Benjamin Mandl

CUES, BELIEFS, AND MEMORY

STOCKHOLM SCHOOL OF ECONOMICS

CUES, BELIEFS, AND MEMORY

This doctoral thesis contains four self-contained chapters.

"Not so irrelevant alternatives: How cue informativeness results in cue effects" studies when and why decoys and defaults affect choice behavior.

"Overestimation of information demand" demonstrates systematic overestimation of the amount of information that is considered before making a decision.

"Motivated beliefs and climate attitudes" studies whether changes in the cost of a pro-environmental action affect the perceived importance of that action.

"Forward looking motivated memory" investigates a potential strategic memory bias.

BENJAMIN MANDL is an economist working on topics in Behavioral and Experimental Economics. He is interested in questions related to human decision making and belief formation. Benjamin holds a B.A. in Economics from Brandeis University and a M.A. in Economics from the European University Institute.

ISBN 978-91-7731-222-2

DOCTORAL DISSERTATION IN ECONOMICS
STOCKHOLM SCHOOL OF ECONOMICS, SWEDEN 2022

Cues, Beliefs, and Memory

Benjamin Mandl

Akademisk avhandling

som för avläggande av ekonomie doktorsexamen
vid Handelshögskolan i Stockholm
framläggs för offentlig granskning
fredagen den 28 januari 2022, kl 13.15,
sal Ragnar, Handelshögskolan,
Sveavägen 65, Stockholm

Cues, Beliefs, and Memory

Benjamin Mandl

STOCKHOLM SCHOOL OF ECONOMICS

Dissertation for the Degree of Doctor of Philosophy, Ph.D.,
in Economics
Stockholm School of Economics, 2022

Cues, Beliefs, and Memory

ISBN 978-91-7731-222-2 (printed)
ISBN 978-91-7731-223-9 (pdf)

This book was typeset by the author using LaTeX.

Front cover photo: © Benjamin Mandl

Back cover photo: © Cecilia Hanzon

Printed by: BrandFactory, Gothenburg, 2021

Keywords: Cues, Decoy, Default, Nudge, Prediction, Information Demand, Motivated Beliefs, Memory, Recall, Environment, Emission, Meat.

To Fee

Foreword

This volume is the result of a research project carried out at the Department of Economics at the Stockholm School of Economics (SSE).

This volume is submitted as a doctoral thesis at SSE. In keeping with the policies of SSE, the author has been entirely free to conduct and present his research in the manner of his choosing as an expression of his own ideas.

SSE is grateful for the financial support provided by the Jan Wallander and Tom Hedelius Foundation which has made it possible to carry out the project.

Göran Lindqvist

Director of Research
Stockholm School of Economics

David Domeij

Professor and Head of the
Department of Economics
Stockholm School of Economics

Acknowledgements

It is common knowledge that the acknowledgements are the only part of a thesis that are read thoroughly. Thus, I decided to make them unreasonably long.

Throughout the journey of my Ph.D. thesis, I have had the good fortune to be accompanied by many wonderful people. Those friends and colleagues made the hard parts bearable and the good parts fun. I would like to express my gratitude for their help and support here.

I want to thank Anna Dreber for being the best supervisor I could have asked for. Her unparalleled drive, can-do attitude, and authenticity have been and will continue to be an inspiration. There have been uncountable instances in which I went into an advisor meeting frustrated and stuck and came out motivated and reassured. Anna also managed to create an open and welcoming environment, which was essential for my productivity and happiness as a PhD student. My thanks also go out to my co-supervisor Magnus Johannesson. Magnus is one of the sharpest minds I have had the pleasure to meet and his attention to detail and ability to dissect my research design, hypothesis, or analysis have been both immensely helpful and inspiring. Magnus was very generous with his time and always available for questions. My advisors gave me complete freedom to explore my research interests and supported me every step of the way.

I had the pleasure to work with amazingly smart and pleasant co-authors who I am fortunate to consider my friends. My first chapter was written together with Gustav Karreskog. Gustav's view of human behavior and learning has been a source of important insights within and outside of research, and it has allowed me to see behavioral economics in a new, and frankly more interesting, light. My second chapter was written in collaboration with Jimin Nam. Jimin and I met serendipitously during my research stay at Harvard University in my fourth year of my PhD. Our shared research interests alone could have filled this thesis. Aside from our productive and frictionless collaboration, I greatly enjoyed our countless brainstorming sessions. My third chapter was written with Eva Ranehill. I am very thankful to have had the opportunity to work with Eva on a topic that is close to my heart. I learned a lot from our collaboration and conversations.

During my fourth year I had the opportunity to visit the Vienna University of Economics and Business for three months. I want to thank Rupert Sausgruber for hosting me and Simone Häckl for being super helpful and integrating me into the department.

Following the stay, I visited the Harvard School of Engineering and Applied Sciences for six months. I am grateful for the time I spent at Harvard University and want to thank Yiling Chen for hosting me. I also want to thank Fiery Cushman, the Department of Economics, the Department of Psychology, and Harvard Business School for being extremely welcoming and open to visiting researchers. I am grateful for having been a part of this world class research environment.

I always felt that the best part of being a Ph.D. student is meeting bright and like-minded people. I enjoyed and learned a lot from Anna's other advisees in regular group meetings. I want to thank Binnur, Domenico, Felix, and Nurit especially for the constant mix of helpful feedback, interesting discussions, and plain fun. I want to thank Christofer and Svante for being great office mates and friends throughout my studies. I very much enjoyed sharing my office with Erik, even if it was only for a short time. Special thanks go out to the regular attendees of Behavioral Fika (and later, Behavioral Lunch) which was an integral part of my research environment at SSE. Adam, Aljoscha, Andrea, Andreas, Andrew, Atahan, Bengt, Christine, Claire, Elle, Emil, Evelina, Guilherme, Isak, Joakim, John, Jon, Julian, Malin, Martha, Mattias, My, Niklas, Oliver, Ricardo, Roza, Siri, Sreyashi, Thomas, Viking, Xueping, Yue, and Zoltan: thanks to each of you for shared coffees, lunches, study sessions, parties, and drinks. I also want to thank the faculty of SSE and Stockholm University for being approachable and always having open doors for me. I want to especially thank Tore Ellingsen, Erik Molin, Jon de Quidt, Jörgen Weibull, and Robert Östling for insightful conversations and research help.

SSE is lucky to have outstanding and helpful administrative staff. I want to thank Camilla Elwing Johansson, Ritva Kiviharju, Rasa Salkauskaite, Malin Skanelid, and Lyudmila Vafaeva for their help and assistance throughout my journey and for shared lunches.

I am also grateful for the generous financial support of SSE and the Tom Hedelius & Jan Wallander Foundation. The financial support has made this research, my international visits, and this thesis possible. I want to additionally mention that I am grateful for SSE's parental support. SSE was very helpful and flexible in allowing me to stay home with both of my daughters without administrative hassle.

My time in Stockholm would not have been the same without my friends. I want to especially thank Emma and Thomas, Andrea and Ted, Anette and David, Magdalena and David, Hannah and Jens, and Johanna for filling my weekends and evenings with joyful moments, cultural exchange, and connection to the world outside of academia. I am also thankful for my childhood friend Clemens for all of his support and for showing me that many issues and difficulties that Ph.D. students face are independent of the subject.

My parents Lucia and Christian's help allowed my wife and I to find an apartment in Stockholm and to visit the U.S. as a family of four, and I am deeply grateful for their assistance. I am also thankful for my parents-in-law Sabine and Alfons, who supported us with moves and visits and who hosted us innumerous times in Vienna. I want to thank my father for encouraging me and supporting me throughout my Ph.D. journey. To my own surprise I now agree with him that doing a Ph.D. is an important and satisfying endeavour for one's own personal development, even if one does not ultimately continue to work in academia.

My biggest thanks of all goes out to the love of my life, Fee. Her unwavering support, unconditional love, and endless patience in discussing research ideas, problems, and frustrations have been the backbone of this thesis. Moving to Sweden together has been the best decision of my life. Our time in Stockholm will always have a special place in my heart during which we became parents to our wonderful daughters Maya and Olivia and got married. Life's challenges and joys are best encountered as a team and I am so grateful to have Fee as my partner.

Stockholm, November 30, 2021
Benjamin Mandl

Contents

Introduction

This doctoral thesis is a collection of four distinct essays all studying different aspects of human behavior. Each chapter centers around an experimental investigation of human biases. Chapter 1 studies when and why contexts affect decision making, Chapter 2 demonstrates a bias in the ability to predict the information demand of others, Chapter 3 investigates self-serving belief formation in an pro-environmental decision setting, and Chapter 4 studies self-serving memory accuracy in an environmental policy relevant setting.

Abstracts for the four different chapters follow below.

* * *

Not so irrelevant alternatives: How cue informativeness results in cue effects.
joint with Gustav Karreskog

We study whether the informational content of a contextual cue affects the strength of cue effects on behavior for two contexts: choice decoys and defaults. Participants in a pre-registered, repeated choice experiment are asked to choose one of three options with payoffs that are difficult to estimate. In each choice problem, they are given subtle cues in the form of decoy choices or pre-chosen default options. A decoy is defined as a choice that is easily compared to another choice in the bundle but dominated by that comparable choice. A default is a pre-chosen choice option. Cues differ in informativeness by treatment condition. Informativeness is a measure of how helpful the cue is to making better decisions. We find a statistically significant effect of informativeness of the default cue but not the decoy cue on choice behavior. We run a second experiment that employs an updated measure and randomization procedure of informativeness and find that informativeness of the cue has a causal effect on choice behavior for both decoys and defaults, indicating that people learn to trust cues when cues are beneficial. These findings might indicate that the scope of nudging and behavioral policy depend on whether the choice architects' interests are aligned with the decision makers' preferences.

* * *

Overestimation of information demand
joint with Jimin Nam

We show in an online experiment that people have an inaccurate estimate of the amount of information that a decision maker considers before making a final decision. Participants are randomized to be either decision makers in a fully incentivized experiment or to be predictors. Decision makers receive piecewise, free information about their task and can submit their decision after each piece of information. Predictors are incentivized to predict the number of pieces of information the decision makers consider before submitting their final choice. We find that predictors overpredict the demand of information of the decision makers by a significant margin.

* * *

Motivated beliefs and climate attitudes
joint with Eva Ranehill

We study whether motivated cognition causes the perceived benefits of climate friendly actions to rise as the associated costs decrease. In our experiment, respondents are offered to donate to plant a tree after being randomized to receive either a high or a low discount to the cost of donation. Before respondents make a final decision to donate or not, we elicit their perceived importance of planting trees, how much CO_2 they believe planted trees sequester from the atmosphere, and to what extent they agree with the statement that their actions contribute to climate change. We find that respondents randomized to a low discount – and hence a higher cost – state a lower perceived importance of planting trees and agreement with individual responsibility for climate change. Our overall results are mixed as we do not find statistically significant differences in the quantitative measure. Our study highlights the possibility that motivated cognition contributes to the slow response to climate challenges and how related policies may impact beliefs and attitudes in important ways.

* * *

Forward looking motivated memory

This research tests a hypothesis of a strategic forward looking bias of memory. It asks whether subjects in a lab experiment systematically misremember information that is dissonant with an upcoming action. In the experiment subjects first learn about negative consequences of consuming meat (such as environmental or ethical issues of the production of meat) and are asked to accurately recall the information before tasting a sample of cured beef. I do not find any evidence of systematic memory distortions in our prespecified outcome measures. I discuss a range of potential explanations for our non-significant results and propose alternative research directions for investigating memory distortions.

Chapter 1

Not so irrelevant alternatives: How cue informativeness results in cue effects

Gustav Karreskog

Benjamin Mandl

Abstract

We study whether the informational content of a contextual cue affects the strength of cue effects on behavior for two contexts: choice decoys and defaults. Participants in a pre-registered, repeated choice experiment are asked to choose one of three options with payoffs that are difficult to estimate. In each choice problem, they are given subtle cues in the form of decoy choices or pre-chosen default options. A decoy is defined as a choice that is easily compared to another choice in the bundle but dominated by that comparable choice. A default is a pre-chosen choice option. Cues differ in informativeness by treatment condition. Informativeness is a measure of how helpful the cue is to making better decisions. We find a statistically significant effect of informativeness of the default cue but not the decoy cue on choice behavior. We run a second experiment that employs an updated measure and randomization procedure of informativeness and find that informativeness of the cue has a causal effect on choice behavior for both decoys and defaults, indicating that people learn to trust cues when cues are beneficial. These findings might indicate that the scope of nudging and behavioral policy depend on whether the choice architects' interests are aligned with the decision makers' preferences.

We are greatly thankful to Anna Dreber and Magnus Johannesson for insightful comments. We are especially grateful for comments and contributions by Nurit Nobel. We are also thankful to seminar participants at the

2021 Arne Ryde Workshop and at SSE for helpful comments and suggestions. We gratefully acknowledge funding by the Jan Wallander and Tom Hedelius Foundation.

1.1. Introduction

Choice architecture design is a widely used tool to bring about a desired behavior by adding contextual cues to a choice problem (Halpern, 2015). Two common context interventions that are used by choice architects to affect choices of their target group are having one of the possible options pre-chosen as a default or adding an inferior choice to a bundle as a decoy to make the desired choice more attractive (the former being more commonly used by behavioral policy makers and the latter by marketeers). Why and when these interventions work are not fully understood: theories to explain why nudges work include the idea that they increase salience for the desired choices (Bordalo et al., 2013), they provide a reference point that induces loss aversion (Tversky and Kahneman, 1991), or that they constitute (implicit) recommendations (Jachimowicz et al., 2019). Similarly, it is unclear why nudges fail. Sunstein (2017) lists strong antedecent preferences and potential counternudges by involved parties as main causes for failures of nudges. In this paper, we put forward and test the idea that choice contexts' (and thus nudges') effectiveness depends on whether the context is helping the decision maker make a better choice. We randomly vary the informativeness of the decoy and default cues and find that people generally follow the cue when the cue has been informative so far.

To illustrate, consider the following examples for an effective and an ineffective default nudge. The Swedish pension system allows everyone working in Sweden to personally choose into which investment fund they want to invest a part of their pension savings, called the Premium pension. By default, a low-fee, public fund is chosen. This fund has outperformed the private funds since inception. Most working age Swedes choose the default fund (Cronqvist et al., 2018). Our example of an ineffective nudge comes from personal experience: if you ever booked a Ryan Air flight, you will have experienced the multiple rounds of expensive, pre-selected additional purchases such as extra travel insurance or seat upgrades. At least for us, these nudges have absolutely zero effect on our purchasing behavior. Why do these nudges seem to differ in their effect? One important difference between these two situations is whether the decision maker can trust the choice architect to have their best interest at heart[1]: the implicit recommendation effect is clear for both situations but Swedes would have learned to trust the recommendations by the Swedish government and we learned to distrust the recommendations by a profit maximizing airline.

In this paper we present evidence from two experiments that decision makers' responses to choice contexts depend on whether the choice context helps to make a better

[1]While we focus on the alignment of incentives between choice architect and decision maker, there are other important reasons that may result in the pension default to be especially effective: actively opting out of the default plan requires a few active decisions such as logging in on the pension platform and making an active fund choice. Social norms or salience of the public pension fund might also play a role. We assume that these reasons affect the strength of defaults independently.

decision. In other words, choice contexts such as defaults or decoys act as implicit recommendations. The decision maker learns the informativeness of the recommendations and whether the recommendations are aligned with the decision maker's personal interests over the course of the experiment. In particular we show that people in an online choice experiment are able to determine whether a nudge in the form of a default or a decoy helps them make good decisions and thus whether to incorporate that nudge into their decision making model.

In both experiments, the participants' task is to select one out of three options. Each option is described by a combination of a solid area of different shapes and a price. Participants receive payoffs equal to the area of the option's shape minus the option's price. They repeat the task 40 times. Participants are told the areas and the possible payoff of each option after each round.

In addition to this general setup, participants are randomized to get one of two cues: default choices or decoy options. Participants in the former group have one option chosen as the default. Participants in the latter group always have choice options where two of the three options have the same shape and one of the two option is dominating the other (from here on we say the dominating option "has a decoy" and the dominated option "is the decoy").

The informativeness of the cues is randomly assigned. For the default group, the informativeness of the cues corresponds to a pre-determined probability for the best option to be the default. For the decoy group, the informativeness of the cue is determined by randomizing the probability that the first ranked options has the decoy. The two experiments differ in the randomization procedure of informativeness. In the first experiment, participants are first randomized into a cue type and then into either the informative or the uninformative treatment. The two treatments differ in the probability for the best option to be the default and for the best option to have a decoy for the two cue types, respectively. Due to the chosen operatlization of the informativeness in the first experiment, the informativeness differences between the informative and the uninformative treatments is relatively large for the default cue, and relatively little for the decoy cue.

We find that the informativeness of the signal has a strong effect on whether the cue affects choice in the default but not in the decoy treatment. Presumably, we fail to find a significant effect because of the small difference in the probability for the best option to have decoy between the informative and the uninformative decoy group. We run a second, pre-registered experiment in which we remedy this design choice and randomize informativeness and test the effects of informativeness more directly. In particular, we assign every participant a random probablity between 0 and 1 that determines the probability for the best option to be the default or to have a decoy. In the second experiment, we find statistically significant treatment effects for both cue types. This implies that decision makers are able to learn to trust and follow cues depending on whether it helped them make better decisions in the past. Notably, a completely uninformative cue does not result in a negative effect on whether the cue affects choice, as could theoretically be expected. Lastly, depending on the informativeness of the cue, the effects of decoy and default are surprisingly similar in size.

Research into the decoy effect (also called attraction or compromise effect) has had a special attraction because the behavioral finding constitutes a violation of the independence of irrelevant alternatives, which states that the relative choice preference between any pair of options is unaffected by adding or removing other options in a bundle (Luce, 2012, von Neumann and Morgenstern, 2007). Failures to replicate the decoy effect outside of few well defined situations led to doubt of the importance and robustness of the finding (Frederick et al., 2014, Yang and Lynn, 2014); additionally, most prior research used non-incentivized research designs (Lichters et al., 2017).

More recent studies have been able to demonstrate the decoy effect: Müller et al. (2014) demonstrate the decoy effect in an incentivized laboratory experiment. Subjects choose between two different ballpoint pens in the main condition and choose the target that has a decoy more often. Crosetto and Gaudeul (2016) use areas and prices in a within participant design to estimate the strength of the effect with respect to the monetary utility cost of choosing the choice that has a decoy vs the originally preferred choice. Kaptein et al. (2016) demonstrate in a large scale online experiment with hypothetical choices that failure to replication might relate to the distance of the attribute values between target and decoy. Farmer et al. (2017) introduce a strict expected value ordering of the choices to show that the decoy effect persists with that change in design. Lichters et al. (2017) show that the strength of the decoy effect increases when the choices are implemented and economically meaningful. Evangelidis et al. (2018) propose and provide evidence that the decoy effect is stronger for the option that is second ranked than for an option that is first ranked on the most important attribute. Padamwar et al. (2019) shows that increasing the range of the attributes of the choices in the bundle increases the decoy effect. Even though the decoy effect does exist, the reasons and mechanisms of the effect are not fully understood.

There are multiple possible explanations for why a decoy effect can be observed: a decoy could give the decision maker a reason to choose one option over the other (Shafir et al., 1993, Dietrich and List, 2016, Gomez et al., 2016), a decoy could change the weighting of the different attributes (Huber et al., 1982, Tversky and Simonson, 1993), or the decoy could affect the choice process itself (Ariely and Wallsten, 1995). Castillo (2020) investigates these potential explanations. The author finds a decoy effect using gambles and additionally describes what he calls a "range effect", which has the opposite effect to the weighting explanation above. This range effect results in increasing the relative weighting of these attributes for which the decoy increases the total range. Natenzon (2019) provides an alternative account for the decoy effect. In the model by Natenzon (2019) the decoy effect is rationalized by describing the decision maker as a Bayesian learner: a decoy acts as a signal for the relative value between the option that has a decoy (i.e. the target) and the decoy and the decision maker chooses the option with the highest posterior mean belief (see also Gerasimou, 2016 for a similar model).

We contribute to the decoy literature in two ways: one, we demonstrate a clear decoy effect in options with an objective, absolute preference ranking. Second, we provide evidence in line with the explanations by Natenzon (2019) for why there even should be a decoy effect: a decoy is a cue that can help make better decisions. Third, we can directly compare the effect size of the decoy effect with a default nudge, combining the study of two context effects that were previously only studied in isolation. To our knowledge we are the first to study the two seemingly unrelated cues of default and decoy choices and provide a common explanation for the two effects. Thus, it is also conceivable that other nudges such as social comparison messages or reminders, prominent placement of the target choice, etc. can be modelled and studied in the same framework as defaults and nudges. We leave this task up to future research.

Our work also contributes to our understanding of why nudges work and when they fail. Recent failures to replicate or bring about persistent behavior change of behavior nudge policies has led to the identification of a few possible causes: e.g. strong antecedent preferences and counternudges (Sunstein, 2017). Our research provides direct evidence that the informativeness of the nudge (or cue) might influence the effectiveness of the intervention. This finding is in line with with an earlier meta study by Jachimowicz et al. (2019) who underline that defaults are more effective when they operate through endorsements. Related to this, Tannenbaum and Ditto (2012) argue and provide evidence for that agents who are nudged try to infer the beliefs and goals of the nudger and that this assessment influences the effectiveness of the nudge. This assessment seems similar to how our participants infer whether the cues are helpful over time and then learn to trust or ignore them. Through this framework, the following studies that describe failed nudges can more easily be understood: Beshears et al. (2010) show that unusually large default contribution rate leads to the majority of employees to opt out and shift to a lower contribution rate. Bronchetti et al. (2011) show a failure of a default of investing tax refunds into saving funds, presumably because the decision makers already had plans on how to use the refund. Altmann et al. (2019) show that defaults in a donation experiment do not result in larger aggregate donations. In these studies it seems that the incentives of the choice architects, to increase savings or donation rates, do not align with the decision makers' preferences and thus these cues are ignored. Lastly, Löfgren et al. (2012) find no default effect among experienced field practitioners. This can be explained in our framework if cues are seen as information carriers in choice value estimations: if decision makers are very well informed, choice cues do not provide additional, choice relevant information and will be ignored.

Lastly, our findings are related to a recent discussion about the use of "bad" nudges by private businesses or nudges that appear partisan to influence the consumers' or voters' behaviors. This increasing use has been seen as potentially problematic (Tannenbaum, Fox, et al., 2017)[2]. Examples for bad nudges include the often-times default cookie settings in GDPR compliant cookie banners, the default newsletter subscriptions, or add-on insurance purchases. Our research shows that decision makers and consumers learn about the intentions and helpfulness of these nudges even under small stakes. Our findings can be seen as a first step to mitigate these growing concerns: in accordance with the argument by Gigerenzer (2018) that many of the biases behavioral economists find might actually confuse intelligent inferences with logical errors, it is possible that we do not need to be saved of potential traps and bad nudges that induce irrational behavior because we are actually learning quickly whether we can trust a signal. Unfortunately, this might also imply that positive behavior change that is not fully aligned with personal preferences might often require more than a default nudge to bring about consistent change in behavior. As an example for such a failure, Hagmann et al. (2019) demonstrate that nudges in the context of pro-environmental policy can crowd out support for more restrictive policies while providing little benefit.

