experience of the pilot and my own judgment, I believe that the structure and the design of the experiment are more likely to be inappropriate to measure the expected effects. I will thus present the potential solutions and improvements that would be advisable to follow if a similar experiment had to be performed in the future,

as well as their limitations.

2.6   Potential Solutions and Improvements

Due to the problems experienced during the pilot, the structure and the design of the experiment would need to be drastically improved or modified if we want to test these effects again in later research.

A first potential improvement is to set up the following approach. Everything would stay the same except for the boxes. Instead of asking people to put their money in one of the two boxes, we could make people choose more actively by forcing them to answer some questions beforehand. These questions could be raised through a computer for example. At the end of this small questionnaire, they would have to select the project they would like to support (still via the computer so that their attention does not fade away). This approach is expected to solve the problems due to lack of attention. Nonetheless, it would probably not seem very natural for charity donations. Moreover, it could potentially add new problems such that many respondents would be reluctant to answer the questions or that they would withdraw themselves from the process as soon as we would tell them that they have to handle some questions before making their donation. This fear is well-founded knowing how reluctant donors were when we simply asked them to put their money into one of the two projects. Furthermore, we would have to make sure that the questionnaire does not annoy or bother people such that they do not want to take the time to read the two messages carefully anymore. The questionnaire should therefore be long enough to capture their attention without being too long so that they do not get bored at the end of it. Otherwise, this would bring us back to the initial problem of people not entirely focusing on the messages.

Another approach which requires drastic changes to the present structure of the experiment could also be the following. Instead of displaying two messages we could expose

participants to one message at a time and see what are the impacts on donations. For instance, we would not attract people by inviting them to donate but we would just wait and see how attractive the message is. We would thus display a large banner displaying in a very clear and visible way one message at a time, and we would interchange it with its opposite version (i.e., negative versus positive) every hour or so. This approach would allow us to measure the impacts of the negativity bias either via the ratio between the number of donors and the number of passersby depending on the type of message, that is, we would only look whether more people stop attracted by the negative message. Or it could also be measured via the amount of money collected by donor and also depending on the type of message, that is, we would only measure how much they donate because perhaps they donate more when the message is negative.

These two approaches would probably eliminate the problems due to lack of attention, as people would donate only if they read the message (otherwise they would not know that we are collecting money). Moreover, they would obviously solve the problem of compensation across donors as only one message would be displayed. However, two important problems would certainly arise regarding those two approaches. First, during the pilot we figured out that if some people were standing near our stand, reading the messages or were actually giving money, more passersby were in turn inclined to stop by and to donate. Whereas if no one was considering to donate, people tended to avoid us and keep walking. This problem would obviously bias the ratio between the number of donors and the number of passersby. Second, those two approaches would also add a problem due to the environment. It might be that the weather influences people when it comes to donations. For instance, people may give a higher amount of money or may be more likely to participate in a fundraising campaign when the weather is nice rather than when it is raining. This is clear that those two main issues would bias the results if not controlled.

One could hope the effects of the negativity bias to be so strong that it would overcome the potential noises of this method of one message at a time. However, because these last

issues would be hard to control, we thought that a second pilot with these approaches was too risky and complex given the resources and time we have. Indeed, we believed that there would have been too many unstable factors to draw sound conclusions.

2.7   Discussion

As we have seen through the present thesis, many pieces of evidence of the negativity bias and the framing effect have been demonstrated in the previous literature. Unfortunately, we were not able to draw conclusions from our findings since the design of the experiment needs to be reviewed. However, a second field experiment with a better structure should allow to observe whether the negativity bias occurs in the domain of charity donations and if it does, to what extent.

If we were able to find a design that is appropriate to measure the negativity bias, two outcomes would arise. Either (i) the real observations are in line with the hypotheses. In this case it would mean that the negativity bias exists and that it has a significant impact on individuals' behaviors. Or (ii) the real observations are not in line with the hypotheses. In this case it would probably be one of the first study to show that the negativity bias does not hold in charity donations.

A possible reason why charity donations is a field where negativity bias could potentially not work is that it engages real money. The negativity bias and framing effect have been studied so far mostly in hypothetical choices, or choices without any real consequence for the subject. Therefore, it may be that those effects exist only when the subjects are *allowed* to choose instinctively, without thinking or balancing the pros and cons too much. What is indeed peculiar about the charity domain is that it involves choices that will have a real impact on donors. In hindsight, it may seem an insignificant impact for most of us but it may also be that people weight potential consequences too heavily given their proportions. As Wilson and Gilbert [2003] have shown, people overestimate the impact of future events

on their lives. This overestimation could justify the fact that people think in a different way (e.g., more thoroughly or simply focusing on something else) when they have to make a donation. This different thinking may in turn justifies the nonexistence of negativity bias and framing effect in this domain.

Implications in the real world are multiple if we consider that the negativity bias and framing effect hold for charity donations. On the one hand, the results would enable to draw conclusions on the type of messages that charity institutions should use to attract as many donors as possible. Indeed, it would give them insightful details on how to write their key message or slogan. These (almost) costless improvements would help them to differentiate from other institutions in a very competitive market and help them to reach their fundraising goals more easily. Charity institutions that understand and apply these issues in their fundraising campaign are more likely to achieve their goals.

On the other hand, apart from being used by NGOs, I think that the findings can shed light on some behaviors we already observe in the real world. Many people indeed often wonder why many messages we see on television, radio and newspapers are framed negatively, it could well be (and certainly is the case of large NGOs with expert marketing managers) that organizations are themselves already aware of this behavioral pattern and are consciously using it, while donors are not aware.

To a broader extent, the results will also demonstrate whether charitable giving are merely the consequence of altruism or not. If the assumptions mentioned above actually hold, it means that donations to charity are not only ruled by altruism. Indeed, together with altruism, things such as economic laws, peer pressure, direct environment and other factors that still need to be discovered play a role.

Conclusion

This work began few months ago with a literature study that unveiled some of the research that has already been done in different domains that are related to the subject of this Master's thesis. The specific domains that are mentioned are negativity bias, framing effect, charity donations, impact of message framing on effectiveness, positive-negative asymmetry, framing of decisions, field experiment, behavioral economics, realism.

This thesis attempted to shed light on the role and impact of negativity bias and framing effect and thereby contributes to the relevant policy discourse. Indeed, the academic literature has shown that the negativity bias and framing effect have a significant impact on people's choices and preferences. This thesis was intended to apply these concepts to charity donations through a field experiment involving real monetary donations.

After analyzing the results of the pilot, we decided to not perform the planned full-scale experiment. Although this decision was hard to conceive at first, it was probably the best one given the time and resources at disposal. Indeed, given the issues we encountered during the pilot, a much larger sample than we initially planned to have would have been needed to compensate the noise.

This decision must in any case not been seen as a failure. We were not able to highlight why there are inconsistencies in the messages coming from organizations as well as from the widespread observation that charity organizations and fundraising agencies often rely on catch-phrases and pictures that present a gloomy future if action is not taken (and donation not made), rather than a positive image if action is taken. However, this thesis

meant to give an actual contribution (whatever small it might be) to the research agenda of the Katholieke Universiteit Leuven. The idea was to add to the knowledge of the research community why we observe this phenomenon. Is it because people are more attracted by negative messages? Is it because it is more efficient in attracting donations? Beyond these answers, we were aiming at finding out an approach to test it. We tried an approach that probably did not worked perfectly for the way it was designed. I think there is definitively room for improvement and that it could be worth keep pushing for this experiment if we manage to solve the problems we discussed. Insofar as it had worked perfectly, this would have been a clear answer to these important questions. As it is now, it remains an open debate which I suppose happened many times to every researcher. More research will therefore be needed to isolate the impacts of those effects in real monetary donations.

Beyond this fact, we have learnt a lot through this work. First, the structure of a field experiment is essential in research. We have seen through this thesis that the design of the experiment has to capture all particularities of the real life in order to possibly demonstrate some effects. Second, thanks to the proliferation of research on the subject, we have learnt that the basic assumptions of rationality and consistency in economic theory may be mitigated in some domains, namely, choices, decisions, preferences, and charity donations. Furthermore, this thesis was important in a sense that it described some behavioral patterns that are common among us, while we are not always aware of them. The objectives of this thesis were indeed to vulgarize the concepts and to demonstrate them through a field experiment. Towards that end, it will always be meaningful for individuals that are curious to know what may unconsciously influence our decisions and choices. People that have learnt these concepts will thus now be able to detect when someone or something is playing with their "predictable irrationality" [Ariely, 2008]. Whether we decide ourselves to be consciously led by this or to counteract it must remain a choice for each of us. At least, people aware of these patterns may choose more objectively and more consciously.

I have learnt a lot from a personal point of view too. A Master's thesis was the main opportunity to illustrate what I have acquired during my bachelor and master programs in economics. Of course, I will never fit into the thesis everything I learned. The technical knowledge that I actually used in the thesis is rather likely to be very small, but it is the overall package that matters: how I learnt to moved through the literature (which means whether I managed to find relevant papers, whether I understood the patterns through different works, and how I eventually present it), how I learnt to reason before, during and after the implementation, and how I present and discuss the results in an "academic way". This thesis can also be considered as a way of presenting myself for future work, as it shows how I faced challenges in the implementation of the project, how I managed to deal with deadlines, etc. These all give a very important signal about my knowledge, ability, reasoning, and commitment.

Finally, the implications of this thesis immediately suggest a host of promising avenues for further investigation. I highlight three of them. First, due to the fact that people think more thoroughly when making a choice involving real monetary donations, these fields may rely on theories that are slightly different from what have been studied so far. Second, the behavior patterns we observed during the pilot could potentially be the subject of other research. Third, the fact that the comparison between A-B- and A+B+ is puzzling deserves further research because it could result in questions we have not really thought to ask in the first place, referred by Rumsfeld [2012] as "unknown unknowns".[43] The interpretation of A+B+ and A-B- may have been an unknown unknown because we have not really thought to ask questions about it, may be a known unknown today because we asked ourselves a question but could not come up with an answer, and may become a known known if further research is undertaken in this domain [Silver, 2013, p. 420].

---

[43]Rumsfeld [2012] defines things we know we know by *known knowns*, things we know we do not know by *known unknowns*, and things we do not know we do not know by *unknown unknowns*.

Bibliography

Joshua Akey. Lecture 7: Hypothesis testing and ANOVA, April 2008. URL `http://www.gs.washington.edu/academics/courses/akey/56008/lecture/lecture7.pdf`. Accessed April 29, 2014.

Alberto Alesina, Rafael Di Tella, and Robert MacCulloch. Inequality and happiness: are europeans and americans different? *Journal of Public Economics*, 88(9):2009–2042, 2004.