The rest of the paper proceeds as follows. Section 1.2 and Section 1.3 describe our two choice experiments, including the experimental design, hypotheses, procedures, and results. Our findings are discussed in Section 1.4 and Section 1.5 concludes.

1.2. Experiment 1

We study the effect of informativeness of cues on decision making in a preregistered online experiment. The experimental design, standard error corrections, hypotheses, excluded observations, significance levels, sample size, and analyses were pre-registered on OSF[3] prior to our data collection. We mention any deviations from the pre-analysis plan. In the experiment, participants are given 40 variations of a simple decision task. Participants are randomized into one of two cue types which we call the decoy and the default treatment. Furthermore, participants within the cue types are randomized to receive cues with two different levels of informativeness.

1.2.1. Experimental Design

The decision task is conceptually based on Crosetto and Gaudeul (2016) and on Trueblood et al. (2013). Participants are asked to choose one of three options. Each option is described by colored area of a few pre-determined shapes (rectangles, squares, circles, and ellipses) on a standardized canvas and a corresponding price. Each problem must be considered for at least 5 seconds before the participant can submit a choice in order to ensure each

[2] see also https://www.nytimes.com/2015/11/01/upshot/the-power-of-nudges-for-good-and-bad.html, accessed September 14 2021.

[3] see https://osf.io/q8y9b/.

decision problem is given some attention and consideration. The participant receives a payoff in experimental units that is equal to the area minus the price of the chosen option. After each choice, the participant receives feedback on the values of the areas and their total payoffs. At the end of the experiment, the experimental units are converted into GBP with a predetermined exchange rate. Examples of the task interface and feedback screens are provided in the appendix as part of the pre-analysis plan.

We generalize the design of earlier decoy and default studies by generating each option value by drawing from predetermined random distributions: first, three option values are drawn from the same normal distribution and then ranked. Area and price are then determined post hoc to equal the previously drawn option values, i.e. they are determined such that they generate the different treatments the participants are randomly assigned to.

Informative decoy

In this treatment, two of the three options are randomly paired to become a comparable pair. Comparable pairs always have the same shape, which is inspired by the discussion by Natenzon (2019) who argues that areas of figures with the same shape are easier to rank by size. In addition to two options having the same shape, the feature values of the higher ranked options are chosen such that they weakly dominate the feature values of the lower ranked option, i.e. the area (price) of the dominating option is higher (lower) or equal to the dominated option, with at least one attribute strictly so. We call the dominated option the "decoy" and say that the dominating option "has a decoy". In this setting, the option that has the decoy is the best ranked option two thirds of the time on average. Thus, the decoy is informative.

Non-informative decoy

In the non-informative treatment, we restrict the possible pairings of the three options to be between the worst option and either the best or the second best option, with equal probability. The lowest ranked option of the three is always the decoy. In this setting, having a decoy is uninformative about the relative ranking of two best options.

Informative default

In this treatment, the option features area, price and shape are determined randomly. One of the options is pre-selected to be the default. With probability 75%, the highest value option is pre-selected. With the remaining probability of 25%, each of the three options has a 33% chance of being pre-selected.

Non-informative default

In this treatment, the option features area, price, and shape are determined randomly. Each of the three options has a 33% chance of being pre-selected.

1.2.2. Hypotheses

We identify cue effects by subsetting our data to situations that are directly comparable. If there is an effect of informativeness on whether participants follow the cues, we should observe differences in choice behavior in directly comparable choice situations. We determine directly comparable choice situations as follows. Choice problems in the non-informative decoy treatment are functionally equivalent to choice problems in the informative decoy treatment where the worst option is randomly selected to be the decoy, which comprises 2/3 of the cases. Similarly, the choice problems encountered by subjects in the non-informative default treatment are functionally equivalent to the 25% of choice problems in the informative default treatment where each option has the same uniform probability of being the default.

Furthermore, we divide our hypotheses into determining first, as a sanity check of sorts, whether there is a cue effect within the informative treatments across the trials and our strict, main hypothesis tests.

Existence Hypothesis Decoy:

There is a decoy effect within the informative decoy treatment, i.e. the propensity to choose the option that has a decoy is higher than the propensity to choose the option without decoy when the decoy is the lowest ranked option.

To test this hypothesis the data is subsetted to all subjects in the informative decoy treatment group and to only those trials where the worst option is a decoy, which is roughly 67% of the trials in the treatment. We regress the indicator of whether the option was chosen on an indicator of whether it has a decoy. We predict a positive coefficient of the decoy indicator using a two-sided t-test and a significance level α of 0.05.

Existence Hypothesis Default:

There is a default effect within the informative default treatment , i.e. the propensity to choose the option that has a default is higher than the propensity to choose the option without default when the default is uniformly randomly chosen option.

To test this hypothesis the data is subsetted to contain only those trials where one of the options is pre-selected with a uniform probability over the three options, i.e. we consider only 25% of the data collected in this treatment. We regress the indicator of whether the option was chosen on an indicator of whether it was the default. We predict a positive coefficient of the default indicator using a two-sided t-test and a significance level α of 0.05.

Our main hypotheses establish that subjects in the informative treatment group choose an item with a context effect more often than subjects in the uninformative treatment group in exactly comparable choice situations.

Main Hypothesis Decoy:

Participants in the informative decoy treatment follow the decoy more often than participants in the non-informative decoy treatment

To test this hypothesis the data is subsetted to contain only those trials where the worst option is a decoy, which is roughly 67% of the trials in the informative decoy treatment and 100% of the trials in the non informative decoy treatment. Then we regress the frequency of choosing the option that has a decoy on an informative treatment indicator. We predict a positive coefficient of the treatment indicator using a two-sided t-test and a significance level α of 0.05.

Main Hypothesis Default:

Participants in the informative default treatment follow the default more often than participants in the non-informative default treatment

To test this hypothesis the data is subsetted to contain only those trials where one of the options is pre-selected with a uniform probability over the three options, i.e. we consider only 25% of the data collected in this treatment. We use 100% of the data points in the non-informative default treatment. Then we regress the frequency of choosing the option that is the default on an informative treatment dummy. We predict a positive coefficient of the treatment indicator using a two-sided t-test and a significance level α of 0.05.

1.2.3. Procedures

Data was collected from May till June 2021 on Prolific Academic, a popular online social science laboratory (see Palan and Schitter (2018) and Peer et al. (2017) for overviews of the prolific platform and the participants).

658 participants started the experiment and 625 completed the study. The average age of our respondent is 27 and 38% of the respondents are women. We preregistered a sample size of at least 150 participants per cell. Our final sample size per cell is 161 in the informative decoy cell, 156 in the non-informative decoy cell, 158 in the informative default cell, and 150 in the non-informative default cell.

Participants read and signed an informed consent form prior to beginning the study. After giving consent, all participants received the same instructions and were then able to practice the task for three rounds. The practice rounds were generated according to the participants' respective treatment assignments. After practice, all participants needed to pass a short quiz with 4 questions. If they answered incorrectly more than 10 times, they were unable to proceed with the study. Participants collected points in each of the 40 rounds. They received points equal to the payoff of the choice they chose, i.e. the area minus the price. After the experiment, the points were converted into GBP according to an exchange rate known to the subjects. The average payment was 2.9 GBP.

1.2.4. Results

The main results are summarized in Table 1.1. Columns (1) and (3) describe our preliminary tests and show that participants in the informative treatment cells responded to the contexts: Whether an option has a decoy (or is the default) strongly predicts whether it is chosen more often both in the informative decoy group (column (1)) and the informative default group (column (3)).

In our main analysis, we subset the data to those comparable situations cases where the worst of the three options happens to be a decoy. We do this because the subsetted cases are directly comparable between the informative and the uninformative decoy group (in all our cases that we analyse, the decoy is the worst option. We can exactly compare behavior between treatment and control in these cases and see whether participants use the decoy cue more often in the *informative decoy* group.) When the worst option is the decoy, the second best and the best option are equally likely to have a decoy.

We can then test whether the chosen options had a decoy more often in the *informative decoy* treatment or in the *non-informative decoy* treatment, in other words, we compare mean likelihoods across two groups. We run an OLS regression on a binary indicator that is equal to 1 if the chosen option has a decoy and 0 otherwise on the treatment indicator. The result is non-signficant ($p = 0.0552$), see Table 1.1.

We run a similar regression for our default analysis. We regress whether the chosen option was the default option across treatments. Here we also subset to only those 25% of the cases in the *informative default* treatment in which all three options have the same probability of being a default. We preregistered to regress whether the chosen option was the default option on the treatment assignment indicator, which is significant ($p < 0.0001$).

We ran additional, non-preregistered analyses to understand why the pre-registered decoy analysis turned out to be non-significant, while there seems to be a clear effect in the default case. In our experiment design, we varied which of the three options would be paired to become a comparable pair (i.e. a decoy and a target): in the informative treatment, each combination of the three options is possible and in the non-informative treatment, one of the two top ranked options is paired with the worst option. Effectively, we directly varied the probability that the best option in the choice triplet has a decoy. In our *non-informative decoy treatment* cell the best and second best option are equally likely

Table 1.1. OLS regressions Experiment 1

	Dependent variable:			
	Chosen option	Chosen has decoy	Chosen option	Chosen is default
	(1)	(2)	(3)	(4)
Constant	0.419***	0.560***	0.248***	0.341***
	(0.008)	(0.008)	(0.008)	(0.007)
Has decoy	0.161***			
	(0.011)			
Is default			0.255***	
			(0.014)	
Treatment = Informative		0.021		0.163***
		(0.011)		(0.020)
Observations	8,368	10,186	4,551	7,517
R^2	0.026	0.0004	0.065	0.0182

Notes: Standard errors in parentheses, robust and clustered on individual level. Column (1) and Column (3) observations are options within the informative treatment for the decoy and default cue type, respectively. Column (2) and Column (4) observations are chosen options in comparable situations over both informativeness treatments for the decoy and default cue type, respectively.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

to have a decoy every round but in the *informative decoy treatment* cell, the probability for the best option to have a decoy is 67%. We measure this in a new variable called informativeness which equals the share of previously experienced trials in which the best option has a decoy in a given round. To calculate this measure, we first define a trial to be informative if the best option has the decoy.

$$\text{trial informative} = \begin{cases} \text{1 if best option has a decoy} \\ \text{0 otherwise} \end{cases}$$

We calculate the share of such informative trials so far for each round in a variable called informativeness.

$$\text{informativeness}_r = \sum_{i=1}^{r-1} \frac{\text{trial informative}_i}{r-1}$$

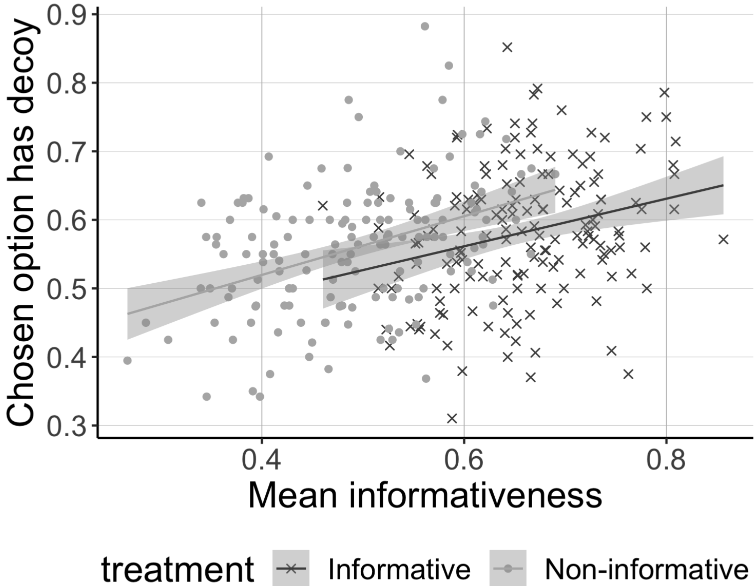
Where r is rounds from 1 to 43, rounds 1 to 3 are practice rounds, and rounds 4 to 43 are the experimental rounds.

To illustrate, if a participant is in her fourth round and in two of her previous rounds the best option had a decoy and in one of her previous rounds the second best option had a decoy, then the cumulative informativeness in round 4 is 0.67. The cumulative informativeness captures whether the realized sequence of decoy signals so far "taught" the participant to follow the decoy.

In our non-preregistered exploratory analysis, we show that informativeness measured in this way strongly predicts whether participants choose the option that has a decoy more often. However, the probability of getting informative signals measured in this way (i.e. where the best option has a decoy) is very similar between our two treatment groups (50% vs 67%). Additionally, the participants play only 40 rounds. Thus, the informativeness of the signals vary a lot across individuals, but not as much across treatments.

This problem is illustrated in Figure 1.1 which plots whether the chosen option has a decoy over the average cumulative informativeness per subject. Multiple participants in the *non-informative decoy treatment* cell had, by chance, options draws that resulted in more informative treatments than some of the participants in the *informative decoy treatment* cell (i.e. consider all the circles points in the scatterplot that had an average cumulative informativeness around .6). Vice versa, some participants in the *informative decoy treatment* cell experienced draws that mirrored the *non-informative decoy treatment* cell options. In other words, considering informativeness directly, our pre-registered regression analysis where we regress whether the chosen option has a decoy on treatment is akin to an intention to treat design where the implemented treatment effect is quite variable.

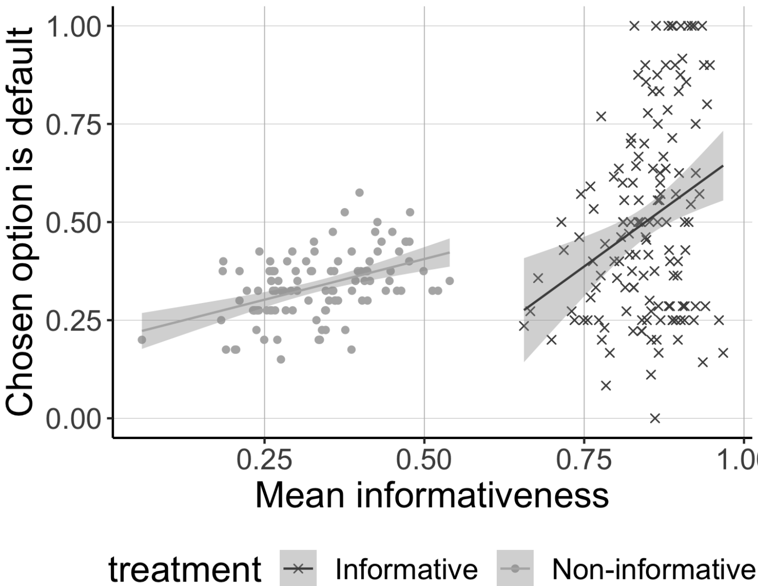
Figure 1.1. Scatterplot of choice with decoy chosen and informativeness



The ramifications of this design are less problematic in the default treatments, see Figure 1.2. We measure informativeness in the default group by the share of previously experienced trials in which the best option was the default. In this design, participants had a more differing experience between treatments: in our design, the best option was pre-selected in the *non-informative default treatment* cell in about 1/3 of the trials, and the best option was pre-selected in the *informative default treatment* cell in about 83% of trials. There is no overlap between informativeness experiences between the participants in the informative and non-informative default cells.

It seems that the reason for a non-significant treatment effect between the informative and the non-informative decoy treatments is due to the identified problems in the experiment design. We decided to run a follow up experiment with an updated design to test whether participants in a more informative cue treatment follow the cue more often.

Figure 1.2. Scatterplot of choice with default chosen and informativeness



1.3. Experiment 2

We pre-registered a new experiment where we randomize informativeness directly and ensure larger differences in treatments between participants on OSF[4]. In our new design, each participant is assigned a probability constant which describes the expected share of rounds in which the best option has a decoy (or is the default, respectively), i.e. we directly randomize the cumulative informativeness over all rounds.

1.3.1. Experimental design

The design is almost identical to the first experiment. Participants are recruited online to perform the task of choosing one of three options for 40 rounds. Participants are randomized into one of two cue types: decoy or default. We introduce a new randomization that is different to the first experiment: each participant is assigned a hidden random real number between 0 and 1 that describes the probability of receiving informative cues (called *assigned informativeness*) throughout the experiment. The participants do not observe their *assigned informativeness*. We call this the *Individual Treatment Level*. In 28 of the rounds, the cue is generated according to the *Individual Treatment Level*. The

[4] see https://osf.io/ds4xp

average number of rounds that are informative equals the assigned informativeness which varies on participant level. These 28 rounds can be understood as cue learning rounds. In the remaining 12 rounds, the cue is on average uninformative for all participants and identically generated for all participants with the same cue type. We will use these 12 *Comparison Rounds* to test our hypotheses. Each participant experiences a random order of 40 choices that are either part of the learning rounds or the *Comparison Rounds*. The way the decoy choices are constructed and the analysis is also updated relative to the first experiment and described below.

Decoy

In the 28 learning rounds, the options features are determined as follows. With probability p equal to the assigned informativeness, the best option is chosen to have a decoy. Then one of the remaining options (either the second or the third best) is chosen with 50% probability to be the best option's decoy. We call the best option and the decoy option a *Comparable Pair*. *Comparable Pairs* have the same *Shape*. One of the options in the *Comparable Pair* either has a larger area, a lower price, or both, i.e. it is dominating the other option in the *Comparable Pair*. We call the dominated option in the randomly chosen pair the decoy. With probability 1-p, the second best option is chosen to have a decoy. The worst option is the decoy, i.e. the second best and the worst option are a *Comparable Pair*. For the remaining 12 comparison rounds, the option features are determined as follows. First, either the best or the second best option is determined to have a decoy. This option is then paired with the lowest value option which is the decoy. So either the best or the second best option and the worst option are a *Comparable Pair* with equal probability.

Default

In the 28 learning rounds, the options features are determined as follows: The option features are determined randomly. The highest value option is pre-selected with a probability p equal to the assigned informativeness. In 1-p of cases, one of the other options is pre-selected with a uniform probability over the two non-best options. In the 12 comparison rounds, the options features are determined as follows: The option features are determined randomly. One of the options is pre-selected by default with a uniform probability over the three options.

1.3.2. Hypotheses

Our hypotheses establish whether the realized informativeness over all rounds predicts the propensity to follow the cue in exactly comparable situations.

Main Hypothesis - Decoy:

Participants who experience a higher realized informativeness follow the decoy more often in the 12 comparison rounds.

To test this hypothesis the data in the decoy treatment is subsetted to contain only the 12 comparison rounds where one of the two best options are paired randomly with the worst option to be a Comparable Pair. Then we regress the binary variable of whether the option with a decoy was chosen on the informativeness. We predict a positive coefficient of the informativeness measure using a two-sided t-test and a significance level α of 0.05.

Main Hypothesis - Default:

Participants who experience a higher realized informativeness follow the default more often in the 12 comparison rounds.

To test this hypothesis the data of the participants in the default treatment is subsetted to contain only the 12 comparison rounds where one of the options is pre-selected with a uniform probability over the three options. Then we regress the binary variable of whether the option that is the default was chosen on the informativeness. We predict a positive coefficient of the informativeness measure using a two-sided t-test and a significance level α of 0.05.

1.3.3. Procedures

Data was collected from September till October 2021 on Prolific Academic. We had a rolling recruitment with the pre-registered goal of having at least 300 participants for each cue type. We recruited a total of 701 participants for our study. The average age of our respondent is 27 and 52% of the respondents are women. Not all participants finished the study. As pre-registered, we also removed all answers that were not made using a desktop PC. The final sample comprises of 302 in the decoy cell and 306 in the default cell.

Identically to our first study, participants gave informed consent, received identical instructions, and were able to practice the task for three rounds. The practice rounds were generated according to the participants' respective *Individual Treatment Levels*. Participants then needed to pass a quiz with 4 questions. Participants who answered incorrectly more than 10 times needed to exit the experiment. The average payment was 2.9 GBP.

1.3.4. Results

We find a strong and statistically significant treatment effect in our main regressions for both the decoy and the default cue types, see Table 1.2. An increase of realized cumulative informativeness by 10 percentage points results in an increase of 1.29% probability of choosing the option with a decoy or a 1.58% probability of choosing the option with the default.

Table 1.2. OLS regressions Experiment 2

	Dependent variable:	
	Option with decoy was chosen	Option with default was chosen
	(1)	(2)
Constant	0.476***	0.283***
	(0.021)	(0.020)
Informativeness	0.129***	0.158***
	(0.037)	(0.039)
Observations	3,622	3,670
R^2	0.004	0.007

Notes: Standard errors in parentheses, robust and clustered on individual level. Observations are chosen options in comparable situations.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

To further visualize our results, we can graph the relationship between the mean choice probability of choosing the option with a decoy or that has the default on the observed informativeness, see Figure 1.3. A few things are important to note here. First, the decoy effect results in choosing the option that has a decoy starting from an informativeness of 0.33 and increases with informativeness. This can be rationalized in a very simple observation. If a decision maker knows nothing about the choice triple except that one choice dominates another she knows that choosing the dominating choice can never be the worst choice. Acting only on this insight is a profitable strategy: as long as informativeness is above 1/3, the expected value for following the decoy and choosing the target is positive.

This relationship between historic informativeness and the average payoff of choosing either the one with a decoy (blue) or the one without (red) is graphed in Figure 1.4. Even without considering the other choice features, choosing the option with the decoy has a higher expected payoff after historic informativeness crosses 0.33. For lower informativeness values, choosing the choice without a decoy has a higher expected payoff.