Mike Allen and Raymond W. Preiss. Comparing the persuasiveness of narrative and statistical evidence using meta-analysis. *Communication Research Reports*, 14(2):125–131, 1997.

Enriqueta Aragones. Negativity effect and the emergence of ideologies. *Journal of Theoretical Politics*, 9(2):189–210, 1997.

Dan Ariely. *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. HarperCollins, 2008.

James R. Averill. On the paucity of positive emotions. In *Assessment and modification of emotional behavior*, pages 7–45. Springer, 1980.

Maya Bar-Hillel. On the subjective probability of compound events. *Organizational behavior and human performance*, 9(3):396–406, 1973.

Roy F. Baumeister, Ellen Bratslavsky, Catrin Finkenauer, and Kathleen D. Vohs. Bad is stronger than good. *Review of general psychology*, 5(4):323–370, 2001.

Behavioural Insights Team. Applying behavioural insights to charitable giving. May 2013. URL https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/203286/BIT_Charitable_Giving_Paper.pdf. Accessed November 22, 2013.

Neeli Bendapudi, Surendra N. Singh, and Venkat Bendapudi. Enhancing helping behavior: An integrative framework for promotion planning. *The Journal of Marketing*, 60:33–49, 1996.

Baler Bilgin. Losses loom more likely than gains: Propensity to imagine losses increases their subjective probability. *Organizational Behavior and Human Decision Processes*, 118(2):203–215, 2012.

Sarah Borgloh, Astrid Dannenberg, and Bodo Aretz. Small is beautiful–experimental evidence of donors' preferences for charities. *Economics Letters*, 120(2):242–244, 2013.

Jordi Brandts and Christiane Schwieren. *Frames and games.* Springer, 2007.

Judson S. Brown. Gradients of approach and avoidance responses and their relation to level of motivation. *Journal of comparative and physiological psychology*, 41(6):450, 1948.

Thomas C. Brown, Patricia A. Champ, Richard C. Bishop, and Daniel W. McCollum. Which response format reveals the truth about donations to a public good? *Land Economics*, pages 152–166, 1996.

Christopher D. B. Burt and Kenneth T. Strongman. Use of images in charity advertising: Improving donations and compliance rates. *International Journal of Organisational Behaviour*, 8(8):571–580, 2005.

John T. Cacioppo and Gary G. Berntson. Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological bulletin*, 115(3):401, 1994.

John T. Cacioppo, Wendi L. Gardner, and Gary G. Berntson. Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, 1(1):3–25, 1997.

John T. Cacioppo, Wendi L. Gardner, and Gary G. Berntson. The affect system has parallel and integrative processing components: Form follows function. *Journal of personality and Social Psychology*, 76(5):839, 1999.

Colin F. Camerer, George Loewenstein, and Matthew Rabin. *Advances in behavioral economics*. Princeton University Press, 2011.

Earl R. Carlson. The affective tone of psychology. *The Journal of General Psychology*, 75 (1):65–78, 1966.

Ziv Carmon and Dan Ariely. Focusing on the forgone: How value can appear so different to buyers and sellers. *Journal of Consumer Research*, 27(3):360–370, 2000.

Patricia A. Champ, Richard C. Bishop, Thomas C. Brown, and Daniel W. McCollum. Using donation mechanisms to value nonuse benefits from public goods. *Journal of environmental economics and management*, 33(2):151–162, 1997.

Chun-Than Chang and Yu-Kang Lee. Effects of message framing, vividness congruency and statistical framing on responses to charity advertising. *International Journal of Advertising*, 29(2):195–220, 2010.

Charities Aid Foundation. World giving index 2012, December 2012. URL `http://www.cafonline.org/PDF/WorldGivingIndex2012\\WEB.pdf`. Accessed November 27, 2013.

Charities Aid Foundation and National Council for Voluntary Organisations. UK giving 2012, November 2012. URL `https://www.cafonline.org/PDF/UKGiving2012Full.pdf`. Accessed May 1, 2014.

Eileen Y. Chou and Keith J. Murnighan. Life or death decisions: Framing the call for help. *PloS one*, 8(3), 2013.

Ronald H. Coase. The problem of social cost. *JL & econ.*, 3:1, 1960.

Charles Darwin. On the origins of species by means of natural selection. *London: Murray*, 1859.

Enny Das, Peter Kerkhof, and Joyce Kuiper. Improving the effectiveness of fundraising messages: The impact of charity goal attainment, message framing, and evidence on persuasion. *Journal of Applied Communication Research*, 36(2):161–175, 2008.

Pierre Desmet and Fred M. Feinberg. Ask and ye shall receive: The effect of the appeals scale on consumers' donation behavior. *Journal of Economic Psychology*, 24(3):349–376, 2003.

Kay Dickersin, S. S. Chan, T. C. Chalmersx, H. S. Sacks, and H. Smith Jr. Publication bias and clinical trials. *Controlled clinical trials*, 8(4):343–353, 1987.

James N. Druckman. Evaluating framing effects. *Journal of Economic Psychology*, 22(1): 91–101, 2001.

Aymeric Duclert. Aide-mémoire sur R, ANOVA, December 2013. URL `http://www.duclert.org/Aide-memoire-R/Statistiques/ANOVA.php`. Accessed April 29, 2014.

John W. Duffield and David A. Patterson. Field testing existence values: comparison of hypothetical and cash transaction values. *Benefits and Costs in Natural Resource Planning, Oregon State University*, 1992.