Figure 1.3. Scatterplot of choice with decoy chosen and informativeness

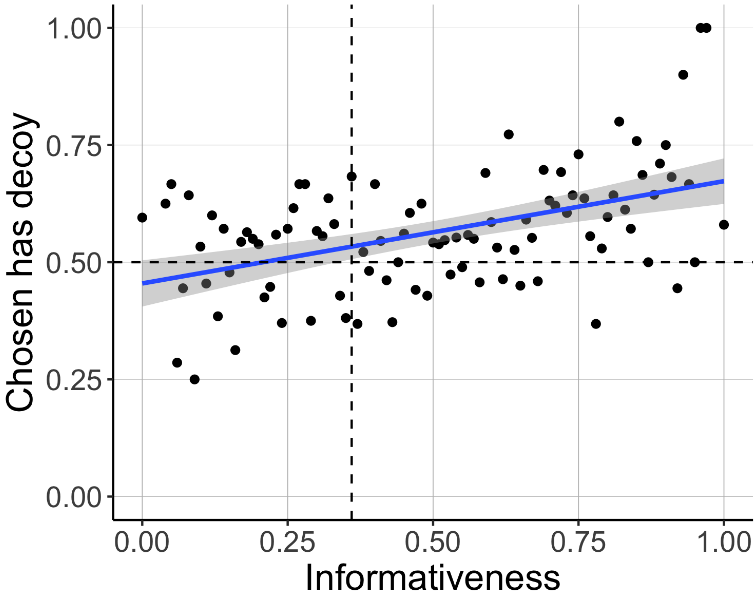
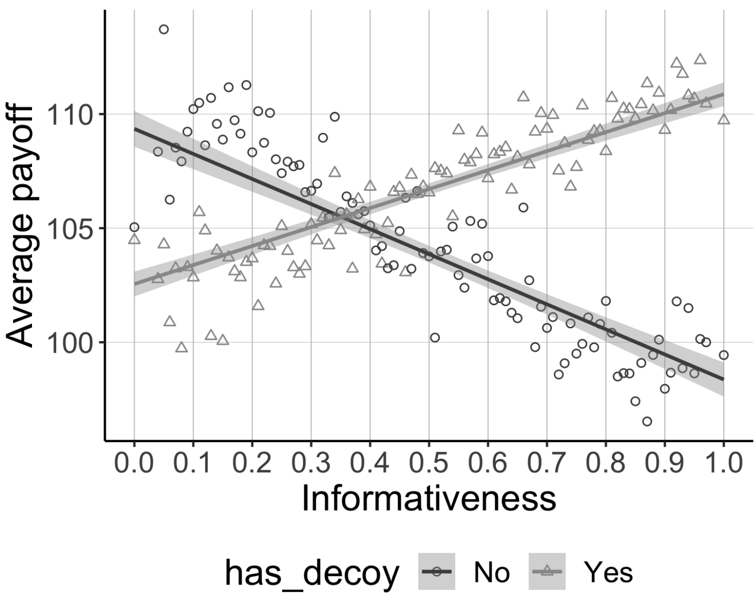
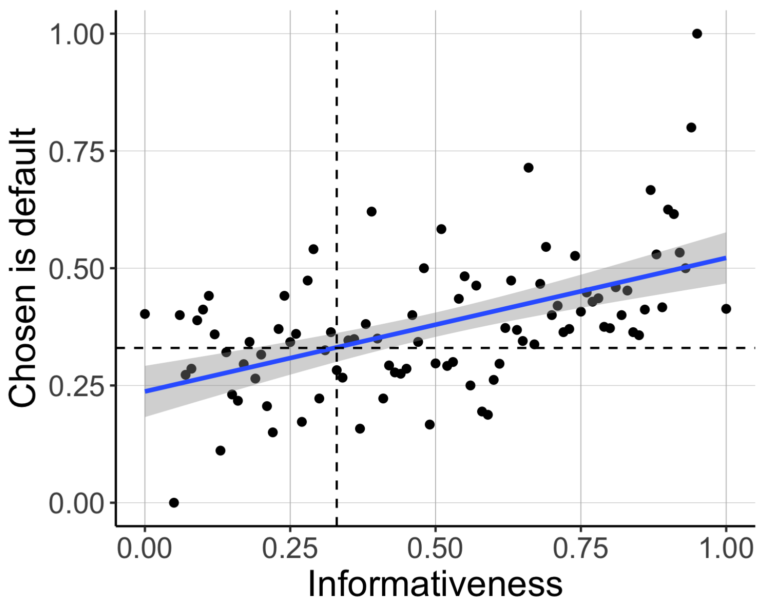


Figure 1.4. Relationship between informativeness and average payoffs of option with/without decoy



Graphing the default effect in Figure 1.5, we can observe that decision makers are surprisingly responsive to the historic informativeness. In particular, when historic informativeness is 1/3, i.e. when the signal has been historically uninformative, the default is also chosen 1/3 of the times. However, it is possible that this exact relationship is a coincidence that may not replicate. In fact, given earlier robust demonstrations of the default effect, one would expect the relationship between historic informativeness and choice probability to be upwards biased relative to the rational benchmark and it is not clear why this is not observed here. A future replication of the results might show a more differentiated relationship more in line with those expectations. Lastly, one can see that for extremely low values of informativeness, there seems to be negative default effect: options that are the default are chosen less often than 1/3. Unfortunately, our choice dataset is not rich enough to explore this relationship in detail but leave open the possibility of exploring negative default effects to future research.

Figure 1.5. Scatterplot of choice with default chosen and informativeness



1.4. Discussion

Our results provide insights to understanding the relationship between different context effects. First, it seems that there might be an underlying commonality between seemingly different context effects. To our knowledge, we are the first to study decoy and default effects in the same framework (and even using the same randomization procedure in Experiment 2). This improvement leads to an interesting observation: the effect sizes of the two context effects seem relatively similar in size. This is quite surprising, as we do not know of an overarching theory so far that would explain why two seemingly unrelated context effects might have a similar effect size. It remains to be seen whether this relationship between two different cues is robust to replication and whether it can be extended to other commonly seen choice context treatments. For example, it is possible that social information cues, reminders, or the choice presentation could be studied in the same or a similar framework to see whether their effect sizes are comparable and follow the same causal links as we identified for decoy and default effects.

Our results give important insights into how and why choice contexts affect choice. In our experiment, decision makers are very responsive to changes in informativeness and seem to be surprisingly rational in incorporating the context's information content into their choice. Thus, decoy and default related choice behavior that has been considered irrational or biased may be, at least in part, be explainable by an underlying, rational decision making process Gigerenzer (2018): people may regularly rationally incorporate contextual cues into their decision. Context choice behavior that has traditionally been identified as irrational biases may stem from being put in unusual situations when participating in choice experiments and not from underlyind irrationality. Modelling this choice behavior by extending Natenzon (2019) could provide fruitful future research opportunities.

Our finding that people seem to be very responsive to the helpfulness of cues sheds light onto why and when nudges work and gives important insights for choice architects. In our experiment, decision makers do not fall prey to uninformative cues repeatedly. In fact, they are surprisingly responsive to changes in informativeness. It is possible, that this sophisticated choice behavior can be extended outside the laboratory into real world decisions: people might follow all sort of nudges more easily and consistently when those nudges contain contextual information that is of functional importance to helping people make better choices. However, when the goals of the nudge are not aligned with decision maker's preferences, consistent and lasting effects of nudges might be difficult to achieve. In other words, choice architects should consider that there might be plenty of scope to help people make choices that are in their constituents' personal interests, but much less scope to trick them into choices they do not actually want. For example, helping people navigate difficult retirement fund choices where goals are clearly aligned might prove more fruitful than nudging people to reduce their meat consumption (assuming people do not actually want to reduce their meat consumption). Lastly, it is important to note that our experimental results come from relatively abstract, online choice experiments and might thus lack in external validity. Future work should investigate the relationship between informativeness and nudge effectiveness in a field experiment or an experimental design that is more externally valid towards the classical nudging literature.

1.5. Conclusion

Choice contexts are a common tool to affect choice behavior. Its use has become increasingly common among policy makers to nudge people towards desired decisions, and has long been used by marketeers to increase revenues. Two commonly used context cues are decoys, a dominated option that is similar to the dominating target, and defaults, a pre-selected choice. In this paper, we test the hypothesis that choice contexts affect behavior when the cues have informative value that helps people make better decisions. In our first experiment, we randomize informativeness in the decoy setting with a relatively small difference in probability of the best option to have a decoy between the informative and the uninformative treatment groups, which unfortunately did not allow us to reject our null hypothesis. In our follow up experiment, we adapt how we randomize informativeness and find strong treatment effects for both the decoy and the default cues: by randomizing informativeness understood as the share of the experienced rounds in which the best option is the target, we also make the study of decoy and default choices directly comparable. Surprisingly, the effects are similar in size, leading us to conjecture that the two seemingly unrelated context effects of decoys and defaults might be closer related than previously considered. We leave the relationship between the context effects studied here and other commonly used nudges for future research.

Our finding has multiple interesting implications. First, we see that people are very reactive to changes in informativeness of subtle cues. Their surprisingly rational behavior indicates a reason for when and why nudges affect choice: when the nudge helps the decision maker make better decisions that are aligned with their preferences. This leads us to conjecture that choice architectures' scope might be bounded. When the nudge helps navigate difficult decisions towards the decision makers' goals it might be more effective and show more consistent results than when the nudge is used to trick people or push them towards costly behavior such as pro-environmental decision making. Due to our limited external validity, exploring and validating this conjecture is left as an important direction of future research.

1.A Appendix

1.A.1. Pre-analysis plan uploaded on OSF - Experiment 1

Cued Decision Making: Pre-Registration Document for Experiment Data Analysis

Authors: Gustav Karreskog, Benjamin Mandl

Introduction:

This paper studies whether decision makers use contextual information sophisticatedly to make better decisions. Context effects (e.g. attraction effect, compromise effect or choice defaults) have been studied before but it's unclear why they persist.This paper aims to provide the evidence that decision makers are able to learn the informational content of contexts and incorporate contexts to improve their decision making.

Participants are recruited on Prolific to participate in an experiment on decision making. Subjects will be randomized into one of four treatment cells: **informative decoy, non-informative decoy, informative default, non-informative default.**

Experimental procedure:

Participants are recruited on Prolific and randomized into one of four treatments. Only participants who speak English fluently and who have at least 10 prior submissions and an approval rate of 95% are recruited. Participants are restricted to desktop users (i.e. no mobile or tablets users). The experiment is coded in OTree. The procedure is as follows:

- Participants receive the same instructions regardless of treatment. They are given information on the task and their goals, and an example of the decision task.
- The Participants' task in the experiment is to select one out of three **options**. Each option is a combination of a solid area of different shapes and a price (see Figure 1 below). Participants are told that their payoff consists of the *Area* of the option's shape minus the option's *Price*. Participants receive the option value as a payoff in **experimental units**. They repeat the task **40 times** and need to consider each problem for at least 5 seconds, i.e. participants cannot submit an answer within the first 5 seconds of starting a decision problem. Participants receive feedback after each round (see Figure 2). The experimental units are transformed into GBP at the end of the experiment.

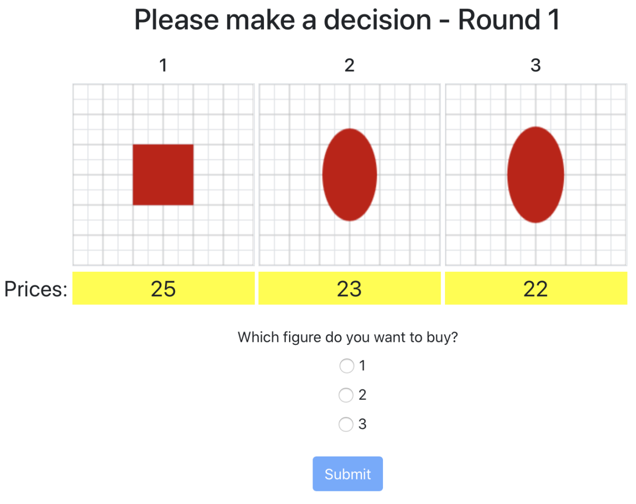


Figure 1: Decision table: Participants can choose one of three options which differ by area and price .

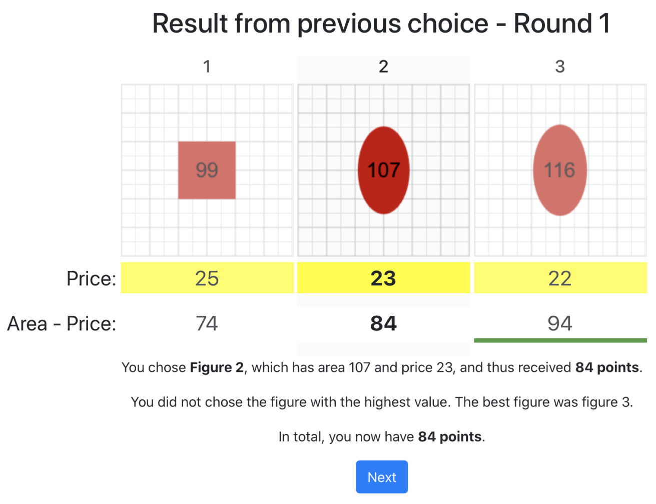


Figure 2: Decision feedback: Participants receive feedback on the bonus points for each option, their total bonus points and if they chose the figure with the highest value (green line)

- The Options' *Areas and Prices* are drawn randomly in each round and each participant. First, three **option values** are drawn from the same normal distribution and then ranked. *Area and Price* are then determined post hoc to equal the previously drawn values. The *Area and Price* are determined in special ways to generate the different treatments the participants were assigned to, which is described below.
- Participants are able to practice the task for three rounds. These rounds are not payoff relevant and are randomly generated according to treatment.
- Following the practice rounds, participants need to correctly answer 4 Quiz questions to proceed. In case they answer incorrectly more than 10 times, they are unable to proceed with the study and need to exit.
- Participants are then randomized into one of four treatments which affects how the option features are determined:
    - **Informative decoy:** In this treatment, the option features are created such that two options are randomly chosen to be a *Comparable Pair*. *Comparable Pairs* have the same *Shape*. One of the options in the *Comparable Pairs* either has a smaller *Area*, a higher *Price*, or both, i.e. it is dominated by the other option in the *Comparable Pair*. We call the dominated option in the randomly chosen pair *the decoy*.

  - **Non-informative decoy:** in this treatment, the option features are created such that one of the two top valued options is paired randomly with the lowest value option. Then, we make the lowest value option *a decoy* for the other randomly chosen option. In other words, we choose the *Area and Price* of the lowest valued option such that they are weakly dominated by the other option.
  - **Informative default:** in this treatment, the option features are determined randomly. The highest value option is pre-selected with a chance of 75%. In 25% of cases, one of the options is pre-selected with a uniform probability over the three options.
  - **Non-informative default:** in this treatment, the option features are determined randomly. One of the options is pre-selected by default with a uniform probability over the three options.
- After finishing 40 trials participants exit the survey.

All study material will be uploaded to OSF with the final pre analysis plan before the start of the data collection.

Standard error corrections

T-tests will assume unequal variances and all regressions will use heteroskedasticity robust standard errors. T-tests are always two sided. We test at a significance level of 5%. Regressions are OLS unless otherwise noted. Regressions will cluster standard errors on the participant level because we expect to see heterogeneity in the treatment effects across individuals.

List of variables for regressions:

The following variables will be used in our regressions described below:

- Chosen_option: equals 1 if the option was chosen by the participant. 0 otherwise.
- Option_has_decoy; equals 1 if the option chosen by the participant is weakly dominating another option in the choice set. 0 otherwise.
- Option_is_default; equals 1 if the option chosen is the default choice. 0 otherwise.
- Decoy_informative: equals 1 if participant is in the informative_decoy treatment. 0 otherwise.
- Default_informative: equals 1 if participant is in the informative_default treatment. 0 otherwise.

Tests for the existence of cue effects:

We will first test whether we can identify our cue effects within the informative decoy/default treatments. We run **Ordinary Least Square regressions** with **clustered standard errors** on participant levels for the hypothesis tests. We remove all choices of the worst option in the informative and uninformative decoy treatment. If the worst option is a decoy and is chosen we expect the choice to be mainly due to input error or inattention. If the worst option is not a decoy, it will not be relevant for our comparison in our main test.

1. Hypothesis:
    a. *There is a decoy effect within the **informative decoy** treatment , i.e. the propensity to choose the option that has a decoy is **higher** than the propensity to choose the option without decoy when the decoy is the lowest ranked option.*
        i. Test of Hypothesis:
            1. The data is subsetted to all subjects in the informative decoy treatment group and to only those trials where the worst option is a decoy, which is roughly 67% of the trials in the treatment.
            2. We run the following regression::
            3. Chosen_option ~ $\beta1 + \beta2$ Option_has_decoy + error
            4. Test: $\beta2 > 0$, t-test.
2. Hypothesis:
    a. *There is a default effect within the **informative default** treatment , i.e. the propensity to choose the option that has a default is **higher** than the propensity to choose the option without default when the default is uniformly randomly chosen option.*
        i. Test of Hypothesis:
            1. First, the data of the participants in the **informative default** treatment is subsetted to contain only those trials where one of the options is pre-selected with a uniform probability over the three options, i.e. we consider only 25% of the data collected in this treatment.
            2. We run the following regression:
            3. Chosen_option ~ $\beta1 + \beta2$ Option_is_default + error
            4. Test: $\beta2 > 0$, t-test.

Main Hypothesis Tests:

Our main hypotheses establish that subjects in the informative treatment group choose an item with a context effect more often than subjects in the uninformative treatment group in exactly comparable choice situations.

3. Hypothesis:

a. *Participants in the **informative decoy** treatment follow the decoy more often than participants in the **non-informative decoy** treatment.*
    i. Test of Hypothesis:
        1. First, the data is subsetted to contain only those trials where the worst option is a decoy, which is roughly 67% of the trials in the informative decoy treatment and 100% of the trials in the non informative decoy treatment.
        2. Participants in the informative decoy treatment choose the option that has a decoy more often than the participants in the non-informative decoy treatment.
        3. Regressing the frequency of choosing the option that has a decoy on an informative_treatment dummy, the treatment coefficient is significantly different from 0.
        4. [On subsetted data:] Chosen_option_has_decoy ~ $\beta 1$ + $\beta 2$ decoy_informative + error
        5. Test: $\beta 2 > 0$, t-test.

4. Hypothesis:
    a. *Participants in the **informative default** treatment follow the default more often than participants in the **non-informative default** treatment.*
        i. Test of Hypothesis:
            1. First, the data of the participants in the **informative default** treatment is subsetted to contain only those trials where one of the options is pre-selected with a uniform probability over the three options, i.e. we consider only 25% of the data collected in this treatment. We use 100% of the data points in the non-informative default treatment.
            2. Our hypothesis is operationalized as follows: we predict that participants in the informative default treatment choose the default option on the subsetted data more often than the participants in the non-informative default treatment.
            3. Regressing the frequency of choosing the option that has a default on an informative_treatment dummy, the treatment coefficient is significantly different from 0.
            4. [On subsetted data:] Chosen_option_has_default ~ $\beta 1$ + $\beta 2$ default_informative + error
            5. Test: $\beta 2 > 0$, t-test.

Excluded observations:

If there are missing observations or unfinished experiments due to e.g. computer, server, or network errors, the affected observations will be removed from analysis. As mentioned above, we will remove all choices of the worst ranked option in the informative and uninformative decoy treatment. In Prolific, participants are given a maximum time to complete the experiment, otherwise they can’t submit their answer. We will exclude the data from participants who exceed the maximum time. No other participants will be excluded.

Significance levels

All tests will be performed at a 95% confidence level.

Power and randomization procedure

We will collect at least 600 samples in four treatment cells, at least 150 subjects per treatment. Since we are only able to use a subset of our data from our informative treatments to test our hypotheses (see hypotheses section)we will be able to use 67% of the informative decoy decisions and 25% of informative default decisions.

Each subject will perform their assigned task 40 times (each round will be called a trial). So, restricting our achieved sample to the decision that can be used to test our hypotheses we have:

- **Informative decoy: 150 * 40 * ⅔ = 4000**
- **Non-informative decoy: 150 * 40 = 6000**
- **Informative default: 150 *40 * ¼ = 1500**
- **Non-informative default: 150 * 40 = 6000**

To determine the power of our experiment a priori we use simulated data. We generate our data by creating choice problems as in our actual experiment and then using a softmax operator to determine which choice is chosen by the simulated participant. We use a lambda of 0.08 in our softmax operation. We operationalize our cue effect as a perceived increase in value of an option prior to the softmax operation. The option values (OV) are sampled from OV ~ N(100,15), the cue effect (TE) in the treatment group is sampled from TE ~N (10,10) and we assume a smaller cue effect (CE) in the control group, CE ~ N(5,10). Based on these assumptions, we sample 6000 participants from which we draw 150 participants per cell to create 1000 different datasets on which we run our main hypotheses tests. According to our simulation, our experiment has 83% power to find an effect at the 5% level in the default treatment and 91% power in the decoy treatment. In other words, our experiment seems well powered to find small cue effects.

Final Questions

We attest that no data have been collected at the time of writing.

1.A.2. Pre-analysis plan uploaded on OSF - Experiment 2

Follow up experiment to "Cued Decision Making: Pre-Registration Document for Experiment Data Analysis"

Authors: Gustav Karreskog, Benjamin Mandl

Introduction:

This document describes a follow up experiment to the original study called "Cued Decision Making: Pre-Registration Document for Experiment Data Analysis" available at https://osf.io/murjf/ . The relevant changes to the first experiment are underlined.

For the new experiment, participants are recruited on Prolific to participate in an experiment on decision making. Subjects will be randomized into experiencing one of two cue types: **decoy and default.**

Experimental procedure:

Participants are recruited on Prolific. Then participants are randomized into one of two cue types. Only participants who speak English fluently and who have at least 10 prior submissions and an approval rate of 95% are recruited. Participants are restricted to desktop users (i.e. no mobile or tablets users). The experiment is coded in OTree. The procedure is as follows:

- Participants receive the same instructions regardless of treatment. They are given information on the task and their goals, and an example of the decision task.
- The Participants' task in the experiment is to select one out of three **options**. Each option is a combination of a solid area of different shapes and a price (see Figure 1 below). Participants are told that their payoff consists of the *Area* of the option's shape minus the option's *Price*. Participants receive the option value as a payoff in **experimental units**. They repeat the task **40 times** and need to consider each problem for at least 5 seconds, i.e. participants cannot submit an answer within the first 5 seconds of starting a decision problem. Participants receive feedback after each round (see Figure 2). The experimental units are transformed into GBP at the end of the experiment.
- Each participant is assigned a hidden random real number between 0 and 1 that describes the probability of receiving informative cues (called "assigned informativeness") throughout the experiment. The participants do not observe their assigned informativeness. We call this the *Individual Treatment Level.*
- In 28 of the rounds, the cue is randomly generated according to the treatment level. So in these 28 rounds, participants experience the informativeness of the cues differently. In the remaining 12 rounds, the cue is on average uninformative for all participants and identically generated for all participants with the same cue type. We will use these 12 *Comparison Rounds* to test our hypotheses.

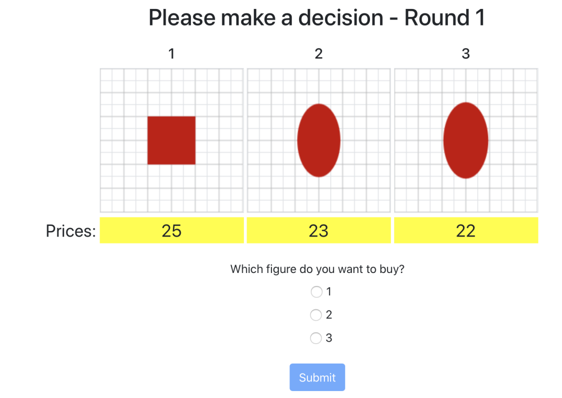


Figure 1: Decision table: Participants can choose one of three options which differ by area and price .

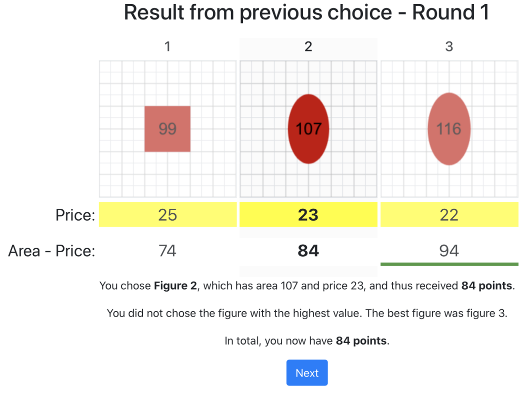


Figure 2: Decision feedback: Participants receive feedback on the bonus points for each option, their total bonus points and if they chose the figure with the highest value (green line)

- The Options' *Areas and Prices* are drawn randomly in each round and each participant. First, three **option values** are drawn from the same normal distribution and then ranked. *Area and Price* are then determined post hoc to equal the previously drawn values. The *Area and Price* are determined in special ways to generate the different treatments the participants were assigned to, which is described below.
- Participants are able to practice the task for three rounds. These rounds are not payoff relevant and are randomly generated according to the individual assigned informativeness.
- Following the practice rounds, participants need to correctly answer 4 Quiz questions to proceed. In case they answer incorrectly more than 10 times, they are unable to proceed with the study and need to exit.
- Participants are then randomized into one of two treatments which affects how the option features are determined:
    - **Decoy:**
        - For 28 out of 40 rounds, the options features are determined as follows:
            - With probability p equal to the assigned informativeness, the best option is chosen to have a decoy. Then one of the remaining options (either the second or the third best) is chosen with 50% probability to be the best option's decoy. We call the best option and the decoy option a *Comparable Pair*. *Comparable Pairs* have the same *Shape*. One of the options in the *Comparable Pairs* either has a larger *Area*, a lower *Price*, or both, i.e. it is dominating the other option in the *Comparable Pair*. We call the dominated option in the randomly chosen pair *the decoy*.
            - With probability 1-p, the second best option is chosen to have a decoy. The worst option is the *decoy*, i.e. the second best and the worst option are a *Comparable Pair*.
        - For 12 out of 40 rounds, the option features are determined as follows:
            - First, either the best or the second best option is determined to have a decoy. This option is then paired with the lowest value option which is the decoy. So either the best or the second best option and the worst option are a *Comparable Pair* with equal probability.
    - **Default:**
        - For 28 out of 40 rounds, the options features are determined as follows:
            - The option features are determined randomly. The highest value option is pre-selected with a probability p equal to the assigned informativeness. In 1-p of cases, one of the other options is pre-selected with a uniform probability over the two non-best options.
        - For 12 out of 40 rounds, the options features are determined as follows:
            - The option features are determined randomly. One of the options is pre-selected by default with a uniform probability over the three options.
- After finishing 40 rounds participants exit the survey.

Standard error corrections

T-tests will assume unequal variances and all regressions will use heteroskedasticity robust standard errors. T-tests are always two sided. We test at a significance level of 5%. Regressions are OLS unless otherwise noted. Regressions will cluster standard errors on the participant level because we expect to see heterogeneity in the treatment effects across individuals.

List of variables for regressions:

The following variables will be used in our regressions described below:

- Chosen_option: equals 1 if the option was chosen by the participant. 0 otherwise.
- Option_with_decoy_chosen; equals 1 if the option that has a decoy was chosen by the participant. 0 otherwise.
- Option_is_default_chosen; equals 1 if the option that is the default choice was chosen. 0 otherwise.
- Treated_round: equals 1 the cue was generated using the individual treatment p, and zero otherwise.
- Realized_informativeness; equals 1 if the best option has a decoy/the best option has a default.
- Cumulative_informativeness; share of the previous rounds that had realized_informativeness (r is rounds from 1 to 43, rounds 1 to 3 are practice rounds, rounds 4 to 43 are the experimental rounds):

$$\text{cumulative informativeness}_r = \sum_{i=1}^{r-1} \frac{\text{informativeness}_i}{r-1}$$

Main Hypothesis Tests:

Our main hypotheses establish that the realized informativeness predicts the propensity to follow the context effect in exactly comparable situations.

1. Hypothesis:
    a. *Participants who experience a higher realized informativeness follow the decoy more often.*
        i. Test of Hypothesis:

1. First, the data is subsetted to contain only the 12 comparison rounds where one of the two best options are paired randomly with the worst option to be a Comparable Pair.
2. Regressing the binary variable of whether the option with a decoy was chosen on the cumulative_informativeness, the treatment coefficient is significantly different from 0.
3. [On subsetted data:] Option_with_decoy_chosen ~ β1 + β2 cumulative_informativeness + error
4. Test: β2 > 0, t-test.

ii.

2. Hypothesis:
   a. *Participants who experience a higher realized informativeness follow the default more often.*
      i. Test of Hypothesis:
         1. First, the data of the participants in the **default** treatment is subsetted to contain only the 12 comparison rounds where one of the options is pre-selected with a uniform probability over the three options.
         2. Regressing the binary variable of whether the option is the default was chosen on the cumulative_informativeness, the treatment coefficient is significantly different from 0.
         3. [On subsetted data:] Option_is_default_chosen ~ β1 + β2 cumulative_informativeness + error
         4. Test: β2 > 0, t-test.

Excluded observations:

If there are missing observations or unfinished experiments due to e.g. computer, server, or network errors, the affected observations will be removed from analysis. In Prolific, participants are given a maximum time to complete the experiment, otherwise they can't submit their answer. We will exclude the data from participants who exceed the maximum time and from those who we identify as not using a computer. No other participants will be excluded.

Significance levels

All tests will be performed at a 95% confidence level.

Sample size

We will collect at least 600 samples in two cells, at least 300 subjects per cue type.

Each subject will perform their assigned task 40 times, 12 of which are directly comparable. Restricting our achieved sample to the decision that can be used to test our hypotheses we have the following number of observations per cue type:

- **<u>Decoy: 300 * 12 = 3600</u>**
- **<u>Default: 300 *12 = 3600</u>**

Final Questions

We attest that no data for this experiment have been collected at the time of writing.

1.B References

Altmann, S. et al. (2019). "Defaults and Donations: Evidence from a Field Experiment." In: *The Review of Economics and Statistics* 101.5, pp. 808–826.

Ariely, D. and T. S. Wallsten (1995). "Seeking Subjective Dominance in Multidimensional Space: An Explanation of the Asymmetric Dominance Effect." In: *Organizational Behavior and Human Decision Processes* 63.3, pp. 223–232.

Beshears, J., J. J. Choi, D. Laibson, and B. C. (C. Madrian (2010). *The Limitations of Defaults*. NBER. URL: `https://www.nber.org/programs-projects/projects-and-centers/retirement-and-disability-research-center/center-papers/rrc-nb10-02` (visited on 10/19/2021).

Bordalo, P., N. Gennaioli, and A. Shleifer (2013). "Salience and Consumer Choice." In: *Journal of Political Economy* 121.5, pp. 803–843.

Bronchetti, E., T. Dee, D. Huffman, and E. Magenheim (2011). "When a Nudge Isn't Enough: Defaults and Saving Among Low-Income Tax Filers." In: *National Tax Journal* 66.

Castillo, G. (2020). "The Attraction Effect and Its Explanations." In: *Games and Economic Behavior* 119, pp. 123–147.

Cronqvist, H., R. H. Thaler, and F. Yu (2018). "When Nudges Are Forever: Inertia in the Swedish Premium Pension Plan." In: *AEA Papers and Proceedings* 108, pp. 153–158.

Crosetto, P. and A. Gaudeul (2016). "A Monetary Measure of the Strength and Robustness of the Attraction Effect." In: *Economics Letters* 149, pp. 38–43.

Dietrich, F. and C. List (2016). "Reason-Based Choice and Context-Dependence: An Explanatory Framework." In: *Economics and Philosophy* 32.2, pp. 175–229.

Evangelidis, I., J. Levav, and I. Simonson (2018). "The Asymmetric Impact of Context on Advantaged versus Disadvantaged Options." In: *Journal of Marketing Research* 55.2, pp. 239–253.

Farmer, G. D., P. A. Warren, W. El-Deredy, and A. Howes (2017). "The Effect of Expected Value on Attraction Effect Preference Reversals: Effect of Expected Value on Attraction Effect." In: *Journal of Behavioral Decision Making* 30.4, pp. 785–793.

Frederick, S., L. Lee, and E. Baskin (2014). "The Limits of Attraction." In: *Journal of Marketing Research* 51.4, pp. 487–507.

Gerasimou, G. (2016). "Partially Dominant Choice." In: *Economic Theory* 61.1, pp. 127–145.

Gigerenzer, G. (2018). "The Bias Bias in Behavioral Economics." In: *Review of Behavioral Economics* 5.3-4, pp. 303–336.

Gomez, Y., V. Martínez-Molés, A. Urbano, and J. Vila (2016). "The Attraction Effect in Mid-Involvement Categories: An Experimental Economics Approach." In: *Journal of Business Research* 69.11, pp. 5082–5088.

Hagmann, D., E. H. Ho, and G. Loewenstein (2019). "Nudging out Support for a Carbon Tax." In: *Nature Climate Change* 9 (June).

Halpern, D. (2015). *Inside the Nudge Unit: How Small Changes Can Make a Big Difference*. Random House. 402 pp.

Huber, J., J. W. Payne, and C. Puto (1982). "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis." In: *Journal of Consumer Research* 9.1, pp. 90–98.

Jachimowicz, J. M., S. Duncan, E. U. Weber, and E. J. Johnson (2019). "When and Why Defaults Influence Decisions: A Meta-Analysis of Default Effects." In: *Behavioural Public Policy* 3.2, pp. 159–186.

Kaptein, M. C., R. Van Emden, and D. Iannuzzi (2016). "Tracking the Decoy: Maximizing the Decoy Effect through Sequential Experimentation." In: *Palgrave Communications* 2.1, p. 16082.

Lichters, M., P. Bengart, M. Sarstedt, and B. Vogt (2017). "What Really Matters in Attraction Effect Research: When Choices Have Economic Consequences." In: *Marketing Letters* 28.1, pp. 127–138.

Löfgren, Å., P. Martinsson, M. Hennlock, and T. Sterner (2012). "Are Experienced People Affected by a Pre-Set Default Option—Results from a Field Experiment." In: *Journal of Environmental Economics and Management* 63.1, pp. 66–72.

Luce, R. D. (2012). *Individual Choice Behavior: A Theoretical Analysis*. Courier Corporation. 172 pp.

Müller, H., V. Schliwa, and S. Lehmann (2014). "Prize Decoys at Work — New Experimental Evidence for Asymmetric Dominance Effects in Choices on Prizes in Competitions." In: *International Journal of Research in Marketing* 31.4, pp. 457–460.

Natenzon, P. (2019). "Random Choice and Learning." In: *Journal of Political Economy* 127.1, pp. 419–457.

Padamwar, P. K., J. Dawra, and V. K. Kalakbandi (2019). "The Impact of Range Extension on the Attraction Effect." In: *Journal of Business Research*, S0148296319307830.

Palan, S. and C. Schitter (2018). "Prolific.Ac—A Subject Pool for Online Experiments." In: *Journal of Behavioral and Experimental Finance* 17, pp. 22–27.

Peer, E., L. Brandimarte, S. Samat, and A. Acquisti (2017). "Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research." In: *Journal of Experimental Social Psychology* 70, pp. 153–163.

Shafir, E., I. Simonson, and A. Tversky (1993). "Reason-Based Choice." In: *Cognition* 49.1-2, pp. 11–36.

Sunstein, C. R. (2017). "Nudges That Fail." In: *Behavioural Public Policy* 1.1, pp. 4–25.

Tannenbaum, D. and P. H. Ditto (2012). "Information Asymmetries in Default Options." In: p. 20.

Tannenbaum, D., C. R. Fox, and T. Rogers (2017). "On the Misplaced Politics of Behavioural Policy Interventions." In: *Nature Human Behaviour* 1.7, p. 0130.

Trueblood, J. S., S. D. Brown, A. Heathcote, and J. R. Busemeyer (2013). "Not Just for Consumers: Context Effects Are Fundamental to Decision Making." In: *Psychological Science* 24.6, pp. 901–908.

Tversky, A. and D. Kahneman (1991). "Loss Aversion in Riskless Choice: A Reference-Dependent Model*." In: *The Quarterly Journal of Economics* 106.4, pp. 1039–1061.

Tversky, A. and I. Simonson (1993). "Context-Dependent Preferences." In: *Management Science* 39.10, pp. 1179–1189.

Von Neumann, J. and O. Morgenstern (2007). *Theory of Games and Economic Behavior: 60th Anniversary Commemorative Edition*. Princeton University Press. 774 pp.

Yang, S. and M. Lynn (2014). "More Evidence Challenging the Robustness and Usefulness of the Attraction Effect." In: *Journal of Marketing Research* 51.4, pp. 508–513.

Chapter 2

Overestimation of information demand

Benjamin Mandl

Jimin Nam

Abstract

We show in an online experiment that people have an inaccurate estimate of the amount of information that a decision maker considers before making a final decision. Participants are randomized to be either decision makers in a fully incentivized experiment or to be predictors. Decision makers receive piecewise, free information about their task and can submit their decision after each piece of information. Predictors are incentivized to predict the number of pieces of information the decision makers consider before submitting their final choice. We find that predictors overpredict the demand of information of the decision makers by a significant margin.

We are thankful to Anna Dreber, Magnus Johannesson, Michael I. Norton, Joakim Semb, Felix Schafmeister, and Binnur Balkan for insightful comments, as well as to seminar participants at SSE, Harvard University, and the Vienna University of Economics and Business for helpful comments and suggestions. We gratefully acknowledge funding by the Jan Wallander and Tom Hedelius Foundation.

2.1. Introduction

In this paper we show that people overestimate the amount of information that decision makers consider before making a decision in an incentive compatible online experiment. There are many situations in which one has to estimate the amount of information that is required before making a decision: an academic might need to estimate how many analyses or how much data she might ned to convince a referee, a sales representative might need to decide on how many arguments for why her product is great she wants to emphasize before a pitch, or a decision maker herself might need to decide whether she requires more information to make an informed decision. In each of these situations, our results seem to indicate that the amount of information that is needed to make a final decision is less than estimated, even if all involved parties are incentivized to make the personally best decision.

In our experiment, participants are randomized to be decision makers or predictors. The former are given piecewise information on payoffs of two different assets and are asked to choose one of the two assets to get a bonus. The decision makers can look at up to 40 data points before making a decision. The predictors are asked to predict how many data points an average decision maker will look at before submitting their final choice.

We find that predictors overestimate the number of data points decision makers look at before making a final choice by a large margin, despite having tried out the task themselves and being incentivized to give an accurate prediction. Furthermore, the majority of decision makers make a decision before looking at all data points. This happens even though information is free, there is no way to speed up the experiment by submitting an earlier choice, and participants generally make better decisions when looking at more data points. This behavior may be irrational, as standard economic theory of information would predict that additional information would never have negative value: either it benefits the decision maker or it could be ignored after it is received Stigler (1961).

Our findings are thus in line with related research by Klein and O'Brien (2018) who find that people use less information than they think to make up their minds in multiple experiments. In these previous experiments, decisions are unincentivized and ask the subjects to report and estimate information demand in decisions that are difficult to have rational expectations of: for example, in one experiment, participants are asked to decide and predict preferences of a style of art, in another participants predict amounts of consuming novel foods. In our experiment, there is a clear, rational answer to the decision problem, and an obvious, monetary incentive for the predictors to get it right. Since we find that this overestimation effect still exists, we believe that it is important to further investigate the ramifications of overestimating the information demand of decision makers, for example in the field.

To the best of our knowledge, there has not been any prior research into testing the prediction accuracy of information demand or cost function in a fully incentive compatible experiment despite the obvious importance of being able to correctly predict one's opponent's beliefs and actions in economic decision making. The most related research pertains to the literature of general misprediction: a range of papers present evidence of systematic mispredictions of taste (Loewenstein and Adler, 1995), one's own future utility (Frey and Stutzer, 2014; Loewenstein, O'Donoghue, et al., 2003), one's own likelihood of going to the gym (Della Vigna and Malmendier, 2006) or one's own time preferences (Augenblick and Rabin, 2019).

In the domain of predicting others, prior research has for example demonstrated difficulties in predicting trade partners' endowment effects which leads to inefficient offers in an experimental goods market (Van Boven et al., 2003). Frederick (2012) shows that there is pluralistic ignorance about others' willingness to pay for a good, Kurt and Inman (2013) extend the research by showing that priming and high empathy for others can lead to smaller willingness to pay estimation gaps[1]. Another, related strand of literature has investigated the ability to predict others' preferences under risk. The results are mixed. Hsee and Weber (1997) show that participants predict more risk seeking behavior than themselves whereas Faro and Rottenstreich (2006) show that experiment participants anticipate that decision makers behave as described in prospect theory but predict a lower risk seeking than actually observed. In a more psychology relevant domain, research on predicting the behavior of others has shown that people tend to believe they know others better than others know them (Pronin et al., 2001). Our findings extend this field by adding that people also seem to believe that decision makers consider more information before making a decision than they actually do.

In addition to misprediction, our paper is related to research on exlaining variations in information demand. Most notably, almost all of our decision makers made a decision before looking at all data points and our predictors anticipated that behavior. A potential explanation for this behavior could be non-Bayesian belief updating, which was identified by Ambuehl and Li (2018) as an explanation for overvaluing low quality information and undervaluing high quality information in their experimental subjects. A follow up to our work could investigate how the subjects update their beliefs throughout the experiment to see whether non-Bayesian updating could account for relatively low information demand. Furthermore, a recent model of knowledge acquisition by Golman et al. (2021) (forthcoming) puts forth the idea that information demand (and avoidance) depends on the perceived importance of the decision, the salience, and anticipated beliefs valence. This model could be tested in our framework by randomizing the payoff differences for our decision makers in our experiment to see whether a larger payoff difference would increase information demand. However, Huber et al. (2008) show that, at least for decisions in an experimental financial market with a wide range of information levels,

[1] given the difficulty to replicate social priming experiments it is unclear whether this result may be robust to a replication attempt (Chivers, 2019)

additional information might not result in better decisions so it is unclear whether increasing the payoff would result in vastly different behavior. Lastly, research by Eliaz and Schotter (2010) shows that people are willing to pay for non-instrumental information to decrease uncertainty about getting bad news. This finding seems to predict a different behavior than what we observed, because in our experiment people could consider (and actually couldn't even avoid) free information to decrease their uncertainty, but made their decision earlier. Follow up research could be helpful in understanding what kind of non-instrumental information has a higher demand than free instrumental information.

The paper proceeds as follows: Section 2.2 describes the experimental design; Section 2.3 describes the results, which are discussed in Section 2.4. Section 2.5 concludes.

2.2. Experimental Design

In our experiment, participants are randomized into one of two conditions: decision maker or predictor. After giving informed consent, all participants were shown the same instructions until they had to submit a decision or a prediction. All participants were given general instructions about the decision task, hereafter called the *Asset Choice Game*. In this game, participants are asked to choose between two assets to determine their bonus for the experiment. An asset in this game is described by a random normal distribution. Payoffs from an asset are determined as random draws from the asset's distribution, and participants are told that one asset has a higher payoff on average. To decide which asset to choose, participants can look at random draws from each asset. We called these random draws *payoff pairs*. Participants can look at up to 40 payoff pairs one-after-one and can report their decision on which asset to choose after each draw. The asset with the higher average payoff also has a higher variance in payoffs which makes it unclear which of the two assets has a higher mean[2]. Lastly, the bonus payment to the participant is equal to the average of 10,000 draws of the chosen asset, minus a constant. The average payoff of the higher asset was set at USD 3 and the lower asset generated an average payoff of USD 1 if chosen.

Participants were asked to answer attention check questions regarding the rules and goals of the experiment to proceed and were dismissed from the study if they incorrectly answered a given question more than twice. At the end of the instructions to the task, participants were also given two rules to keep in mind: first, we asked them to report the very first point at which they made a decision. It was necessary to spell out this instruction to make sure that people report the moment when they feel that additional information would not change their decision. Second, participants had to look at all 40 pieces of information regardless of when they reported the moment. Thus, it was impossible to

[2]in about one third of the draws, the asset with the lower mean has a higher individual draw.

save time by answering early because after a final decision was made, all remaining payoff pairs needed to be viewed one-by-one anyway.

After passing another multiple-choice question to ensure that the rules are understood, all participants could practice the task. They received one payoff pair and experienced the exact prompts and setup of the task. Following the practice round, the participants in the decision maker condition began the task. The predictors however were asked to predict how many payoff pairs they *think* the average respondent will look at. They are incentivized by the binarized scoring rule (Hossain and Okui, 2013) where the true answer is the average number of payoff pairs the decision makers viewed before making a final decision. Under the binarized scoring rule, predictors participate in a lottery where the probability of winning a USD 2 prize is equal to the squared distance of the given answer to the true answer. The binarized scoring rule is incentive compatible, even under the assumption of risk neutrality and is becoming increasingly used for belief elicitation in economics. Participants are given the option to either learn more about the binarized scoring rule before submitting a guess or to submit a guess directly.

Lastly, all participants answer a short questionnaire. We asked participants to report their estimated or experienced difficulty of the task, whether the decison makers would have chosen the same asset if they would have looked at more payoff pairs, whether they believed our statement that the experiment could not be sped up by answering earlier, the probability that the decision makers chose the higher paying assets, as well as demographics (e.g. education, sex, age, income, and employment status).

We chose the design of this game in order to incentivize all participants with a game that is easy to understand and presumably easy to predict. In the most comparable experiment of the studies run by Klein and O'Brien (2018), participants had to decide whether they like a new style of art and predictors had to predict the number of paintings considered before the decision maker makes a final decision. In this setup, there is no rational number of paintings to look at before making a decision, making it very difficult for a rational agent to have proper expectations. Additionally, we removed all possible deception our design by first describing the game generally and then assigning roles. In Klein and O'Brien (2018) participants were first led to believe that they were going to take part in the game before being assigned the role of predictor. Lastly, decision makers could actually save a bit of time by deciding earlier in Klein and O'Brien (2018) because all paintings were shown in one page after the decision maker submitted their final decision. In our experiment, decision makers who decided earlier had to click and pass exactly as many pages to finish the experiment as participants who decided later. Due to these important changes, we believe that our design allowed us to run a precise experiment on testing whether there is systematic misprediction in information demand.

2.2.1. Hypotheses

We prespecified one main hypothesis to establish our result and a number of additional, secondary hypotheses to increase our understanding in the decision making process and potential confounders of our study. All reported hypotheses were prespecified in the pre-analysis plan that was pre-registered on OSF[3].

Main Hypothesis

Our main hypothesis establishes whether predictors predict that decision makers will look at more example payoffs than decision makers actually do before making a final decision in a two sided t-test. We retest the same hypothesis by regressing the number of draws that are predicted/considered on the treatment and a vector of control variables (e.g. education, sex, age, income, employment status) as a robustness check.

Secondary Hypotheses

After establishing our main results, we move to investigate hypotheses that may shed light on the mechanisms. We try to understand the game performance by testing the relationship between the number of draws the decision maker looks at and the probability of choosing the higher paying asset. We also want to understand whether decision makers think they should have looked at more data points in the exit survey. Regret at deciding too early could indicate overconfidence of decision makers during the game.

One reason for misprediction in our game could be an inaccurate estimate of the difficulty of the game. Since predictors get to experience the game during a short trial run before making a prediction, we expect the difficulty estimation to be accurate. We also expected that predictors not only overestimate the number of draws that a decision maker looks at but also the probability of getting the higher paying asset, because the predictor might overestimate the demand and the value of the information given to the decision makers.