Elizabeth W. Dunn, Lara B. Aknin, and Michael I. Norton. Spending money on others promotes happiness. *Science*, 319(5870):1687–1688, 2008.

Phillipa J. Easterbrook, R. Gopalan, J. A. Berlin, and David R. Matthews. Publication bias in clinical research. *The Lancet*, 337(8746):867–872, 1991.

Catherine C. Eckel and Philip J. Grossman. Subsidizing charitable contributions: a natural

field experiment comparing matching and rebate subsidies. *Experimental Economics*, 11 (3):234–252, 2008.

Seymour Epstein. Toward a unified theory of anxiety. *Progress in experimental personality research*, 4:1–89, 1967.

Catrin Finkenauer and Bernard Rimé. Socially shared emotional experiences vs. emotional experiences kept secret: Differential characteristics and consequences. *Journal of Social and Clinical Psychology*, 17(3):295–318, 1998.

Daniel T. Gilbert, Elizabeth C. Pinel, Timothy D. Wilson, Stephen J. Blumberg, and Thalia P. Wheatley. Immune neglect: a source of durability bias in affective forecasting. *Journal of personality and social psychology*, 75(3):617, 1998.

John Gottman. *Why marriages succeed or fail and how you can make yours last.* Simon & Schuster, New York, 1995.

Bent Greve. What is welfare? *Central European Journal of Public Policy*, (1):50–73, 2008.

Carola Gruen and Stephan Klasen. Growth, inequality, and welfare: comparisons across space and time. *Oxford Economic Papers*, 60(2):212–236, 2008.

Benjamin E. Hilbig. Sad, thus true: Negativity bias in judgments of truth. *Journal of Experimental Social Psychology*, 45(4):983–986, 2009.

Carroll E. Izard. The face of emotion. 1971.

Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, pages 263–291, 1979.

Daniel Kahneman and Amos Tversky. Choices, values, and frames. *American psychologist*, 39(4):341, 1984.

Daniel Kahneman, Paul Slovic, and Amos Tversky. *Judgment under uncertainty: Heuristics and biases.* Cambridge University Press, 1982.

Daniel Kahneman, Jack L. Knetsch, and Richard H. Thaler. Experimental tests of the endowment effect and the coase theorem. *Journal of political Economy*, pages 1325–1348, 1990.

Daniel Kahneman, Jack L. Knetsch, and Richard H. Thaler. Anomalies: The endowment effect, loss aversion, and status quo bias. *The journal of economic perspectives*, 5(1): 193–206, 1991.

David E. Kanouse and L. Reid Hanson Jr. Negativity in evaluations. In *Preparation of this paper grew out of a workshop on attribution theory held at University of California, Los Angeles, Aug 1969.* Lawrence Erlbaum Associates, Inc, 1987.

David M. Lane. Online statistics education: An interactive multimedia course of study, n.d. URL `http://onlinestatbook.com/`. Accessed April 29, 2014.

Jeff T. Larsen. Negativity bias. pages 1–4.

Bibb Latane. Field studies of altruistic compliance. *Representative Research in Social Psychology*, 1970.

John A. List. Introduction to field experiments in economics with applications to the economics of charity. *Experimental Economics*, 11(3):203–212, 2008.

John A. List. The market for charitable giving. *The Journal of Economic Perspectives*, 25 (2):157–180, 2011.

Douglas C. Macmillan, Trevor S. Smart, and Andrew P. Thorburn. A field experiment involving cash and hypothetical charitable donations. *Environmental and Resource Economics*, 14(3):399–412, 1999.

Evan Mayo-Wilson, Aamer Imdad, Kurt Herzer, Mohammad Yawar Yakoob, and Zulfiqar A. Bhutta. Vitamin a supplements for preventing mortality, illness, and blindness

in children aged under 5: systematic review and meta-analysis. *BMJ: British Medical Journal*, 343, 2011.

Neal E. Miller. Experimental studies of conflict. 1944.

Thomas Moriarty. Crime, commitment, and the responsive bystander: Two field experiments. *Journal of Personality and Social Psychology*, 31(2):370, 1975.

National Council for Voluntary Organisations. What are the main trends in charitable giving, November 2014. URL `http://data.ncvo.org.uk/a/almanac12/what-are-the-main-trends-in-charitable-giving/`. Accessed May 1, 2014.

Stale Navrud and Knut Veisten. Validity of nonuse values in contingent valuation: An empirical test with real payments. In *Seventh EAERE Conference, Lisbon*, 1996.

Guido Peeters. The positive-negative asymmetry: On cognitive consistency and positivity bias. *European Journal of Social Psychology*, 1(4):455–474, 1971.

Cathy Pharoah and Sarah Tanner. Trends in charitable giving. *Fiscal Studies*, 18(4): 427–443, 1997.

William Revelle. R and analysis of variance, May 2013. URL `https://personality-project.org/r/r.anova.html`. Accessed April 29, 2014.

Milton Rokeach and Gilbert Rothman. The principle of belief congruence and the congruity principle as models of cognitive interaction. *Psychological Review*, 72(2):128, 1965.

Paul Rozin and Edward B. Royzman. Negativity bias, negativity dominance, and contagion. *Personality and social psychology review*, 5(4):296–320, 2001.

Donald Rumsfeld. DoD News Briefing - Secretary Rumsfeld and Gen. Myers - News Transcript, U.S. Department of Defense, February 2012. URL `http://www.defense.gov/transcripts/transcript.aspx?transcriptid=2636`. Accessed May 9, 2014.

Adrian Sargeant, Douglas C. West, and John B. Ford. Does perception matter?: an empirical analysis of donor behaviour. *The Service Industries Journal*, 24(6):19–36, 2004.

Susanne Schwager and Klaus Rothermund. Motivation and affective processing biases in risky decision making: A counter-regulation account. *Journal of Economic Psychology*, 2012.

David O. Sears. Positivity biases in evaluations of public figures. *Paper presented at the annual meeting of the American Psychological Association*, 1976.

David O. Sears and Richard E. Whitney. Political persuasion. *In I. deS. Pool, et al. (eds), Handbook of Communication*, 1973.

Kalle Seip and Jon Strand. Willingness to pay for environmental goods in norway: A contingent valuation study with real payment. *Environmental and Resource Economics*, 2(1):91–106, 1992.

Nate Silver. *The Signal and the Noise: The Art and Science of Prediction*. Puffin Books, 2013.

Paul Slovic, Baruch Fischhoff, and Sarah Lichtenstein. Response mode, framing, and information-processing effects in risk assessment. *New directions for methodology of social and behavioral science: Question framing and response consistency*, 11:21–36, 1982.

Deborah A. Small, George Loewenstein, and Paul Slovic. Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organizational Behavior and Human Decision Processes*, 102(2):143–153, 2007.

Gerald E. Smith and Paul D. Berger. The impact of direct marketing appeals on charitable marketing effectiveness. *Journal of the Academy of Marketing Science*, 24(3):219–231, 1996.

Fujian Song, S. Parekh, L. Hooper, Y. K. Loke, J. Ryder, A. J. Sutton, C. Hing, C. S. Kwok, C. Pang, and I. Harvey. *Dissemination and publication of research findings: an updated review of related biases.* Prepress Projects Limited, 2010.

Doron Sonsino. A note on negativity bias and framing response asymmetry. *Theory and decision*, 71(2):235–250, 2011.

StatSoft, Inc. Electronic statistics textbook, 2013. URL `http://www.statsoft.com/textbook/`. Accessed April 29, 2014.

Johan Swinnen and Pasquamaria Squicciarini. Mixed messages on prices and food security. *Science*, 335(6067):405–406, 2012.

Shelley E. Taylor. Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychological bulletin*, 110(1):67–85, 1991.

Richard Thaler. Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1(1):39–60, 1980.

The Giving Institute. Giving USA 2013, the annual report on philanthropy, 2013. URL `http://www.charitynavigator.org/index.cfm?bay=content.view&cpid=42#.U2o4DK2Sys3`. Accessed May 7, 2014.

David L. Thomas and Ed. Diener. Memory accuracy in the recall of emotions. *Journal of Personality and Social Psychology*, 59(2):291, 1990.

Edward B. Titchener. *An outline of psychology.* Macmillan, 1907.

Amos Tversky and Daniel Kahneman. Belief in the law of small numbers. *Psychological bulletin*, 76(2):105, 1971.

Amos Tversky and Daniel Kahneman. The framing of decisions and the psychology of choice. *Science*, 211:453–458, 1981.

Amos Tversky and Daniel Kahneman. Rational choice and the framing of decisions. *Journal of business*, pages S251–S278, 1986.

UNAIDS. *World AIDS day report 2011*. Geneva, 2011. URL `http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2011/jc2216_worldaidsday_report_2011_en.pdf`. Accessed November 26, 2013.

Timothy D. Wilson and Daniel T. Gilbert. Affective forecasting. *Advances in experimental social psychology*, 35:345–411, 2003.

Robert S. Woodworth and Harold Schlosberg. *Experimental psychology*. Oxford and IBH Publishing, 1954.

Wilhelm M. Wundt and Charles H. Judd. *Outlines of psychology*. W. Engelmann, 1907.

Appendices

Appendix A

BUY A KU
LEUVEN PEN
FOR 1€ AND
SEND 2€ TO
CHARITY!

KU LEUVEN

KOOP EEN KU
LEUVEN PEN
VOOR 1€ EN
SCHENK 2€
AAN HET
GOEDE DOEL!

KU LEUVEN

Appendix B

VOUCHER

Date: ....

Name and Surname (optional):[44] ....

Gender: □ M   □ F

Has donated ....€ for supporting the charity project   □ A   □ B

Katholieke Universiteit Leuven will add 1€ to the donation.

Signature (optional): ....

If you wish to remain informed about the total amount raised and the future development of the projects, please leave us your email: ....

Receipt

Katholieke Universiteit Leuven received ....€ for supporting the charity project □ A □ B and it will add 1€ to this donation.

Date: ....   Signature: ....

If you have any question concerning this fundraising activity or if you want to know more details about the projects, please feel free to contact andrea.guariso@kuleuven.be or mara.squicciarini@kuleuven.be.

---

[44]We left this line optional in case the subject wants to participate anonymously.

Appendix C

BUY A KU
LEUVEN
PEN FOR 1€
AND SEND
2€ TO
CHARITY!

March 18, 2014

Diestsestraat, Leuven

KU LEUVEN

KOOP EEN KU
LEUVEN PEN
VOOR 1€ EN
SCHENK 2€
AAN HET
GOEDE DOEL!