Our last prespecified test makes sure that there are no significant differences in trust in our instructions between the two treatments. A different level of trust in instructions could also give rise to the overprediction: for example, if decision makers do not believe our statement that they will need to look at all data points regardless of when they submit their decision but predictors do believe that statement, it would be natural for the predictions to be inaccurate.

2.2.2. Experimental procedures

We investigate our research question in an online experiment with 440 participants on Prolific Academic, a popular online social science laboratory (see Palan and Schitter (2018)

[3]see https://osf.io/j5sqv

and Peer et al. (2017) for reviews of the prolific platform and the participants) We ran our experiment in April 2021 and paid a show up fee of £1.1 and an average bonus of £2.21 to our participants. After accepting to take the study on Prolific's platform, participants were redirected to Qualtrics, where the study was coded. We recruited fluent English speaking, experienced US residents to participate in an experiment on decision making. The average time to complete the experiment was 10.3 minutes. 428 participants completed the experiment. As preregistered, before any analysis is run, we remove 17 incomplete (at least one covariate missing) surveys as well as one survey by a non-binary participant to easily control for gender using a binary indicator. These removals resulted in a sample size of 411.

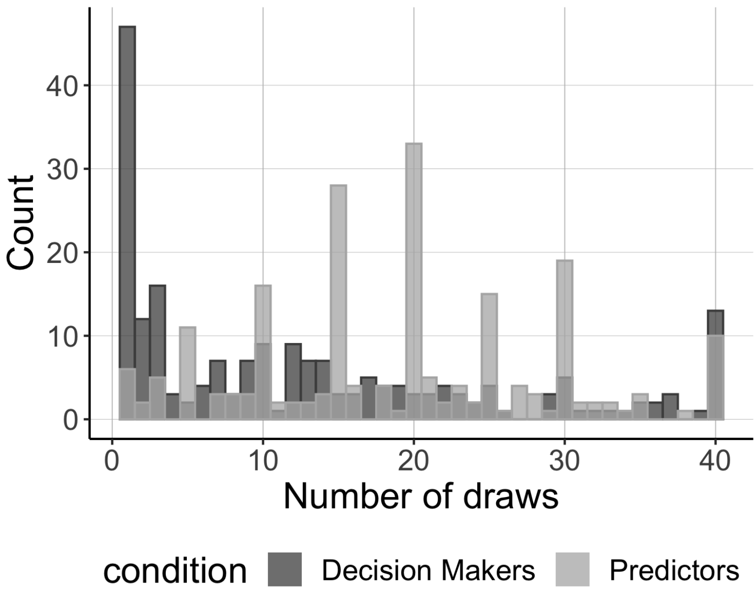
2.3. Results

In our analysis we closely follow our pre-registered analysis plan. Any deviations are explicitly stated. We first present evidence that our incentivized predictors overpredict the average information demand of the decision makers in the Asset Choice Game. Then we go through our secondary hypotheses outlined above to argue that this forecast error seems to stem from a misprediction in difficulty of the decision task rather than misunderstanding of the rules or instructions of the game, or failing trust in the experimenter.

2.3.1. Main result

The average predictions of the number of payoff pairs considered before submitting a final decision is 19.08, which is significantly larger than the actual average number of payoff pairs looked at by decision makers (mean difference = 6.18, $t(390.94) = -5.63$ $p < 0.001$, independent samples T test). Looking at the histogram of predictions and decision makers' information demand shows an interesting pattern. Most predictions center around multiples of 5 with the mode prediction being the median of the possible payoff pairs, 20. In contrast, the majority of decision makers decided immediately after looking at the first payoff pair. The finding in conjunction with Figure 2.1 give some interesting insights: participants in our experiment have difficulties to accurately predict the average information demand of decision makers, and there is quite some heterogeneity in predictions. Also, very few gave the rational but naive prediction of 40. 79.13% choose the higher paying asset. Among those that only looked at one data point, 53.19% get it correct, giving a first indication that participants do get valuable information from looking at more payoff pairs but choose to submit an answer very early.

Figure 2.1. Histogram of number of draws considered/predicted



Our result is robust to the inclusion of control variables[4] in our regression analysis, see Table 2.1, Column (1). None of our controls are significant.

Our experiment seems adequately powered. Based on our standard error our testing significance level alpha of 5%, the minimum detectable effect size in our experiment is a difference in about 3 actual versus predicted considered payoff pairs, a relatively narrow difference considering the histogram above.

2.3.2. Secondary results

In our secondary analysis, we try to understand more about why there is systematic misprediction of information demand. We first ask whether decision makers do better if they look at more information (i.e., whether they look at unreasonably few payoff pairs). We regress the binary indicator for whether the chosen option was the higher paying asset on the number of draws considered, using the same control variables as in the earlier robustness regression. We find that for each additional payoff pair, a decision maker has a significant increase in the probability of choosing the higher paying asset of 0.63% (p =

[4] We regress on the following covariates: high education (defined as having a 4 year degree or more), sex, age, high income (defined as above median), and employed (defined as being either part time or full time employed)

Table 2.1. Main and secondary regressions

	Dependent variable:		
	Number of draws considered	Verdict was optimal	Regret
	(1)	(2)	(3)
Constant	9.608***	0.766***	5.295***
	(2.510)	(0.125)	(0.560)
Condition = predictor	6.033***		
	(1.125)		
Number of draws considered		0.006*	0.030**
		(0.002)	(0.010)
Age	0.054	−0.002	0.006
	(0.045)	(0.002)	(0.010)
High Education	−0.554	−0.096	0.003
	(1.198)	(0.058)	(0.278)
Sex = female	0.701	−0.0005	−0.290
	(1.134)	(0.056)	(0.258)
High Income	0.268	−0.064	0.254
	(1.164)	(0.062)	(0.267)
Employed	0.926	0.139*	−0.161
	(1.277)	(0.068)	(0.301)
Observations	411	206	206
R^2	0.077	0.069	0.056
Adjusted R^2	0.064	0.041	0.028
Residual Std. Error	11.152 (df = 404)	0.399 (df = 199)	1.779 (df = 199)
F Statistic	5.653*** (df = 6; 404)	2.464* (df = 6; 199)	1.971 (df = 6; 199)

Notes: Heteroskedasticity robust standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

0.012), see Table 2.1, column 2. This indicates that decision makers who look at more data points choose the higher paying asset more often.

We examine whether decision makers regret their choice more often when they look at fewer payoff pairs, or whether they realize that they made a suboptimal choice due to deciding too early. We regress a measure for whether the decision maker would choose the same asset if they had waited to look at all data points before making a decision, captured on a 7-point Likert scale, on the number of draws considered and our vector of controls. We find a significant effect of looking at more data points and being content with one's choice ($p = 0.003$). In other words, participants who looked at fewer data points more often regretted their choice. The resulting regressions are summarized in Table 2.1, Column (2) and Column (3), respectively.

We ask whether the reason for the misprediction is that there is a difference in estimated and perceived difficulty in the task, in other words, we test whether predictors can accurately predict the game's difficulty. An overestimation of the difficulty is a potential explanation for our observed effect, as it seems intuitively rational to expect a higher information demand for more difficult decisions. We try to answer this question in two ways. We compare the reported difficulty scales directly. We do not find a significant difference between estimated and reported difficulty in our 7-point Likert scale survey question (mean difference = -0.02, $t(409) = -0.13$ $p = 0.9$). Second, we compare the expected probability of getting the optimal asset with the actual probability of getting the optimal asset. This alternative measure of expected difficulty gives in interesting insight: predictors underestimate the probability of getting the higher paying asset (79.13% of decision makers choose the higher paying asset and predictors estimate that 64.65% are able to choose the higher paying asset, mean difference = -14.48, $t(204) = -11.02$ $p < 0.001$). If predictors believe the task to be more difficult than it really is, they should naturally predict a higher number of draws considered before making a decision.

We ran an additional, not preregistered regression to test whether there is a statistically significant relationship by regressing the predicted probability of getting the correct answer on the number of draws considered, adjusted for heteroskedastic standard errors. We do not find a statistically significant relationship between these two predictions ($p = 0.32$). Since we cannot reject it is still unclear if a wrong estimate of the difficulty of the task may explain the misprediction in information demand. Further research could try to investigate a potential link between estimated difficulty and information demand.

Lastly, we worried that one reason for decision makers to give earlier answers might be because they did not trust our statement that decision makers cannot save time when answering early. Treating the Likert scale questions on trust as continuous, we cannot reject a t-test of difference in mean trust levels (mean difference = 0.24, $t(408.99) = 1.69$ $p = 0.09$). Additionally, we test whether there is a significant difference in whether the decision makers at the end of the experiment report that they actually did make a decision when their mind was made up and the predictors' expectations of whether they think that the decision makers actually followed the goal. 93.7% of all decision makers said they stopped looking at additional data when their mind was made up, significantly more than expected by the predictors, 66.8% ($p < 0.001$). Those decision makers who answered no to the question were asked to give an explanation. Excerpts of the reasons that were given are included in the appendix and typically explain their decision making process (e.g. trying to minimize uncertainty through looking at more pairs). It seems that those who answered no misunderstood the question since their given reasons point towards the fact that they actually made up their minds. However, it is surprising that only two thirds of predictors expected decision makers to submit a decision when their mind was made up. Whether this discrepancy has significant effects on our result should be investigated in future research.

2.4. Discussion

Given our main result, it seems as if people overestimate the information demand of decision makers by a considerable amount. This finding has wide ranging implications. There are many situations in which decision makers are given information for example by academics, salespeople, advisors, consultants, or other subject matter experts before choosing a product or strategy. Our research indicates that the amount of information that is required to make an incentivized decision may be lower than anticipated by the information providers. In other words, there might often be an inefficient information provision situation, where too many resources are spent on collecting and presenting information and arguments than necessary to make a decision. It is unclear why there might be such an inefficiency.

We tested one potential explanation for the existence of our effect. Predictors might have a wrong understanding of the difficulty of the task. Our mixed results show that when rated on a 7-point Likert scale, predictors and decision makers report the same level of difficulty, but when measured by the comparing the estimated and the actual probability of choosing the higher paying asset, predictors overestimate the difficulty of the task. Despite these mixed results, it is easy to imagine that predictors have inaccurate predictions about the difficulty of the task, but at the same time there is no immediately apparent reason for why predictors would overestimate the difficulty. Rather, given the research on overconfidence of own ability, one could reasonably expect an underestimate of the difficulty of the task. Thus, additional research is required to understand whether the overestimation of the difficulty of the task is systematic and whether our effect persists in spite of taking additional measures to make sure the difficulty estimates are accurate.

One final possible explanation for this inefficiency is that predictors have difficulties correctly estimating the cost of processing information. An outsider (as in our experiment) might be unable to imagine the actual cost or might be overconfident in her own ability to bear the information processing costs. An information provider (such as an advisor) might have difficulties imagining how hard it can be to process complicated information that is very familiar to her. We did not test these questions in the present experiment but investigating this path in future research seems promising.

2.5. Conclusion

We present results of an inquiry into the question whether people mispredict the information demand before making a financial decision in an incentive compatible online experiment. We find that there is systematic overprediction of the amount of information that needs to be considered to make a decision, despite being familiar with the task.

This systematic misprediction implies many interesting potential avenues for research. First, we think it would be important to validate these findings in the field. Potential areas of research could be sales or financial advisory situations, where sellers may overpredict the amount of information that is required to make a sale. Many similar situations exist, and a potential study could investigate when and where these situations result in inefficient information provision. Second, we think that the mechanism of misprediction is still not fully understood and requires more work. Theoretical modelling of the prediction in conjunction with lab experiments could help shine a light on why there is a bias in predicting information demand and how subjects could be debiased.

2.A Appendix

2.A.1. Selected reasons for answering no on whether the decision maker submitted their choice when their mind was made up

- Option B was the best choice in most of the rounds.
- I felt like I had enough information
- Probabilty[5] stopping theory suggests that the optimal stopping time is 37%, which is about 14/15 out of 40. By number 12, there were 10 A's and 1 B.
- I wanted to see more pairs to help me be more assured about my decision.
- I made my final decision when I did because I suspected the asset had a good consistent payoff throughout.
- I finally saw enough of the results to estimate how volatile and beneficial the returns on Asset A were.

[5] sic

2.A.2. Pre-analysis plan uploaded on OSF

Asset Choice Game Pre Analysis Plan

Authors: Benjamin Mandl, Jimin Nam

Date: April 16th 2021, 13:00

Introduction:

This project aims to explore whether people overpredict how much information is incorporated into making a decision. In recent research by Klein and O'Brien (2018), participants fail to anticipate how quickly their minds change in multiple, unincentivized experiments, and believe that they (as well as others) will evaluate more evidence before making up their minds than they actually do. We want to test, whether people do also overpredict the consumption of information in a fully incentive compatible research design.

In the study by Klein and O'Brien, both the actual decisions and the overprediction by the predictors in the experiment don't necessarily have detrimental consequences. In our design, participants are recruited online on Prolific and are randomized into two treatments, decision makers and predictors. Both learn that the decision makers' task is to decide to choose one of two different assets that have different random distributions: one asset has a higher mean and a wider variance, and the other has a lower mean and a lower variance. Decision makers receive 40 individual draws of both assets one-by-one. After each draw, the decision maker can make a final decision to choose one of the two assets. Thus, the decision makers' goal is to learn which of the two distributions has a higher mean by looking at many draws. If decision makers decide to choose one asset before looking at all 40 draws, they will still need to view all the remaining draws one-by-one. The decision makers and the predictors are being told that the decision makers cannot save time by submitting a choice earlier. The decision maker's payoff equals the average of many draws of the chosen asset. That means that the decision maker's incentivized task is to determine which asset has the higher mean through looking at individual draws of both assets.

Predictors receive the same information but are tasked to accurately predict how many random draws decision makers will look at before making a decision. Predictors and decision makers will complete one practice trial in full so they have experience in the task. They are incentivized by the binarized scoring rule. We predict that predictors will overpredict how many draws decision makers will look at before making a decision. We hypothesize that the underlying explanation for this misprediction stems from an inability to correctly estimate the decision makers cost function or perceived cost of incorporating information.

To begin, the study participants will read the consent form prior to partaking in the survey. Following the exercise, participants will be asked to answer binary (yes/no) questions about the exercise's instructions and a 7-point scale measure of difficulty with choosing an asset. Participants will conclude by answering a block of demographic questions and feedback on the exercise. The survey will be conducted through the Qualtrics platform and is estimated to take about 10 minutes to complete. Since the task view a series of asset pairs there is no more than minimal risk. There will be no direct identifiers and the Prolific ids will be removed upon downloading the files from Qualtrics, as we do not need it outside the role of attaching participants to their responses and paying them the correct bonuses. We will not collect IP addresses.

To our knowledge, testing the prediction abilities of a decision making task in the way described above has not been explored in economics. However, this research is related to the literature of misprediction: a range of papers present evidence of systematic mispredictions of taste (Loewenstein and Adler 1995), one's own future utility (Loewenstein, O'Donoghue, and Rabin 2003; Frey and

Stutzer 2014), and one's own time preferences (Augenblick and Rabin 2019). In the domain of predicting the utility of others specifically, prior research has focused on endowment effects and the willingness to pay of others. Van Boven, Loewenstein, and Dunning (2003) show difficulties of predicting trade partners' endowment effects, Frederick (2012) shows that there is pluralistic ignorance about others' willingness to pay for a good, Kurt and Inman (2013) extend the research by showing that priming and high empathy for others can lead to smaller willingness to pay estimation gaps. We are not aware of any research on predicting others' decision making abilities, information demand, or cost functions despite the obvious importance of being able to correctly predict one's opponents beliefs and actions in economic decision making. Our planned project aims to fill this gap.

Experimental procedure:

Participants are recruited on Prolific to participate in an experiment on decision making. Subjects will be randomized into one of two treatment cells: **decision maker** or **predictor.**
The experiment is coded in Qualtrics. The procedure is as follows:

- We recruit US residents that are fluent in English and have at least 50 previous assignments and a 95% approval rating to ensure high quality responses.
- Participants are asked whether they give informed consent to participating in this study.
- Participants receive the same instructions regardless of treatment. They are given information on the task and their goals.
- At the start of the study, all participants need to correctly answer 4 quiz questions to be able to continue the study. In the case that they answer incorrectly more than 2 times, they will be ineligible to proceed.
- The decision makers' task in the experiment is to look at example payoff pairs for two assets with different random distributions. One asset has a higher mean and a higher variance, and one a lower mean and a lower variance. They can look at up to 40 random draws per asset one by one. After each draw, they are asked whether they would like to choose one of two assets to be payoff relevant. Their payoff is equal to 10,000 draws of the asset the decision makers chose, divided by 10,000 and subtracted by 100. If they choose the high mean asset, the payoff will be roughly USD 3, and a choice of the low mean asset leads to a payoff of roughly USD 1. The main outcome variable for this treatment is how many payoff draws the decision makers looked at before making a decision.

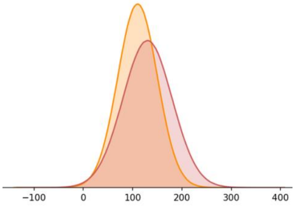


Figure: Asset distributions: The orange shaded distributions' mean is lower than the red shaded distributions mean (not shown to the participants)

- The predictors' task is to imagine being in the position of the decision maker and predict how many rounds of payoffs a decision maker will look at before making a final decision. They are fully informed about the decision makers' task and payoffs. Predictors can guess any number between 1 and 40.
    - Predictors are incentivized by the binarized scoring rule (Hossain and Okui, 2013). Under this scoring rule, predictors predictions generate a score which determines the probability of winning a payoff lottery.
        - In this experiment, the payoff of the lottery is USD 2.
        - The probability of winning the lottery is equal to the following: (1 – (1/1600) * (true answer – given answer)^2), where true answer is the mean number of

rounds of payoffs all decision makers looked at before making a decision (rounded to integers)) and given answer is an integer between 1 and 40.

    - In other words, if a participant guesses 16 and the true answer is 3, their probability of getting USD 2 is (1 – (1/1600) * (3 – 16)^2) = 0.89. The true answer is defined as the average of the number of draws looked at the by decision makers before making a decision (see also list of variables for regressions below). We divide by 1/(40^2) as this converts predictions to the percentage of the maximum of draws a decision maker could look at (a guess of 16 would mean that the decision maker looks at 16/40 = 40% of all data points).

- In the end, participants answer questions on demographics, Likert scale questions about the difficulty of the task, and the quality of the instructions with relation to the task.

All study materials will be uploaded to OSF with the final pre-analysis plan before the start of the data collection.

Standard error corrections

T-tests will assume unequal variances and all regressions will use heteroskedasticity robust standard errors. Regressions are all ordinary least squares. All tests are always two-sided. We test at a significance level of 5%.

List of variables for regressions:

The following variables will be used in our regressions described below:

- Number of draws considered: how many random draws of the assets are looked at before the participants makes a final decision. E.g., the number of draws considered is 3 if a decision maker looks at 3 asset payoff draws, the number of draws considered is 4 if a predictor predicts decision makers to look at 4 asset draws before making a decision.
- Verdict: whether asset A or asset B was chosen.
- Verdict optimal: whether asset A was chosen when A was optimal (or B if B was optimal)
- Difficulty: 7-point likert scale question describing the whether the decision maker experienced difficulty deciding between two assets and whether a predictor expected to experience difficulty deciding between two assets.
- Regret: 7-point likert scale question describing whether the decision maker would have chosen the other asset had she waited to look at all 40 example payoff pairs.
- Trust: 7 point likert scale question:
    - Decision makers: Did you believe our statement in the experiment description that said that you cannot speed up the experiment by choosing an asset early?
    - Predictors: Did you believe that the decision makers believed our statement in the experiment description that said that the experiment cannot be sped up by choosing an asset early?
- Compliance:
    - Decision makers: Did you make your asset decision at the first point at which you made up your mind? [yes/no]
    - Predictors: Do you think the decision maker made the asset decision at the first point at which the mind was made up? [yes/no]
- Explanation:
    - Decision makers: Why did you make the final decision when you did? (free answer)
    - Predictors: Why do you think the decision maker made the final decision when she/he did? (free answer)

- Prediction of getting it right:
    - Predictors: What is the probability that decision makers chose the higher paying asset? [0-100]
- Demographics (participants choose one of each):
    - Education: Less than high school, High school graduate, Some college, 2 year degree, 4 year degree, professional degree, doctorate
        - highEducation: binary indicator variable that is equal to 1 for all participants with 4 year degree or more
    - Biological sex: male, female, rather not say
        - sex: male, female ("rather not say" will be excluded from the dataset)
    - Age: numerical 18-99
    - Income: bracketed in 10k, from 0-10k to 90-99k, 100k to 149k, 150k+
        - highIncome: binary indicator variable that is equal to 1 for those with above median income.
    - Employment: Employed part-time, employed Full-time, Unemployed, Not in the labor force, other (free answer)
        - Employed: employed part-time and employed Full-time

Likert scale questions will be converted to the numbers 1-7 and treated as continuous for hypotheses testing.

Primary/Main Hypothesis and Tests:

Our main hypotheses establish that predictors predict that decision makers will look at more data points before making a decision than decision makers actually do.

1. Hypothesis – main test:
    a. *Predictors predict that decision makers will look at more example payoffs than decision makers actually do before making a final decision*
        i. Test: two sided sample t-test of predicted vs actual number of draws considered, predicted > actual.

Secondary Hypotheses and Tests:

1. Robustness test – OLS regression with control variables:
    a. *Predictors predict that decision makers will look at more example payoffs than decision makers actually do before making a final decision - regression*
        i. We run the following OLS regression:
        ii. Number of draws considered = a0 + a1Predictors + a2'X + error
        iii. X is vector of control variables (highEducation, sex, age, highIncome, employed)
        iv. Test: a1 > 0
2. Hypothesis – looking at more information helps makes better decisions:
    a. *Decision makers' probability of making the optimal verdict increases with numbers of data points looked at.*
        i. We run the following OLS regression:
        ii. Verdict optimal= a0 + a1Number of draws considered + a2'X + error
        iii. Test: a1 > 0
3. Hypothesis – looking at more information helps curtail regret:
    a. *Decision makers' regret decreases with numbers of data points looked at.*
        i. We run the following OLS regression:
        ii. Regret = a0 + a1Number of draws considered + a2'X + error
        iii. Test: a1 < 0
4. Hypothesis – predictors have an accurate estimate of the difficulty of the task:
    a. *Estimated and reported difficulty does not differ significantly*

        i. T-test of difficulty of the task reported by decision makers vs difficulty of the task estimated by the predictors is non significant.
5. Hypothesis – predictors are overconfident in decision makers getting it correct
    a. Predicted probability of getting the optimal asset is higher than actually getting the optimal asset:
        i. T-test predicted probability of getting the optimal asset vs probability of getting the optimal asset
        ii. Test whether the predicted probability is higher than the actual probability
6. Hypothesis – predictors and decision makers had similar levels of trust in the instructions:
    a. *Estimated and reported trust of veracity of instructions does not differ significantly*
        i. T-test of trust in instructions reported by decision makers vs trust in instructions by the predictors is non significant.

Excluded observations:

If there are missing observations/unfinished surveys (due to e.g., computer errors), the affected observations will be removed from analysis. If control variables are missing, the affected observations will be removed from all analyses (including the regressions and the t-tests). Participants who "rather not say" their gender will be removed from analysis to allow for a binary gender indicator in regressions. No other participants will be excluded for our main conclusions.

Significance levels

All tests will be performed at a 95% confidence level.

Power and randomization procedure

We aim to analyze at least 200 subjects per treatment. Given the very large effect size of d=1.45 in the article by Klein and O'Brien 2018 at a sample size of n=206, we recruit a larger sample to ensure high power.

We will collect 420 data points. Then we calculate the number of usable observations after excluding participants in line with the rules outlined above. Based on how many participants we need to remove we will recruit more participants until at least 200 usable observations per treatment are collected.

Additional reported statistics

In the paper, the MDE will also be reported in natural units, based on the observed standard error. For this, we will transform the minimum detectable effect size (MDE) by multiplying the standard error by a transformation constant of 2.8 for testing at $p<0.05$.

We will also report the number of participants who answered yes to the compliance question (Did you make your asset decision at the first point at which you made up your mind?/Predictors: Do you think the decision maker made the asset decision at the first point at which the mind was made up?) as well as report sample answers given in the corresponding free answer section.

Final statement

We attest that no data have been collected at the time of writing.

References

Augenblick, Ned, and Matthew Rabin. 2019. 'An Experiment on Time Preference and Misprediction in Unpleasant Tasks'. *The Review of Economic Studies* 86 (3): 941–75. https://doi.org/10.1093/restud/rdy019.

Frederick, Shane. 2012. 'Overestimating Others' Willingness to Pay'. *Journal of Consumer Research* 39 (1): 1–21. https://doi.org/10.1086/662060.

Frey, Bruno S., and Alois Stutzer. 2014. 'Economic Consequences of Mispredicting Utility'. *Journal of Happiness Studies* 15 (4): 937–56. https://doi.org/10.1007/s10902-013-9457-4.

Hossain, Tanjim, and Ryo Okui. n.d. 'The Binarized Scoring Rule', 28.

Klein, Nadav, and Ed O'Brien. 2018. 'People Use Less Information than They Think to Make up Their Minds'. *Proceedings of the National Academy of Sciences of the United States of America* 115 (52): 13222–27. https://doi.org/10.1073/pnas.1805327115.

Kurt, Didem, and J. Jeffrey Inman. 2013. 'Mispredicting Others' Valuations: Self-Other Difference in the Context of Endowment'. *Journal of Consumer Research* 40 (1): 78–89. https://doi.org/10.1086/668888.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin. 2003. 'Projection Bias in Predicting Future Utility'. *The Quarterly Journal of Economics* 118 (4): 1209–48. https://doi.org/10.1162/003355303322552784.

Van Boven, Leaf, George Loewenstein, and David Dunning. 2003. 'Mispredicting the Endowment Effect:: Underestimation of Owners' Selling Prices by Buyer's Agents'. *Journal of Economic Behavior & Organization* 51 (3): 351–65. https://doi.org/10.1016/S0167-2681(02)00150-6.

2.B References

Ambuehl, S. and S. Li (2018). "Belief Updating and the Demand for Information." In: *Games and Economic Behavior* 109, pp. 21–39.

Augenblick, N. and M. Rabin (2019). "An Experiment on Time Preference and Misprediction in Unpleasant Tasks." In: *The Review of Economic Studies* 86.3, pp. 941–975.

Chivers, T. (2019). "What's next for Psychology's Embattled Field of Social Priming." In: *Nature* 576.7786 (7786), pp. 200–202.

Della Vigna, S. and U. Malmendier (2006). "Paying Not to Go to the Gym." In: *American Economic Review* 96.3, pp. 694–719.

Eliaz, K. and A. Schotter (2010). "Paying for Confidence: An Experimental Study of the Demand for Non-Instrumental Information." In: *Games and Economic Behavior* 70.2, pp. 304–324.

Faro, D. and Y. Rottenstreich (2006). "Affect, Empathy, and Regressive Mispredictions of Others' Preferences Under Risk." In: *Management Science* 52.4, pp. 529–541.

Frederick, S. (2012). "Overestimating Others' Willingness to Pay." In: *Journal of Consumer Research* 39.1, pp. 1–21.

Frey, B. S. and A. Stutzer (2014). "Economic Consequences of Mispredicting Utility." In: *Journal of Happiness Studies* 15.4, pp. 937–956.

Golman, R., G. Loewenstein, A. Molnar, and S. Saccardo (2021). *The Demand for, and Avoidance of, Information*, p. 118.

Hossain, T. and R. Okui (2013). "The Binarized Scoring Rule." In: *The Review of Economic Studies* 80 (3 (284)), pp. 984–1001.

Hsee, C. K. and E. U. Weber (1997). "A Fundamental Prediction Error: Self-Others Discrepancies in Risk Preference." In: *Journal of Experimental Psychology* 126.1, pp. 45–53.

Huber, J., M. Kirchler, and M. Sutter (2008). "Is More Information Always Better?" In: *Journal of Economic Behavior & Organization* 65.1, pp. 86–104.

Klein, N. and E. O'Brien (2018). "People Use Less Information than They Think to Make up Their Minds." In: *Proceedings of the National Academy of Sciences of the United States of America* 115.52, pp. 13222–13227.

Kurt, D. and J. J. Inman (2013). "Mispredicting Others' Valuations: Self-Other Difference in the Context of Endowment." In: *Journal of Consumer Research* 40.1, pp. 78–89.

Loewenstein, G. and D. Adler (1995). "A Bias in the Prediction of Tastes." In: *The Economic Journal* 105, pp. 929–937.

Loewenstein, G., T. O'Donoghue, and M. Rabin (2003). "Projection Bias in Predicting Future Utility." In: *The Quarterly Journal of Economics* 118.4, pp. 1209–1248.

Palan, S. and C. Schitter (2018). "Prolific.Ac—A Subject Pool for Online Experiments." In: *Journal of Behavioral and Experimental Finance* 17, pp. 22–27.

Peer, E., L. Brandimarte, S. Samat, and A. Acquisti (2017). "Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research." In: *Journal of Experimental Social Psychology* 70, pp. 153–163.

Pronin, E., J. Kruger, K. Savitsky, and L. Ross (2001). "You Don't Know Me, But I Know You: The Illusion of Asymmetric Insight." In: *Journal of Personality and Social Psychology* 81.4, pp. 639–656.

Stigler, G. J. (1961). "The Economics of Information." In: *Journal of Political Economy* 69.3, pp. 213–225.

Van Boven, L., G. Loewenstein, and D. Dunning (2003). "Mispredicting the Endowment Effect:: Underestimation of Owners' Selling Prices by Buyer's Agents." In: *Journal of Economic Behavior & Organization* 51.3, pp. 351–365.

Chapter 3

Motivated beliefs and climate attitudes

Benjamin Mandl

Eva Ranehill

Abstract

We study whether motivated cognition causes the perceived benefits of climate friendly actions to rise as the associated costs decrease. In our experiment, respondents are offered to donate to plant a tree after being randomized to receive either a high or a low discount to the cost of donation. Before respondents make a final decision to donate or not, we elicit their perceived importance of planting trees, how much CO_2 they believe planted trees sequester from the atmosphere, and to what extent they agree with the statement that their actions contribute to climate change. We find that respondents randomized to a low discount – and hence a higher cost – state a lower perceived importance of planting trees and agreement with individual responsibility for climate change. Our overall results are mixed as we do not find statistically significant differences in the quantitative measure. Our study highlights the possibility that motivated cognition contributes to the slow response to climate challenges and how related policies may impact beliefs and attitudes in important ways.

We are greatly thankful to Fredrik Carlsson, Anna Dreber, Magnus Johannesson, and Roberto Weber for insightful comments, as well as to participants at several conferences and seminars for helpful comments and suggestions. We thank the Center for Collective Action Research (CeCAR) for generous financial support and the Jan Wallander and Tom Hedelius Foundation, and Knut and Alice Wallenberg Research Foundation for funding.

3.1. Introduction

During the last decades, scientific advances have enhanced our understanding of climate change, stressing the need to adjust our societies and lifestyles to become environmentally sustainable. Still, around the world, the transition to less carbon-intense production is slow and global greenhouse gas emissions are still increasing[1]. In this context, understanding determinants of individual climate related attitudes and beliefs to enhance public support of climate friendly policies appears paramount. In this paper, we explore whether motivated cognition is one mechanism contributing to a poor response to climate change.

There are several reasons why motivated cognition may be important with respect to climate related attitudes and behaviors. Previous research indicates that motivated reasoning is enhanced by "moral wriggle room" arising with the possibility (and incentives) to recruit and evaluate evidence in favor of desired beliefs (e.g., Festinger, 1962, Dana et al., 2007, Epley and Gilovich, 2016; Di Tella et al., 2015). Individuals who wish to maintain a positive self-image as climate friendly may be inclined to engage in motivated reasoning due to high incentives (e.g., the cost of no longer flying precluding visiting relatives living far away) and enabled by factual insecurity, availability of contradictory information and ambiguous risks[2].

To study the importance of motivated beliefs applied to climate attitudes, we implement an experiment exploring whether respondents' subjective beliefs about the importance of a climate friendly action depend on its cost. The experiment comprised about 2,500 respondents randomized to 5 conditions. In our two main conditions, LowDiscount and HighDiscount, respondents are informed that they will have the opportunity to donate to plant a tree at the end of the study. All respondents are informed of the price for planting a tree (USD 1), and that they will be randomized into one of two groups to receive either a low discount (5 cents) or a high discount (95 cents) to that price. Transparent instructions and control questions ensure that the large majority of participants are also aware of this information – preventing impeding that prices or subsidy sizes act as a relevant signal for donation importance. Additionally, we implemented two corresponding conditions with hypothetical decisions and a baseline condition simply measuring respondents' beliefs absent any donation decision.

Our measures of interests were elicited after participants had answered a couple of control questions, but before they made their donation decisions. We pre-registered two main outcome variables – participants' perception about the usefulness to plant trees to combat global warming, and their quantitative beliefs about the amount of CO_2

[1] see https://ourworldindata.org/co2-emissions?country=#year-on-year-change-in-global-co2-emissions, accessed August 13th 2021.

[2] While scientific evidence on climate change today is clear, the exact consequences and their timing is debated. Further, judging the total impact of, e.g., different consumption choices is difficult. For example, it is unclear whether replacing a gasoline car with an electric one is carbon positive if electricity is produced from fossil fuels.

sequestered by the planted tree from the atmosphere in a year[3]. Participants could submit an answer between 0 and 200 lbs and were informed that the average American's carbon footprint is about 100 lbs CO_2 per day. As a secondary measure, we asked participants to what degree they believed their actions contribute to global warming to assess perceptions of personal responsibility. Finally, at the end of the experiment, participants made their donation choice. We find that participants who face a higher cost of planting a tree assign less importance to planting trees to combat global warming and indicate lower agreement with the statement that their actions contribute to global warming compared to participants who face a lower cost. The effects are significant at the 1 and 5 percent level and are moderate in size – representing about 16% of a standard deviation. However, we do not find an impact of experimental condition on participants beliefs about the CO_2 sequestered from the atmosphere.

This paper contributes to the literatures on motivated cognition, climate attitudes and ethical consumption, showing that exogenous variation in the experienced costs associated with climate friendly behavior impacts stated beliefs both about the impact of that behavior and personal responsibility for climate change. The setting we generate also illustrates how motivated beliefs can be forward looking and adjusted to subsequent behavior. Our results are consistent with the predictions of Hestermann et al., 2020, who develop a theory to account for the impact of prices on cognitive dissonance in markets for goods with negative externalities (see also Di Tella et al., 2015 and Ging-jehli et al., 2019 for an application to settings with strategic interaction). According to their theory, higher prices decrease anticipated consumption which in turn mitigates the need to engage in self-deception. As a result, their theory predicts that motivated beliefs increase price elasticity.

Closely related to our paper is also Pace and van der Weele (2020). While the main research question explored in their experiment focuses on how uncertainty about the CO_2 emissions associated with a product impacts consumption demand, they also exogenously vary the price of the product in a design that shares features with ours. Contrary to the results presented here though, Pace and van der Weele (2020) find no impact of consumer surplus on emission beliefs. This could be a result of the different sample sizes – Pace and van der Weele (2020) have a considerably smaller sample size per experimental condition than we do, collecting 1000 participants randomized into 9 conditions and pooling some conditions for tests. Their power to find an effect of our size is 49.6%. Further research on this topic would be valuable[4].

Our results have important policy implications. If, on top of the impact of purely economic concerns, motivated reasoning causes individuals to downplay the value of

[3] Our pre-analysis plan is available at https://osf.io/gj8mr/.

[4] Another difference between the studies is that we differentiate between prices and costs, in order to avoid any effect of prices as perceived carriers of information.

political reforms and behavioral changes that are costly, the resulting focus on easy and low-cost measures to combat climate change may slow down necessary adjustment associated with larger benefits. However, the same mechanism also offers an optimistic perspective on the use of economic instruments, indicating that the resulting impact on demand associated with changes in prices may be amplified through motivated reasoning.

The remaining paper proceeds as follows. Section 3.2 discusses related literature. Section 3.3 describes our experiment. Section 3.4 presents the results which are discussed in Section 3.5. We conclude in in Section 3.6.

3.2. Related Literature

Our study contributes to several strands of research. Our findings add to the literature on behavioral environmental economics and in particular to the strand of literature exploring how behavioral biases pose threats to individuals' motivation to combat climate change (see, e.g., Shu and Bazerman, 2010). The papers most related to our experiment in this literature investigate the interplay between costs and pro-environmental beliefs and behavior. Hagmann et al. (2019) implement a nudging experiment and show that people are unwilling to face personal costs to combat climate change, preferring the free (and easy to ignore) nudge versus a more effective tax. Similarly, in a hypothetical choice experiment, Hedlin and Sunstein (2016) show that approval ratings of a green energy default depend on the price of the green option.

Our findings add to the growing literature on motivated cognition, as reviewed in Gino et al. (2016) and Bénabou and Tirole (2016). A range of experiments in this field show that people adjust their beliefs to, for example, justify selfishness in strategic games (e.g., Di Tella et al., 2015), or to protect beliefs about own ability (e.g., Zimmermann, 2020; Exley and Kessler, 2019; Grossman and Owens, 2012; Buser et al., 2016; Heger and Papageorge, 2018; Schwardmann and van der Weele, 2019; Chew et al., 2020), popularity (Eil and Rao, 2011; Gotthard-Real, 2017), or morality (e.g., Dana et al., 2007). We contribute to this literature by showing that motivated reasoning can occur in a forward-looking process, similar to the model proposed in Bénabou and Tirole (2016). In their model, signals about the payoff of a (pro-social) action in an initial period may be obfuscated to affect beliefs about the self in a later period, just as participants in our study show motivated beliefs before deciding to donate (see, also Hestermann et al., 2020 for a similar argument)[5].

Our paper also relates to the strand of empirical papers discussing the impact of prices on information acquisition or strategic ignorance (e.g. Grossman, 2014). Closest to our paper in this literature is perhaps Ambuehl (2017) who uses a strategy that shares

[5] Thus, our results suggest a counterbalancing force to observations crowding out. Contrary to observing monetary incentives crowd out, we find that participants who receive larger subsidies to the cost of donating donate more often and believe it to be more effective.

features with ours to investigate the impact of prices on information acquisition and belief formation in the context of aversive consumption decisions. He finds that incentives (sometimes rationally) skew information acquisition and indirectly impact demand.

In psychology and marketing, there is also a literature of the effects of identity, motivated beliefs, and self-deception on pro-environmental behavior. Examples in this field include Wade-Benzoni et al. (2007) who show an increased willingness in to act pro-environmentally when induced to identify as strong environmentalists. Farjam et al. (2019) show experimentally that the correlation between pro-environmental attitudes and behavior only hold in low cost situations, whereas there exists a gap between stated pro-environmental beliefs and behavior when costs are large. Our results indicate that this gap to some extent is endogenous, i.e., that stated beliefs itself depend on the cost.

3.3. Experiment Design

We explore our research question—whether the cost of an environmentally friendly action impacts its perceived benefit—in an online experiment. In the experiment, a total of 2,536 participants were recruited via Amazon Mechanical Turk (MTURK) and randomized to take part in 1 out of 5 experimental conditions. All conditions were collected simultaneously, and every participant had an equal probability of being randomized into one of the five conditions. Table 3.1 provides an overview of the number of participants in the different conditions[6].

Table 3.1. Donation decision by experimental condition

Condition	Observations	Donation frequency (%)
Baseline	464	-
Low Discount	488	32.2
High Discount	488	80.1
Low Discount Hypothetical	482	70.3
High Discount Hypothetical	499	91.6
Total	2,421	

Notes: In total, 2,536 participants took part in the survey. As pre-registered, 115 participants who reported having googled answers during the experiment were dropped prior to running our analysis.

[6] The experiment design and results sections follow the pre-analysis plan without deviations.

3.3.1. Experimental conditions

The 5 experimental conditions implemented comprise 2 main conditions and 3 secondary conditions. In this section, we first explain the survey structure for the two main conditions, and then explain how the secondary conditions differed. The full instructions for all conditions are available in the online appendix.

At the onset of the survey, participants in the two main conditions were informed that they would be given the opportunity to donate part of their payment to plant a tree at the end of the survey, and that the price for the donation was USD 1. The trees were to be planted via the organization One Tree Planted, which plants trees in US national forests. Having been informed about the subsequent donation decision and the price of the donation, participants were told that they would be randomly assigned to receive a subsidy of 5 (the LowDiscount condition) or 95 cents (the HighDiscount condition) to the price of USD 1.

The instructions thus explicitly explained the existence of the two conditions, and that participants would be randomized to the one or the other. Two attention checks controlled that all participants knew the price of planting a tree absent a subsidy, as well as the cost they would face (USD 0.05 or 0.95) when asked to donate at the end of the survey. These control questions were implemented to assure that all participants were aware of what experimental condition to which they had been randomized, as well as to avoid that participants failed to distinguish between the price of the donation and their individual cost. Failure to distinguish between the price and the individual cost may be problematic in our setting if the price (or cost) is presumed to carry information about aspects of the donation such as, for example, it's effectiveness to counteract climate change. Therefore, as pre-registered, the 13% of eligible participants who failed these attention checks 3 times had to exit the survey before submitting beliefs and donation decisions.

After the attention checks we elicited our two main measures of participants' beliefs about the effectiveness of planting trees to address climate change. The first measure asked participants to what extent they considered planting trees to be "an important activity to mitigate emissions of CO2 and combat global warming" and thus asked participants to make a qualitative statement about the effectiveness of a possible donation. The second measure asked participants to make a quantitative statement – asking how many lbs of Co2 participants estimated "a planted tree binds per year". As an additional, secondary outcome variable we measured attitudes to personal responsibility for climate change, asking participants to what extent they agreed with a statement that their actions contribute to global warming and climate change. Finally, participants were asked whether they wanted to donate to plant a tree and whether they wanted to sign up for a newsletter from OneTreePlanted.

The survey ended with a brief questionnaire collecting information about respondent age, gender, gross annual income, highest completed education, and political orientation. We also asked participants whether they googled to find an answer regarding the amount of carbon dioxide sequestered by a tree. (The survey questions as well as the generated variables are described in the pre-analysis plan and for convenience in Table A.1 in the Appendix.)

Two of the three secondary conditions were very similar to the main conditions. The main difference was that instead of real choices, the two secondary conditions asked participants for hypothetical answers. Participants in the HypLowDiscount and the HypHighDiscount conditions were asked to imagine having the opportunity to donate a tree with a subsidy of USD 0.05 or 0.95 respectively, and to answer as if the choice were real. In all other aspects the survey instructions were kept as similar as possible. The inclusion of the control conditions was made to further explore the possibility that participants interpret the costs associated with planting a tree as informative of the impact also absent an actual donation opportunity. A significant impact of the cost of the donation across the hypothetical conditions may indicate that, despite our efforts, individual costs are interpreted as informative about the impact of planting a tree on climate change.

Finally, we implemented a Baseline condition. In the Baseline condition participants were not asked to donate any money, and only answered the set of questions about their beliefs about the impact of planting trees. The Baseline condition was implemented to elicit beliefs about the impact of planting trees absent any donation decision or information about prices or costs.

3.3.2. Hypotheses

Based on the outcome variables elicited in the two main conditions we formulate the following main hypothesis:

Main hypothesis:

Participants who were randomly allocated to face a cost of planting a tree of USD 0.05 will state a higher importance of planting trees to combat global warming than those who face a price of USD 0.95.

We first use regression analysis to test this hypothesis for our two main outcome measures of qualitative and quantitative belief, and for our secondary measure of personal responsibility. We regress the main outcome variables on the treatment indicator for High Discount and the control variables we prespecified. Those are Age, Gender, High Education which is 1 if the participant has at least some college education, High Income which is income above the collected median as well as political orientation relative to being aligned with the Republican party. We then explore our secondary, more exploratory, hypotheses regarding the robustness of our results and implement a heterogeneity analysis. First, we run heterogeneity analyses to test for relationships between our control variables and the treatment effect as well as for their interactions. While we pre-registered the regression, we did not pre-register any directional predictions for these analyses.

3.3.3. Experiment procedures

The survey was implemented using Qualtrics and took about 3 minutes to fill in. Participants earned USD 0.5 as a show up fee and received an additional USD 1 if they correctly answered the 2 attention checks.

We collected a total of 2,536 observations. The sample size is based on the power calculation presented in our pre-analysis plan. Because we did not know the propensity of MTurkers to donate at different costs, our pre-analysis plan stipulated to collect 100 observations, randomized across the High- and the LowDiscount conditions, in a pilot. If the results from the pilot indicated a difference in donation rates between the two conditions of at least 30 percentage points, we would continue to collect 900 observations for our main conditions and 500 observations for our two hypothetical conditions and the baseline, and otherwise reconsider our experimental design[7]. The pilot data and the 900 observations subsequently collected were pooled for our analysis.

All observations were collected during the first week of December 2020 via MTURK. The average age of the respondents in this study is 40.3 years. 51.6% of our participants are female, and the median income is between $40,000 - $49,999 (11% of sample) and the median education level is a college degree (45% of sample).

3.4. Results

In this section, we first present evidence that our treatment variation of different discounts to donate to plant a tree results in large effects in donation frequency. Next, we describe our main results exploring whether the exogenous variation in donation costs influences participants' beliefs about climate impact of planting trees. Finally, we use our preregistered exploratory analysis to better understand the mechanisms underlying our findings.

[7] This was done to determine the impact of our main treatment on donation behavior, and to be able to conduct a preliminary power analysis. The pilot comprised 99 participants randomly assigned to the High- and LowDiscount conditions. In order to ensure enough statistical power, we pre-registered to continue the data collection only if donation rates differed by at least 30 percentage points between the two conditions (because prospective donation rates is what we predict will lead to motivated beliefs). The outcome of the pilot was a difference of about 58 percentage points (14 out of 53 in the HighDiscount group donated, while 39 out of 46 participants in the LowDiscount group donated), and we therefore proceeded to collect the remaining data. The pre-registration specified a pilot comprising 100 participants. The resulting 99 instead of 100 answers is due to some participants not finishing the survey.