18 Maart 2014

Diestsestraat, Leuven

KU LEUVEN

Appendix D

```
Tukey multiple comparisons of means
    95% family-wise confidence level


Fit: aov(formula = var ~ categ, data = fr)


$categ

                          diff        lwr       upr     p adj
a-(a-b+)-a-(a-b-) -0.12500000 -1.4006224 1.1506224 0.9999858
a+(a+b-)-a-(a-b-) -0.41176471 -1.5501318 0.7266024 0.9464812
a+(a+b+)-a-(a-b-) -0.80000000 -2.2151760 0.6151760 0.6400285
b-(a-b-)-a-(a-b-) -0.20000000 -1.6151760 1.2151760 0.9998305
b-(a+b-)-a-(a-b-) -0.35294118 -1.4913083 0.7854260 0.9766731
b+(a-b+)-a-(a-b-)  0.37500000 -0.9006224 1.6506224 0.9827452
b+(a+b+)-a-(a-b-)  0.40000000 -1.0151760 1.8151760 0.9862471
a+(a+b-)-a-(a-b+) -0.28676471 -1.2461234 0.6725940 0.9810231
a+(a+b+)-a-(a-b+) -0.67500000 -1.9506224 0.6006224 0.7123345
b-(a-b-)-a-(a-b+) -0.07500000 -1.3506224 1.2006224 0.9999996
b-(a+b-)-a-(a-b+) -0.22794118 -1.1872999 0.7314175 0.9951398
b+(a-b+)-a-(a-b+)  0.50000000 -0.6187949 1.6187949 0.8529964
b+(a+b+)-a-(a-b+)  0.52500000 -0.7506224 1.8006224 0.8988065
a+(a+b+)-a+(a+b-) -0.38823529 -1.5266024 0.7501318 0.9607010
b-(a-b-)-a+(a+b-)  0.21176471 -0.9266024 1.3501318 0.9989643
b-(a+b-)-a+(a+b-)  0.05882353 -0.7086634 0.8263105 0.9999974
b+(a-b+)-a+(a+b-)  0.78676471 -0.1725940 1.7461234 0.1859130
b+(a+b+)-a+(a+b-)  0.81176471 -0.3266024 1.9501318 0.3449347
b-(a-b-)-a+(a+b+)  0.60000000 -0.8151760 2.0151760 0.8838503
```

```
b-(a+b-)-a+(a+b+)   0.44705882 -0.6913083 1.5854260 0.9191425

b+(a-b+)-a+(a+b+)   1.17500000 -0.1006224 2.4506224 0.0924105

b+(a+b+)-a+(a+b+)   1.20000000 -0.2151760 2.6151760 0.1549454

b-(a+b-)-b-(a-b-)  -0.15294118 -1.2913083 0.9854260 0.9998792

b+(a-b+)-b-(a-b-)   0.57500000 -0.7006224 1.8506224 0.8474161

b+(a+b+)-b-(a-b-)   0.60000000 -0.8151760 2.0151760 0.8838503

b+(a-b+)-b-(a+b-)   0.72794118 -0.2314175 1.6872999 0.2693267

b+(a+b+)-b-(a+b-)   0.75294118 -0.3854260 1.8913083 0.4422811

b+(a+b+)-b+(a-b+)   0.02500000 -1.2506224 1.3006224 1.0000000
```

Note: All lines have to be read as follows: for instance, the first line `a-(a-b+)-a-(a-b-)` represents the difference of means between the first condition (`a-(a-b+)`) and the second condition (`a-(a-b-)`). Moreover, the condition `a-(a-b+)` corresponds to the mean of donors that choose the message A when they were faced with the combination A-B+. This method of reading applies to the whole output.