3.4.1. The impact of cost on donation behavior

Table 1 tabulates donation rates for the different conditions, indicating that our treatment had the intended effect on donation rates. In the LowDiscount condition, only about 32 percent donated whereas more than 80 percent did so in the HighDiscount condition. The corresponding numbers for the hypothetical conditions are 70 and 96 percent respectively.

3.4.2. Main results

In this section, we test our main hypothesis, whether the cost of an action – and thereby the likelihood of engaging in it – impacts the perceived consequences associated with the action. An overview of the average stated beliefs and attitudes among the participants in the different conditions is presented in Table 3.2. This direct comparison of stated beliefs in the main conditions relative to those in Baseline indicates small differences between the Baseline and the LowDiscount condition and a tendency for somewhat more positive beliefs in the HighDiscount condition[8].

Table 3.2. Donation decision by experimental condition

Condition	Importance planting trees	CO_2 sequestered	Responsibility
Baseline	6.19	93.26	5.55
Low Discount	6.23	95.00	5.47
High Discount	6.37	94.76	5.66
Low Discount Hypothetical	6.23	94.13	5.37
High Discount Hypothetical	6.30	90.25	5.56

Notes: Participants qualitative beliefs and perceived responsibility were elicited on a scale ranging from 1 (completely disagree) to 7 (completely agree). Their qualitative beliefs about amount of sequestered CO_2 was elicited through a slider from 0-200lbs.

Our pre-specified main tests are presented in the first two columns of Table 3. They present OLS regressions of our two outcome main variables – participants' qualitative beliefs about the importance of planting trees and their quantitative estimate of the number of pounds of CO_2 that a planted tree binds per year – regressed on a dummy variable for experimental condition and a vector of individual controls (age, gender,

[8] P-values based on two.sided Mann-Whitney tests indicate that the comparison between beliefs in the Baseline and HighDiscount condition is significant for both the quantitative beliefs and feelings of responsibility (p-values = <0.001 and 0.038, respectively). For the quantitative beliefs also the difference between the Baseline and the hypothetical high discount condition was significant (p-value = 0.026). These analyses were not pre-registered. Since the baseline condition is absent all donation considerations, shorter, and paid less than the other conditions, it is unclear how well baseline beliefs and discount condition beliefs can be compared.

education, and income). As indicated in the first column, we find that participants who face a high discount agree with the statement that planting trees "is an important measure to combat global warming" to a higher degree than participants who face a low discount. The estimated impact corresponds to 0.15 of a step on the 7 degree scale, or 0.16 of a standard deviation, with a p-value of 0.010. We do not, however, find an impact of experimental condition on the amount of CO_2 participants estimate a tree sequesters from the atmosphere. Hence, we find only mixed support for our hypothesis.

Table 3.3. Impact of experimental condition on participants beliefs, OLS regressions

	(1) Importance planting trees	(2) CO_2 sequestered	(3) Personal Responsibility
Constant	6.055***	93.06***	5.370***
	(0.138)	(9.129)	(0.202)
High Discount	0.154**	0.390	0.199*
	(0.0592)	(3.958)	(0.0878)
Age	-0.000204	0.0993	-0.00760*
	(0.00260)	(0.165)	(0.00384)
Female	0.210***	14.55***	0.329***
	(0.0605)	(3.962)	(0.0896)
Gender (other)	0.120	15.34	0.368
	(0.225)	(18.48)	(0.370)
High Education	0.0557	-8.586	0.185
	(0.0931)	(6.805)	(0.144)
High Income	0.0445	-5.592	0.137
	(0.0614)	(4.109)	(0.0920)
Observations	976	976	976
R-squared	0.019	0.018	0.026

Notes: Heteroskedasticity robust standard errors in parenthesis.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Columns 4 and 5 repeat the analysis in the first two columns using instrumental variables regressions where the dependent variable is the beliefs about the usefulness to plant a tree, the donation decision is the endogenous variable, and experimental condition is the instrumental variable. All control variables are the same. This approach allows us to more precisely measure how a change in the cost impacts beliefs through the probability of a donation. The IV results are qualitatively similar to the OLS – a higher discount increases the perceived use of planting a tree using the qualitative (column 3) but not the quantitative measure (Column 4). Estimating the change in beliefs associated with a change in the probability of donating instead of across conditions, the estimated effect size is about twice as large. The IV analysis can be understood as a scaling of the effect by the treatment difference in the fraction that donated. With a difference of about 50 percentage in donation rates, we scale the treatment coefficients in the IV by about 2.

Interestingly, we also find a significant effect exploring our secondary measure of perception of personal responsibility for climate change. Estimating the same regressions also for this measure, we find a high discount associated with a stronger perception of personal responsibility both when estimating the effect using OLS and IV regression (see columns 3 and 6 in Table 3; β OLS = 0.199, p-value = 0.024; β IV = 0.408, p-value = 0.020).

3.4.3. Heterogeneity

In Table 4 we explore whether the environmental beliefs and attitudes elicited vary by the collected sociodemographic characteristics as outlined in our pre-analysis plan. No directions of the hypothesized effects were pre-registered. Women express a more positive attitude to planting trees and indicate a higher perception of individual responsibility than men do, but do not donate to a higher degree. Further, self-reported democrats express a more positive belief about the general usefulness of planting trees than republicans (the omitted category). They also express a higher perception of individual responsibility and donate to a larger degree. However, they do not state a higher belief in terms of the quantity CO_2 sequestered[9].

3.4.4. Hypothetical decisions

We also implemented two control conditions eliciting hypothetical decisions, and a baseline condition without any donation decision. The hypothetical conditions were kept as close to the main High- and LowDiscount conditions as possible, with the difference that instead of making their donation decision for real, participants were informed that their decision to donate would be recorded but would neither be carried out at the end of

[9] Table A.2 in the appendix presents the same analysis as in Table 4 interacting the sociodemographic variables with treatment. The only significant interaction in this analysis is a negative interaction between the gender category "Other" and treatment with respect to beliefs about the general usefulness of planting a tree. However, only 14 individuals in our sample identified as "Other" and this interaction is therefore based on a very small number of observations.

Table 3.4. Impact of experimental condition on participants beliefs, IV regressions

	(1) Importance planting trees	(2) CO_2 sequestered	(3) Personal Responsibility
Constant	6.041***	93.02***	5.352***
	(0.135)	(9.199)	(0.195)
High Discount	0.316**	0.802	0.408*
	(0.118)	(8.108)	(0.176)
Age	-0.00172	0.0955	-0.00955*
	(0.00260)	(0.167)	(0.00380)
Female	0.194**	14.51***	0.308***
	(0.0594)	(3.950)	(0.0881)
Gender (other)	0.0419	15.14	0.267
	(0.232)	(18.56)	(0.346)
High Education	0.0345	-8.640	0.158
	(0.0894)	(6.804)	(0.139)
High Income	0.0465	-5.587	0.140
	(0.0596)	(4.096)	(0.0896)
Observations	976	976	976
R-squared	0.075	0.018	0.072

Notes: Heteroskedasticity robust standard errors in parenthesis.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3.5. Environmental beliefs and sociodemographic characteristics

	(1) Importance planting trees	(2) CO_2 sequestered	(3) Personal Responsibility	(4) Donation
Age	0.000579	0.0973	-0.00568	0.00420**
	(0.00249)	(0.167)	(0.00352)	(0.00128)
Female	0.158**	15.47***	0.231**	0.0138
	(0.0588)	(3.971)	(0.0837)	(0.0319)
Gender (other)	0.0612	13.89	0.229	0.314
	(0.178)	(15.85)	(0.366)	(0.173)
High Education	-0.00866	-6.894	0.0508	0.0381
	(0.0905)	(6.784)	(0.142)	(0.0539)
High Income	0.0302	-4.686	0.115	-0.0236
	(0.0601)	(4.125)	(0.0859)	(0.0329)
Pol. Orientation (Prefer not to say)	-0.215	24.70*	-0.402	-0.217*
	(0.217)	(11.25)	(0.322)	(0.0894)
Independent	0.0374	0.459	0.212	0.00720
	(0.0995)	(5.785)	(0.141)	(0.0457)
Democrat	0.481***	-6.220	1.090***	0.0932*
	(0.0820)	(5.553)	(0.115)	(0.0419)
Constant	5.966***	92.94***	5.022***	0.323***
	(0.144)	(10.03)	(0.209)	(0.0791)
Observations	976	976	976	976
R-squared	0.077	0.027	0.153	0.030

Notes: OLS regressions with robust standard errors. 70 individuals in our sample stated that their political preferences did not align with any of the categories Independent, Democrat or Republican, or preferred not to say. The baseline category in our regression is Republican, since we found it most interesting to compare democrats and republicans.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

the experiment nor influence their final payoff. In the absence of a decision to donate, there is no reason for the cost to influence beliefs through the type of motivated cognition explored here. We thus predicted that no significant differences in beliefs would arise between the two control conditions. Finding no difference also provides an indication that participants did not interpret the stated costs as informative of the benefits of planting trees.

As predicted, we are unable to reject the null hypothesis, see Table 5, showing the result of OLS regressions comparing the perception of the benefits of planting a tree among participants in the two control conditions. Neither for our main qualitative, nor for the quantitative measure do we find a significant difference in beliefs ($p = .41$ and $p = .21$, respectively).

We included the control conditions to get an indication of the likelihood that our results are be driven by, for example, the size of costs or discounts mentioned in the treatments being interpreted as informative of the effectiveness of trees to mitigate climate change or influence demand effects. While this analysis does not allow us to conclude that there are no differences in beliefs and attitudes across the two hypothetical conditions, it can help us bound it. The upper bound of the possible effect can be derived from the minimum detectable effect: based on the standard error and our t-test alpha of 5% we could detect an effect of .176 steps on the 7-degree Likert scale. We discuss this at more length in the conclusion, but if, e.g., demand effects were important, and correlated with experimental condition, we would expect a difference in beliefs also in the hypothetical case.

Table 3.6. Beliefs in the hypothetical conditions, OLS regressions

	(1) Importance planting trees	(2) CO_2 sequestered
Constant	6.124***	96.47***
	(0.266)	(16.86)
Treatment	0.0654	-3.676
	(0.0629)	(3.970)
ageNum	-0.00395	-0.0765
	(0.00266)	(0.158)
genderInd1	0.109	7.909*
	(0.0649)	(3.982)
genderInd3	-0.237	9.232
	(0.348)	(34.76)
highEducation	-0.0205	8.041
	(0.0894)	(6.755)
highIncome	0.0896	1.037
	(0.0651)	(4.093)
Observations	981	981
R-squared	0.008	0.007

Notes: Heteroskedasticity robust standard errors in parenthesis.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

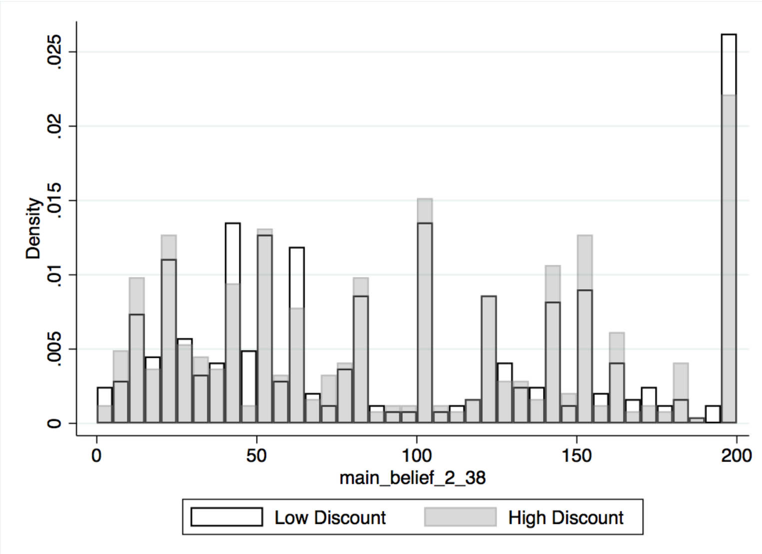
3.5. Discussion

In this study, we implement an online experiment with a total of 2,536 participants to explore whether people adjust their beliefs about the efficacy of a pro-environmental action based on the personal cost of engaging in that action. To do so we randomly assign participants to experimental conditions varying the size of a discount, and hence the cost, to a donation for planting a tree. After informing respondents whether they are randomized to receive a high or a low discount, but before they make a donation decision, we ask respondents about their perception of the usefulness to plant a tree as well how much CO_2 they believe a tree sequesters from the atmosphere in a year. As a secondary outcome variable, we ask participants to what extent they think their actions contribute to climate change.

We find that participants who are randomly assigned a large discount state a higher perceived usefulness of planting trees, but do not estimate a higher amount of CO_2 to be sequestered by trees from the atmosphere, than participants who are assigned a smaller discount. This difference in perceived usefulness is small but significant at the 1 percent level. Our results further indicate that participants in the high discount condition report stronger feelings of personal responsibility – our secondary outcome variable – to fight climate change relative to participants in the low discount group. There may be several reasons why we fail to find an effect on the estimated CO_2 sequestered from the atmosphere. One possible reason is that people may not generally know the CO_2 pound equivalent of various actions and we may have asked our participants to estimate a very unfamiliar quantity. We aimed to anchor our respondents by telling them that the average American's daily CO_2 footprint is 100 lbs per day which was also the midpoint of the scale. We cannot observe a strong anchoring or midpoint effect, see Figure 1. The mode of the quantitative belief responses is the maximum of 200 lbs on the scale for both the low and the high discount conditions, possibly indicating a ceiling effect. Lastly, it is also possible that the question was too complicated to induce quick, intuitive motivated beliefs. For motivated beliefs to occur, our participants would have to immediately see that reporting a higher CO_2 in our question corresponds with a higher effectiveness of trees, and perhaps this wasn't immediately obvious.

Our estimated impact of experienced costs on the stated importance of planting trees corresponds to about 0.16 of a standard deviation. While this is not a large effect, we still believe it is relevant. We believe motivated beliefs are likely to arise with respect to attitudes and beliefs about the environment and climate impact, as outlined in the introduction. However, we were not sure to capture such effects in a short and smaller scale online experiment. Several features of the design may also, if anything, mitigate an impact of motivated beliefs. For example, the experimental design included a discount in both conditions, and the instructions emphasized the random assignment to an exogenously imposed discount to all participants in the main conditions before participants were

Figure 3.1. Histogram of quantitative beliefs



assigned to an experimental condition. This was done to avoid that the discount itself, or any experimenter demand effects associated with a larger discount, impacted comparisons between experimental conditions. Such an effect would cause us to overestimate the impact of cost on beliefs. Further, the experimental instructions emphasized the difference between the price (which was the same across conditions) and costs (which varied across conditions) in order to avoid tendencies to interpret the cost a carrier of information. While we do not find any significant impact of the cost on beliefs and attitudes in the control conditions with hypothetical donation (for our main outcome variables) we note that, if the belief were that the cost signals quality, the effect would be opposite the one hypothesized here. We would then expect participants with high costs to believe in larger impacts. It is also possible that a desire to self-signal provides an incentive to donate when it is expensive – an effect which would also run counter to the one explored here. Self-signaling theories predict that incurring a relatively higher personal cost would increase the signaling value of engaging in a behavior. Thus, our main results provide an initial indication that subjective beliefs about the importance of a pro-environmental action depend at least in part on the personal cost experienced. We find a small effect, but it arises despite our efforts to emphasize the exogeneity of the size of the discount and the fact that the decision maker's beliefs are elicited before the final decision is made.

We believe that this study, while representing only a first step, has potentially important policy implications that should be explored in future work. Our findings that beliefs about, and recognized personal responsibility for, climate related questions vary with expected future costs indicate that communication about climate change and how our societies propose to solve it impact our beliefs about these efforts themselves. Given the difficulties to generate broad political support for sometimes costly climate friendly policies, these are important dimensions to consider. Awareness of these channels may be exploited to decrease resistance to pro-environmental legislation and associated costs. Another indication arising from our research is that lowering the costs and hurdles to engage in environmentally friendly behaviors may positively impact the perceived benefit of those behaviors and result in an increased frequency of these behaviors. In sum, we believe that follow up work replicating our findings and evaluating attitudes to important environmental policy dimensions in other contexts and in more representative samples would be valuable.

3.6. Conclusion

We present the results of an online experiment on the effect of a subsidy to the planting of a tree through a non-profit on the beliefs of the efficacy of a planted tree in combating climate change. For one of our two outcome variables, we find that subjective beliefs are higher when the subsidy to the price of the tree is larger, i.e. when the personal cost for planting a tree is low. We do not find statistically significant differences between treatment groups for the quantitative belief. We show that participants report a higher personal responsibility for combating climate change in the treatment group with a large subsidy. These motivated beliefs may be policy relevant as our findings show that the personal cost that the participant bears to combat climate change also affects her reported beliefs, even before a decision is made. Future work should investigate whether these motivated beliefs can also be identified in the field and the persistence of the effect over time.

3.A Appendix

3.A.1. Additional secondary prespecified analyses

We pre-specified a number of additional analyses for exploratory reasons in our pre-analsis plan. We first report our interaction analysis where we investigate how the treatment effect differs by our collected covariates, see Table 3.7. There are no obvious patterns detectable in the reported regressions.

We also planned to report a combined table of our main and hypothetical treatments, see Table 3.8.

Lastly, we pre-specified comparisons of our main treatment against the baseline, each for the low and high discount groups, see Tables 3.9 and 3.10. It is important to note that the comparison between Baseline and the discount treatments has one caveat: we do not have strict randomization. Due to the way we ran our pilot, we randomized the first 100 participants into the Low and High Discount groups only. The subsequent 2400 recruited participants were randomized into all five treatments. Thus, when comparing the Baseline with either discount treatments, there might be systematic bias stemming from the difference in timing and randomization probability of recruitment. However, we think that this bias is likely to not be very large, as the time between recruitment for the pilot and for the rest of the experiment was very short and because there is no immediately apparent reason for why the subject distribution should have changed substantially between that time.

Perhaps interestingly, the High Discount group has statistically significant higher subjective beliefs than Baseline, but not for the quantitative beliefs we measured. There is no significant difference in the Low Discount group for either belief measure.

Table 3.7. Heterogeneity analysis, robust OLS regressions

	Importance planting trees	CO_2 sequestered	Personal Responsibility	Donation Decision
Age	-0.000212	0.0320	-0.00953	0.00475**
	(0.00394)	(0.235)	(0.00498)	(0.00179)
High Discount	-0.295	2.301	-0.345	0.509**
	(0.469)	(24.75)	(0.691)	(0.185)
High Discount X Age	0.00153	0.0854	0.00800	9.78e-05
	(0.00501)	(0.337)	(0.00711)	(0.00232)
Female	0.0812	15.70**	0.136	0.0727
	(0.0860)	(5.662)	(0.119)	(0.0421)
High Discount X Female	0.162	-0.212	0.203	-0.0631
	(0.117)	(7.935)	(0.166)	(0.0553)
"Non-Binary"	0.492***	5.955	-0.0766	0.191
	(0.0728)	(23.04)	(0.696)	(0.373)
"Non-Binary" X Treatment	-0.622**	4.829	0.444	0.0853
	(0.193)	(30.43)	(0.818)	(0.379)
High Education	-0.0348	-14.28	0.00951	0.0613
	(0.135)	(10.04)	(0.198)	(0.0701)
High Education X Treatment	0.0466	13.70	0.0861	-0.0265
	(0.181)	(13.74)	(0.283)	(0.0942)
High Income	-0.0687	2.708	0.0783	-0.0436
	(0.0844)	(5.922)	(0.119)	(0.0436)
High Income X Treatment	0.211	-14.98	0.0960	0.0689
	(0.120)	(8.295)	(0.173)	(0.0572)
Independent	0.0728	-25.32	0.530	0.160
	(0.282)	(14.49)	(0.396)	(0.0834)
Republican	0.160	-21.32	0.459	0.149
	(0.291)	(15.17)	(0.400)	(0.0866)
Democrat	0.698*	-23.88	1.533***	0.266**
	(0.273)	(14.31)	(0.383)	(0.0832)
Independent X Treatment	0.410	-1.828	0.151	0.0542
	(0.425)	(21.26)	(0.651)	(0.164)
Republican X Treatment	0.130	-10.41	-0.142	-0.00446
	(0.436)	(22.25)	(0.660)	(0.168)
Democrat X Treatment	0.0290	-17.88	-0.0993	-0.0362
	(0.413)	(21.02)	(0.632)	(0.162)
Constant	5.886***	119.8***	4.795***	-0.146
	(0.313)	(16.50)	(0.404)	(0.101)
Observations	976	976	976	976
R-squared	0.095	0.034	0.161	0.268

Table 3.8. Combined table of main and hypothetical treatment regressions

	(1) Importance planting trees	(2) CO_2 sequestered
High Discount	0.150*	0.285
	(0.0594)	(3.958)
Hypothetical	0.00871	-0.395
	(0.0636)	(3.958)
High Discount (Hypothetical)	0.0660	-3.915
	(0.0628)	(3.968)
Age	-0.00218	0.000320
	(0.00186)	(0.114)
Female	0.158***	11.05***
	(0.0442)	(2.814)
Gender (other)	-0.0529	11.63
	(0.211)	(19.80)
High Education	0.0183	0.403
	(0.0645)	(4.824)
High Income	0.0690	-2.078
	(0.0447)	(2.892)
Constant	6.187***	89.34***
	(0.100)	(6.797)
Observations	1,957	1,957
R-squared	0.012	0.009

Notes: Heteroskedasticity robust standard errors in parenthesis.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3.9. OLS regressions between Baseline and Low Discount Treatment

	(1) Importance planting trees	(2) CO_2 sequestered	(3) Personal Responsibility
Low Discount	0.0359	1.341	-0.0911
	(0.0627)	(4.026)	(0.0863)
Age	0.000599	-0.126	-0.0117**
	(0.00252)	(0.154)	(0.00366)
Female	0.131*	12.01**	0.0941
	(0.0632)	(4.048)	(0.0877)
Gender (other)	0.357	-36.74	0.190
	(0.433)	(22.23)	(0.394)
High Education	0.0693	-11.60	0.308*
	(0.120)	(6.763)	(0.138)
High Income	-0.00789	-1.059	0.171
	(0.0637)	(4.116)	(0.0880)
Constant	6.038***	103.2***	5.628***
	(0.169)	(8.919)	(0.189)
Observations	952	952	952
R-squared	0.006	0.015	0.022

Notes: Heteroskedasticity robust standard errors in parenthesis.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3.10. OLS regressions between Baseline and High Discount Treatment

	(1) Importance planting trees	(2) CO_2 sequestered	(3) Personal Responsibility
High Discount	0.0961**	0.594	0.0497
	(0.0306)	(2.013)	(0.0430)
Age	0.000599	-0.0761	-0.0100**
	(0.00219)	(0.155)	(0.00370)
Female	0.208***	10.93**	0.210*
	(0.0602)	(4.021)	(0.0868)
Gender (other)	-0.180	-15.83	0.120
	(0.246)	(25.09)	(0.311)
High Education	0.0357	-4.814	0.226
	(0.116)	(6.728)	(0.145)
High Income	0.0701	-8.317*	0.157
	(0.0631)	(4.112)	(0.0895)
Constant	5.996***	98.82***	5.579***
	(0.163)	(8.898)	(0.204)
Observations	952	952	952
R-squared	0.024	0.015	0.022

Notes: Heteroskedasticity robust standard errors in parenthesis.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.A.2. Pre-analysis plan uploaded on OSF

Pre-registration for "Self-Serving Motivated Beliefs and Environmental Action"

Benjamin Mandl, Stockholm School of Economics
Eva Ranehill, University of Gothenburg
2020-11-16

1) What is the main question being asked or hypothesis being tested in this study?