Appendix E

```
# ANOVA Analysis #


> fr <- data.frame(var = c(2, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 1, 2, 0, 2,
2, 2, 1, 1, 1, 3, 1, 0, 2, 0, 0, 1, 0, 0, 1, 2, 1, 2, 1, 0, 2, 0, 1, 0, 0, 1,
0, 0, 1, 2, 0, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 2, 0,
1),
+ categ = factor(c("a-(a-b-)", "a-(a-b-)", "a-(a-b-)", "a-(a-b-)", "a-(a-b-)",
"b-(a-b-)", "b-(a-b-)", "b-(a-b-)", "b-(a-b-)", "b-(a-b-)", "a-(a-b+)", "a-(a-b+)",
"a-(a-b+)", "a-(a-b+)", "a-(a-b+)", "a-(a-b+)", "a-(a-b+)", "a-(a-b+)", "b+(a-b+)",
"b+(a-b+)", "b+(a-b+)", "b+(a-b+)", "b+(a-b+)", "b+(a-b+)", "b+(a-b+)", "b+(a-b+)",
"a+(a+b+)", "a+(a+b+)", "a+(a+b+)", "a+(a+b+)", "a+(a+b+)", "b+(a+b+)", "b+(a+b+)",
"b+(a+b+)", "b+(a+b+)", "b+(a+b+)", "a+(a+b-)", "a+(a+b-)", "a+(a+b-)", "a+(a+b-)",
"a+(a+b-)", "a+(a+b-)", "a+(a+b-)", "a+(a+b-)", "a+(a+b-)", "a+(a+b-)", "a+(a+b-)",
"a+(a+b-)", "a+(a+b-)", "a+(a+b-)", "a+(a+b-)", "a+(a+b-)", "a+(a+b-)", "b-(a+b-)",
"b-(a+b-)", "b-(a+b-)", "b-(a+b-)", "b-(a+b-)", "b-(a+b-)", "b-(a+b-)", "b-(a+b-)",
"b-(a+b-)", "b-(a+b-)", "b-(a+b-)", "b-(a+b-)", "b-(a+b-)", "b-(a+b-)", "b-(a+b-)",
"b-(a+b-)", "b-(a+b-)")))
#a-(a-b-) denotes that the donor chose A- when he was facing the combination
A-B-, etc.
> summary(fr)                                          #show a summary of the data
      var            categ
 Min.   :0.0   a+(a+b-):17
 1st Qu.:0.0   b-(a+b-):17
 Median :1.0   a-(a-b+): 8
 Mean   :0.8   b+(a-b+): 8
 3rd Qu.:1.0   a-(a-b-): 5
```

```
 Max.   :3.0   a+(a+b+): 5

              (Other) :10
```

```
> fit <- aov(var ~ categ, fr)                    #do the analysis of variance

> summary(fit)                                   #show the summary table

          Df Sum Sq Mean Sq F value Pr(>F)

categ      7   7.65  1.0929   2.148 0.0514 .

Residuals  62  31.55  0.5089

---

Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

#or

> anova(lm(var ~ categ, fr))        #another method for the analysis of variance

Analysis of Variance Table


Response: var

         Df Sum Sq Mean Sq F value  Pr(>F)

categ      7   7.65 1.09286  2.1476 0.05141 .

Residuals 62   31.55 0.50887

---

Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

> print(model.tables(fit,"means"),digits=3)            #report the means and

                                                       the number of subjects/cell

Tables of means

Grand mean


0.8


 categ
```

```
    a-(a-b-) a-(a-b+) a+(a+b-) a+(a+b+) b-(a-b-) b-(a+b-) b+(a-b+) b+(a+b+)
         1    0.875    0.588      0.2      0.8    0.647     1.38      1.4
rep      5    8.000   17.000      5.0      5.0   17.000     8.00      5.0
> boxplot(var~categ,fr)                              #graphical summary
```
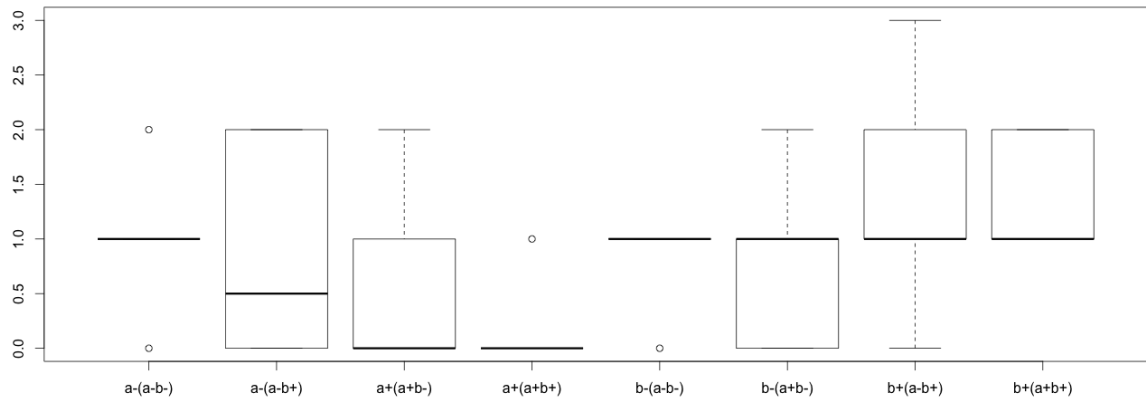


```
#If the ANOVA had shown a significant difference between the means, we can do
a Tukey HSD Test in order to know between which factors there is a difference.
The ANOVA did not show a significant difference (p-value = 0.0514) but we still
do it in order to show the methodology:
> tes <- aov(var ~ categ, fr); TukeyHSD(tes)            #Tukey HSD test with an
                                                    interval confidence of 0.95

  Tukey multiple comparisons of means
    95% family-wise confidence level


Fit: aov(formula = var ~ categ, data = fr)


$categ
```

	diff	lwr	upr	p adj
a-(a-b+)-a-(a-b-)	-0.12500000	-1.4006224	1.1506224	0.9999858
a+(a+b-)-a-(a-b-)	-0.41176471	-1.5501318	0.7266024	0.9464812
a+(a+b+)-a-(a-b-)	-0.80000000	-2.2151760	0.6151760	0.6400285
b-(a-b-)-a-(a-b-)	-0.20000000	-1.6151760	1.2151760	0.9998305
b-(a+b-)-a-(a-b-)	-0.35294118	-1.4913083	0.7854260	0.9766731
b+(a-b+)-a-(a-b-)	0.37500000	-0.9006224	1.6506224	0.9827452
b+(a+b+)-a-(a-b-)	0.40000000	-1.0151760	1.8151760	0.9862471
a+(a+b-)-a-(a-b+)	-0.28676471	-1.2461234	0.6725940	0.9810231
a+(a+b+)-a-(a-b+)	-0.67500000	-1.9506224	0.6006224	0.7123345
b-(a-b-)-a-(a-b+)	-0.07500000	-1.3506224	1.2006224	0.9999996
b-(a+b-)-a-(a-b+)	-0.22794118	-1.1872999	0.7314175	0.9951398
b+(a-b+)-a-(a-b+)	0.50000000	-0.6187949	1.6187949	0.8529964
b+(a+b+)-a-(a-b+)	0.52500000	-0.7506224	1.8006224	0.8988065
a+(a+b+)-a+(a+b-)	-0.38823529	-1.5266024	0.7501318	0.9607010
b-(a-b-)-a+(a+b-)	0.21176471	-0.9266024	1.3501318	0.9989643
b-(a+b-)-a+(a+b-)	0.05882353	-0.7086634	0.8263105	0.9999974
b+(a-b+)-a+(a+b-)	0.78676471	-0.1725940	1.7461234	0.1859130
b+(a+b+)-a+(a+b-)	0.81176471	-0.3266024	1.9501318	0.3449347
b-(a-b-)-a+(a+b+)	0.60000000	-0.8151760	2.0151760	0.8838503
b-(a+b-)-a+(a+b+)	0.44705882	-0.6913083	1.5854260	0.9191425
b+(a-b+)-a+(a+b+)	1.17500000	-0.1006224	2.4506224	0.0924105
b+(a+b+)-a+(a+b+)	1.20000000	-0.2151760	2.6151760	0.1549454
b-(a+b-)-b-(a-b-)	-0.15294118	-1.2913083	0.9854260	0.9998792
b+(a-b+)-b-(a-b-)	0.57500000	-0.7006224	1.8506224	0.8474161
b+(a+b+)-b-(a-b-)	0.60000000	-0.8151760	2.0151760	0.8838503
b+(a-b+)-b-(a+b-)	0.72794118	-0.2314175	1.6872999	0.2693267

```
b+(a+b+)-b-(a+b-)  0.75294118 -0.3854260 1.8913083 0.4422811

b+(a+b+)-b+(a-b+)  0.02500000 -1.2506224 1.3006224 1.0000000


> TukeyHSD(tes, conf.level = 0.90)                          #to change the level of
                                                              confidence to 0.90

   Tukey multiple comparisons of means
     90% family-wise confidence level


Fit: aov(formula = var ~ categ, data = fr)


$categ
                            diff         lwr        upr      p adj
a-(a-b+)-a-(a-b-) -0.12500000 -1.28634522 1.0363452 0.9999858

a+(a+b-)-a-(a-b-) -0.41176471 -1.44815070 0.6246213 0.9464812

a+(a+b+)-a-(a-b-) -0.80000000 -2.08839685 0.4883968 0.6400285

b-(a-b-)-a-(a-b-) -0.20000000 -1.48839685 1.0883968 0.9998305

b-(a+b-)-a-(a-b-) -0.35294118 -1.38932717 0.6834448 0.9766731

b+(a-b+)-a-(a-b-)  0.37500000 -0.78634522 1.5363452 0.9827452

b+(a+b+)-a-(a-b-)  0.40000000 -0.88839685 1.6883968 0.9862471

a+(a+b-)-a-(a-b+) -0.28676471 -1.16017882 0.5866494 0.9810231

a+(a+b+)-a-(a-b+) -0.67500000 -1.83634522 0.4863452 0.7123345

b-(a-b-)-a-(a-b+) -0.07500000 -1.23634522 1.0863452 0.9999996

b-(a+b-)-a-(a-b+) -0.22794118 -1.10135530 0.6454729 0.9951398

b+(a-b+)-a-(a-b+)  0.50000000 -0.51856714 1.5185671 0.8529964

b+(a+b+)-a-(a-b+)  0.52500000 -0.63634522 1.6863452 0.8988065

a+(a+b+)-a+(a+b-) -0.38823529 -1.42462129 0.6481507 0.9607010

b-(a-b-)-a+(a+b-)  0.21176471 -0.82462129 1.2481507 0.9989643
```

```
b-(a+b-)-a+(a+b-)   0.05882353 -0.63990777 0.7575548 0.9999974

b+(a-b+)-a+(a+b-)   0.78676471 -0.08664941 1.6601788 0.1859130

b+(a+b+)-a+(a+b-)   0.81176471 -0.22462129 1.8481507 0.3449347

b-(a-b-)-a+(a+b+)   0.60000000 -0.68839685 1.8883968 0.8838503

b-(a+b-)-a+(a+b+)   0.44705882 -0.58932717 1.4834448 0.9191425

b+(a-b+)-a+(a+b+)   1.17500000  0.01365478 2.3363452 0.0924105

b+(a+b+)-a+(a+b+)   1.20000000 -0.08839685 2.4883968 0.1549454

b-(a+b-)-b-(a-b-)  -0.15294118 -1.18932717 0.8834448 0.9998792

b+(a-b+)-b-(a-b-)   0.57500000 -0.58634522 1.7363452 0.8474161

b+(a+b+)-b-(a-b-)   0.60000000 -0.68839685 1.8883968 0.8838503

b+(a-b+)-b-(a+b-)   0.72794118 -0.14547294 1.6013553 0.2693267

b+(a+b+)-b-(a+b-)   0.75294118 -0.28344482 1.7893272 0.4422811

b+(a+b+)-b+(a-b+)   0.02500000 -1.13634522 1.1863452 1.0000000
```

```
#The p-values show no significant difference between the means for any condition.
```

List of Figures

List of Tables

Abbreviations

ANOVA — **AN**alysis **O**f **VA**riance

CV — **C**ontingent **V**aluation

GDP — **G**ross **D**omestic **P**roduct

HIV — **H**uman **I**mmunodeficiency **V**irus

NCAA — **N**ational **C**ollegiate **A**thletic **A**ssociation

NGO — **N**on-**G**overnmental **O**rganization

Tukey HSD — **Tukey H**onestly **S**ignificant **D**ifference

WTA — **W**illingness-**T**o-**A**ccept

WTP — **W**illingness-**T**o-**P**ay