This study explores motivated beliefs in the area of environmental behavior and attitudes, asking whether the costs associated with an environmentally friendly behavior influences the probability of engaging in the behavior, and as a consequence the perception of its value through motivated beliefs. In particular, we ask whether an exogenous change in the cost of a donation decision affects self-reported beliefs about the impact of the donation.

The environmentally friendly choice we explore is the decision whether to donate to plant a tree. We explore this in an experiment on MTurk, and this pre-analysis plan specifies our research plan. We plan to implement 5 conditions – two main conditions and three secondary conditions. In the two main conditions we vary the size of a randomly assigned subsidy for planting a tree such that participants in the LowDiscount condition get only a small subsidy (USD 0.05 to the price of USD 1) to plant a tree, whereas participants in the HighDiscount condition get a substantially larger subsidy (USD 0.95 to the price of USD 1).

In two of the three control conditions we ask participants to make the same decision as participants in our main conditions but to do so hypothetically. In the last control condition participants are not asked to donate any money, and only answer a set of questions about their beliefs about the impact of planting trees. The control conditions are described in "Section 5. Exploratory analyses" below. These control conditions are implemented in order to study underlying mechanisms. We consider all analysis based on the control conditions exploratory.

Based on our outcome variables elicited in the two main conditions as detailed below we aim to test the following hypothesis:

> H_1: Participants who face a cost of planting a tree of USD 0.05 will state a higher importance of planting a tree to combat global warming than those who face a price of USD 0.95.

against the null hypothesis of no effect:

> H_0: There is no difference in the perceived importance of planting a tree to combat global warming between participants who face a low cost to do so (USD 0.05) and those who face a high cost (USD 0.95).

2) Describe the dependent variable(s) specifying how they will be measured.

Here, we describe both the main dependent variables, as well as the construction of our explanatory variables.

Dependent variables

<u>*Primary outcome variables*</u>

1. Subjective Beliefs: Our first main outcome variable is based on the question: "Please indicate how strongly you agree or disagree with the following statement.

   Planting trees is an important activity to mitigate emissions of CO2 and combat global warming."
   (Answer options: Completely disagree, disagree, somewhat disagree, neither agree nor disagree, somewhat agree, agree, completely agree*)*

   To generate the outcome variable the seven Likert scale answer options will be transformed to the numbers 1-7.

2. Efficacy of planting trees: The second main outcome variable will simply be the response to the question:
   "The average American's carbon footprint is about **100 lbs CO2 per day**. How many lbs C02 would you estimate a **mature US hardwood tree binds per year**?"
   (*Answer options*: Slider from 0-200 lbs)

Secondary outcome variables for exploratory analysis

3. Personal responsibility: A third, exploratory, outcome variable will be based on the question:

   "To what extent do you agree with the following statement? I believe that my actions contribute to global warming and climate change."
   (*Answer options*: Completely disagree, disagree, somewhat disagree, neither disagree nor agree, somewhat agree, agree, completely agree)

   To generate the outcome variable the seven Likert scale answer options will be transformed to the numbers 1-7.

Independent Variables

The independent variable of interest is treatment - the discount that is applied to the cost of planting a tree. We will code this variable with the values 0 (low discount) and 1 (high discount).

Additional variables

In addition, we measure:

4. Whether the participant chooses to donate to plant a tree or not when given the opportunity at the end of the survey.
5. Whether a participant wants to sign up for a newsletter from the organization they can donate to.
6. Socio-demographic variables
   1. Age
   2. Gender
   3. Gross annual income (categorical)
   4. Highest level of education
   5. Political orientation (specified as Republican, Democrat or Independent)

7. Survey measures automatically collected by Qualtrics (e.g. time of day a survey was started and finished, etc.)

In the analysis, age will be used as a continuous variable. The question about gender has 3 options (male, female and other) and the variable will simply be coded using a dummy variable for each category. We ask participants to state their income in brackets of USD 10,000 (with a maximum of USD 150,000). In our analysis, income will be a dummy variable taking the value of 1 for high income based on a median split. Educational attainment is elicited in 5 categories (<high school degree, high school degree or equivalent, Some college but no degree, bachelor degree, graduate degree). Also this variable will used as a dummy variable in the analysis and will take the value 1 for all individuals who state that they have at least some college education. With respect to political orientation participants indicate “Republican”, “Democrat”, or “Independent” and the answer categories will be coded accordingly into three categories.

For our main analysis we will exclude participants who fail any of the two attention check questions more than twice. We also ask participants whether they googled an answer to the question of how much CO_2 a tree binds per year, and whether, in that case, they found an answer. In the main analysis of this question we will exclude all participants who reported googling the answer. For robustness, we will additionally report analyses with all participants.

3) How many and which conditions will participants be assigned to?

Study participants will be randomly assigned to one of 5 conditions (High or Low Discount, High or Low Discount with a Hypothetical Decision, Baseline). Indicator variables are used to indicate treatments in regressions below.

4) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

Using the data from our two main conditions, we will implement the same analysis for both our primary outcome measures (and our secondary outcome variable, as defined in Section 2, in an exploratory analysis).

Our main tests will consist of the following OLS regressions:

1. *Subjective Beliefs = a0 + a1HighDiscount + a2’X + error*
2. *Efficacy of Planting Trees = b0 + b1HighDiscount + b2’X + error*

where *X* includes control variables for age, gender, education, and income. (Political orientation will be studied in an exploratory analysis).

We will test the following hypothesis using a two sided t-test with an alpha of 5%. (All relevant tests will be two-sided (t-tests) with significance level alpha = 0.05. When applicable, standard errors will be corrected for heteroskedasticity using the STATA robust command.):

- H1: a1 > 0
- H2: b1 > 0

As a secondary test we will provide a more precise measure of how the probability to donate impacts beliefs about its importance. To do so we estimate an instrumental variables regression where the dependent variable is the beliefs about the usefulness to plant a tree, the donation decision is the endogenous variable, and treatment (0 or 1) is the instrumental variable. Also this regression will be run with controls for socio-demographic variables (age, gender, education and income).

3. *1st stage: Donation = c0 + c1HighDiscount + c2'X + error*
4. *2nd stage: Beliefs = d0 + d1Donation + d2'X + error*

where X is a vector of the control variables specified above.

- H3: d1 > 0

5) Any secondary analyses?

In addition to the two main conditions described above we will implement 3 control conditions as described above.

We will use the data from our two main conditions as well as from these three secondary conditions for a more exploratory analysis. In these more exploratory analyses, we will study various aspects of our data related to motivated beliefs, the donation decisions, the decision to sign up for the newsletter, and political preferences. More precisely, we will:

1. Do a heterogeneity analysis on the pooled data set for our main conditions, exploring whether climate attitudes and donation decisions differ with sociodemographic characteristics or political orientation. For this analysis, we will run OLS regressions with robust standard errors with our main outcome variables, as well as "Personal Responsibility" and the donation decision as dependent variables and age, gender, income, education and political preferences as independent variables.

5. *Subjective Beliefs = e0 + e1'X + error*
6. *Efficacy of Planting Trees = e0 + e1'X + error*
7. *Personal Responsibility = e0 + e1'X + error*
8. *Donation = e0 + e1'X + error*

2. Explore whether the impact of our main experimental condition interacts with sociodemographic characteristics or political orientation. This analysis will be implemented by running the same regression as described in the previous point, adding condition and an interaction of condition and all control variables.

9. *Subjective Beliefs = e0 + e1HighDiscount + e2'X + e3HighDiscount x age + e4HighDiscount x gender + e5HighDiscount x HighIncome + e6HighDiscount x HighEducation + e7HighDiscount x political_preferencces + error*
   i. *Test: each coefficient e1 to e7 ≠ 0*
10. *Efficacy of Planting Trees = e0 + e1HighDiscount + e2'X + e3HighDiscount x age + e4HighDiscount x gender + e5HighDiscount x HighIncome + e6HighDiscount x HighEducation + e7HighDiscount x political_preferencces + error*

ii. *Test: each coefficient e1 to e7 ≠ 0*

11. *Personal Responsibility = e0 + e1HighDiscount + e2'X + e3HighDiscount x age + e4HighDiscount x gender + e5HighDiscount x HighIncome + e6HighDiscount x HighEducation + e7HighDiscount x political_preferences + error*

iii. *Test: each coefficient e1 to e7 ≠ 0*

12. *Donation = e0 + e1HighDiscount + e2'X + e3HighDiscount x age + e4HighDiscount x gender + e5HighDiscount x HighIncome + e6HighDiscount x HighEducation + e7HighDiscount x political_preferencces + error*

iv. *Test: each coefficient e1 to e7 ≠ 0*

3. Explore whether respondents interpret the costs associated with planting a tree as informative of the impact of the donation, absent any donation decision. To test this, we will run regressions 1 and 2 above for the two hypothetical conditions as described below. A significant effect of treatment also in the hypothetical conditions indicates that costs are interpreted as indicative of impact. While this does not allow us to exclude that such an effect exists if small, this analysis gives us an impression if such an effect appears important. These regressions will be run with controls for socio-demographic variables (age, gender, education, and income).

13. *Subjective Beliefs = a0 + a1HighDiscountHypothetical + a2'X + error*

14. *Efficacy of Planting Trees = b0 + b1HighDiscountHypothetical + b2'X + error*

We will test the following hypothesis using a two sided t-test with an alpha of 5%:

- H1: a1 > 0
- H2: b1 > 0

4. In addition, we are also interested in running the above regression, including the observations from the four conditions with real and hypothetical decisions, as specified below.

15. *Subjective Beliefs = a0 + a1HighDiscount + a2Hypothetical + a3HighDiscountHypothetical + a4'X + error*

16. *Efficacy of Planting Trees = b0 + b1HighDiscount + b2Hypothetical + b3HighDiscountHypothetical + b4'X + error*

5. We are also interested in comparing the dynamics of motivated beliefs relative to the baseline. To do so, we will compare our two main outcome variables in the Baseline and our two main conditions respectively. We will run the following regressions:

17. *Subjective Beliefs = a0 + a1 lowDiscount + a2'X + error*

18. *Efficacy of Planting Trees = b0 + b1 lowDiscount + b2'X + error*

- H1: a1 > 0
- H2: b1 > 0

19. Subjective Beliefs = a0 + a1 highDiscount + a2'X + error

20. Efficacy of Planting Trees = b0 + b1 highDiscount + b2'X + error

- H1: a1 < 0
- H2: b1 < 0

6) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

We will determine the number of observations to collect based on the results of our pilot. For the pilot we will collect data on 100 participants who are randomized into either the high or the low discount conditions. According to our power analysis, we decided to continue to collect an additional 900 observations in our main treatments if the pilot indicates a difference in donation rates of at least 30 percentage points between conditions. If the difference in donation rates in the high and the low cost conditions are not larger than a 30% difference, we will reconsider our experimental design.

The reasoning for our desired sample is as follows. Assuming that the motivated beliefs effect is moderated fully by the effect of the price on the decision to donate, we can only expect to find treatment effects for those subjects who are marginally affected by the different donation prices. We do not know the propensity to donate of MTurkers at different prices. Thus, if we find a larger than 30 percentage points difference in donation rates, our combined sample of 1000 participants allows us to find a minimum detectable effect in motivated beliefs due to different prices of a reasonable Cohen's d: Using G*Power we find that the minimum detectable effect size of a t-test of a difference between two independent means with n=1000 is a Cohen's d of 0.177. If we assume that 30% more people decide to donate in the high discount treatment than in the low discount treatment, and if motivated beliefs depend on the decision to donate, a sample of 1000 participants could detect an effect of .177/.3 = .59. If the differences in donation rates are 50%, we can expect to find an effect of .157/.5 = 0.35. This power analysis is based on simple assumptions to help make us a decision on approximate sample sizes, and serves as a guide, and not a precise estimate.

If the pilot indicates a difference in donation rates in the main conditions of more than 30%, we will also collect 500 observations in each of the control conditions.

7) Anything else you would like to pre-register? (e.g., data exclusions, variables collected for exploratory purposes, unusual analyses planned?)

The survey includes control questions asking respondents to identify their experimental condition and what this imply for the cost of planting a tree. Respondents who fail to answer these questions correctly will be excluded from the analysis more than twice. As mentioned above, participants who stated they googled the answer to our numerical question will also be excluded from that part of the analysis.

8) Have any data been collected for this study already?

A small test sample was collected through CloudResearch, an MTURK sample provider, to test functionality of the provider and integration into the MTURK platform. The data was not used for any analyses.

3.B References

Ambuehl, S. (2017). "An Offer You Can't Refuse? Incentives Change How We Inform Ourselves and What We Believe." In: 6296.

Bénabou, R. and J. Tirole (2016). "Mindful Economics: The Production, Consumption, and Value of Beliefs." In: *Journal of Economic Perspectives* 30.3, pp. 141–164.

Buser, T., L. Gerhards, and J. J. van der Weele (2016). *Measuring Responsiveness to Feedback as a Personal Trait*. SSRN Scholarly Paper ID 2789407. Rochester, NY: Social Science Research Network.

Chew, S. H., W. Huang, and X. Zhao (2020). "Motivated False Memory." In: *Journal of Political Economy* 128.10, pp. 3913–3939.

Dana, J., R. A. Weber, and J. X. Kuang (2007). "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness." In: *Economic Theory*. Vol. 33. 1, pp. 67–80.

Di Tella, R., R. Perez-truglia, and A. Babino (2015). "Conveniently Upset : Avoiding Altruism by Distorting Beliefs about Others ' Altruism †." In: 105.11, pp. 3416–3442.

Eil, B. D. and J. M. Rao (2011). "The Good News-Bad News Effect : Asymmetric Processing of Objective Information about Yourself." In: *American Economic Journal : Microeconomics* 3.2, pp. 114–138.

Epley, N. and T. Gilovich (2016). "The Mechanics of Motivated Reasoning." In: *Journal of Economic Perspectives* 30.3, pp. 133–140.

Exley, C. L. and J. B. Kessler (2019). *Motivated Errors*. Working Paper 26595. National Bureau of Economic Research.

Farjam, M., O. Nikolaychuk, and G. Bravo (2019). "Investing into Climate Change Mitigation despite the Risk of Failure." In: *Climatic Change* 154.3, pp. 453–460.

Festinger, L. (1962). "A Theory of Cognitive Dissonance (Vol. 2)." In: *Stanford university press.*

Ging-jehli, N. R., F. H. Schneider, and R. A. Weber (2019). *On Self-Serving Strategic Beliefs.*

Gino, F., M. I. Norton, and R. A. Weber (2016). "Motivated Bayesians: Feeling Moral While Acting Egoistically." In: *Journal of Economic Perspectives* 30.3, pp. 189–212.

Gotthard-Real, A. (2017). "Desirability and Information Processing: An Experimental Study." In: *Economics Letters* 152, pp. 96–99.

Grossman, Z. (2014). "Strategic Ignorance and the Robustness of Social Preferences." In: *Management Science* 60.11, pp. 2659–2665.

Grossman, Z. and D. Owens (2012). "An Unlucky Feeling: Overconfidence and Noisy Feedback." In: *Journal of Economic Behavior and Organization* 84.2, pp. 510–524.

Hagmann, D., E. H. Ho, and G. Loewenstein (2019). "Nudging out Support for a Carbon Tax." In: *Nature Climate Change* 9 (June).

Hedlin, S. and C. R. Sunstein (2016). "Does Active Choosing Promote Green Energy Use? Experimental Evidence." In: *Ecology Law Quarterly* 43.1, pp. 107–141.

Heger, S. A. and N. W. Papageorge (2018). "We Should Totally Open a Restaurant: How Optimism and Overconfidence Affect Beliefs." In: *Journal of Economic Psychology* 67, pp. 177–190.

Hestermann, N., Y. Le Yaouanq, and N. Treich (2020). "An Economic Model of the Meat Paradox." In: *European Economic Review* 129, p. 103569.

Pace, D. and J. J. van der Weele (2020). "Curbing Carbon: An Experiment on Uncertainty and Information About CO_2 Emissions." In: *SSRN Electronic Journal.*

Schwardmann, P. and J. van der Weele (2019). "Deception and Self-Deception." In: *Nature Human Behaviour* 3.10 (10), pp. 1055–1061.

Shu, L. L. and M. H. Bazerman (2010). "Cognitive Barriers to Environmental Action: Problems and Solutions." In: *Oxford Handbook of Business and the Environment*, pp. 1–26.

Wade-Benzoni, K. A., M. Li, L. L. Thompson, and M. H. Bazerman (2007). "The Malleability of Environmentalism." In: *Analyses of Social Issues and Public Policy* 7.1, pp. 163–189.

Zimmermann, F. (2020). "The Dynamics of Motivated Beliefs." In: *American Economic Review* 110.2, pp. 337–61.

Chapter 4

Forward looking motivated memory

Benjamin Mandl

Abstract

This research tests a hypothesis of a strategic forward looking bias of memory. It asks whether subjects in a lab experiment systematically misremember information that is dissonant with an upcoming action. In the experiment subjects first learn about negative consequences of consuming meat (such as environmental or ethical issues of the production of meat) and are asked to accurately recall the information before tasting a sample of cured beef. I do not find any evidence of systematic memory distortions in our prespecified outcome measures. I discuss a range of potential explanations for our non-significant results and propose alternative research directions for investigating memory distortions.

I am greatly thankful to Anna Dreber, Magnus Johannesson, Rupert Sausgruber, Robert Östling, Gergely Hajdu, Anna Schwarz, and Simone Häckl for insightful comments. I am also thankful to seminar participants at SSE, the Eating Meat 2019 Workshop at DIW Berlin, and at the Vienna University of Economics and Business for helpful comments and suggestions. I am very grateful for research assistance by Anna Schwarz. I thank the Vienna University of Economics and Business for allowing me to collect data in the WULABS. I am grateful for help and assistance regarding lab use by Owen Powell. We thank the Mistra Center for Sustainable Markets (Misum) for generous financial support and we gratefully acknowledge funding by the Jan Wallander and Tom Hedelius Foundation.

4.1. Introduction

Memory has been shown to be subject to motivational biases. People seem to misremember information (e.g. ethical codes (Shu and Gino, 2012)) and personal actions (e.g. degree of selfishness in dictator games (Saucet and Villeval, 2019), presumably to reduce cognitive dissonance (Festinger, 1962) or restore their self-concept (Shu and Gino, 2012). It is unclear whether agents use misremembering strategically before encountering a potentially uncomfortable situation to avoid any arising of cognitive dissonance in the first place. This project aimed to investigate whether people bias their recall accuracy of facts to avoid feeling potential dissonance in a future action. I test this hypothesis in a laboratory experiment. In the experiment, people are first given information about externalities associated with consuming a particular good. Then the participants in the treatment cell are informed that they will be asked to consume that good at the end of the experiment. Before being asked to consume the good, the participants are incentivized to recall the information provided at the beginning of the experiment. In particular, people are asked to recall numerical information about the externalities of eating meat before they are asked to taste a beef sample ('meat treatment'). The recall accuracy is compared to the recall of participants in a control condition who are asked to taste a fruit sample at the end of the experiment ('fruit treatment').

This potential recall bias as been hypothesized in both the economics and the psychology literature. In models of motivated beliefs (see Bénabou and Tirole, 2016), self-deception is modeled as affecting the probability of accurately recalling an informative signal. Agents are aware of the impact of adverse information in an earlier stage on future actions and thus strategically obfuscate the original signal. Similarly, in the seminal article on motivated reasoning by Kunda (1990), biased memory search and belief construction were argued to be the main facilitators for motivated reasoning. Similarly, confirmation bias, i.e. the tendency to interpret information in line with prior beliefs (reviewed in Nickerson, 1998), offers almost the same predictions to the recall accuracy as motivated beliefs. However, there is a small difference. Under confirmation bias, information interpretation is biased towards prior beliefs. Thus, when one has a negative view of one's ethical behavior in the past, confirmation bias would predict a higher recall accuracy of past unethical information or actions. However, the assumption in models of motivated beliefs is that the goal or motivation of the agent is to increase self-esteem (Bénabou and Tirole, 2016). Thus, even people with a negative view of one's ethical behavior would try to obfuscate negative signals in order to build up their self-esteem[1].

Investigating whether there is motivated recall is especially interesting with regards to understanding how factual information does or does not affect decision making. One such relevant decision is whether to eat meat. Meat consumption entails undeniable negative

[1]This automatic optimistic self-deception does seem to be less common or absent in people suffering from depression (Alloy and Abramson, 1979; Korn et al., 2014)

externalities: it is estimated to be one of the biggest contributors of CO_2 emissions from private consumption (Ivanova et al., 2016), leads to the suffering and death of billions of animals, and some meat is suspected to be carcinogenic (Willett et al., 2019). Despite this, global meat consumption is expected to rise or stay stagnant, even in most developed nations (OECD and Nations, 2020). Meat eating is an especially interesting phenomenon considering that a large share of meat eaters also exhibit empathic feelings for animals. The co-occurrence of apathy and empathy with the suffering of animals was coined the 'meat paradox' (Hestermann et al., 2020, Bastian and Loughnan, 2017). One explanation for the meat paradox could be that people misremember the facts when trying to make a decision on whether to eat meat, i.e. they exhibit motivated recall.

Prior studies investigating motivated recall share a specific detail in their design: participants are asked to recall actions that they performed or information that they already received, where these actions or information are hypothesized to induce cognitive dissonance (Festinger, 1962). In one example of this backward looking dissonance reduction study design, participants receive feedback on an IQ test and are asked to recall parts of that feedback after some time (Zimmermann, 2020). Saucet and Villeval (2019) let participants play dictator games and then ask participants to recall the amounts donated to the receivers. They find that dictators remember themselves to be more altruistic than they were. In a similar experiment, Carlson et al. (2018) provide evidence that this misremembering effect is driven by those dictators that give less than what they personally believed was fair. In contrast, participants in the present study are asked to recall information *before* they perform an action. In particular, participants are incentivized to recall objective statements before performing an action that would presumably induce cognitive dissonance if the objective statements were taken into account. Thus, I investigate what I would call forward looking justifications for behavior. To my knowledge, I am the first to do so.

To test whether people actually bias the recall of informative signals when they are about to eat meat, an experiment needs to control for three things: first, only meat eating participants are invited in order to increase compliance with the treatment protocol. Since all participants are randomized in the session to either taste a meat sample or a fruit sample, meat abstaining participants would drop out of the meat treatment but not out of the fruit treatment which could result in sample selection effects. By only inviting meat eating participants, participants do not systematically drop out of the experiment. Second, to test biased recall instead of biased guessing or estimation and to control for prior knowledge, the information that is to be recalled needs to be given to the participant within the experiment. Third, to test whether biased recall affects the decision to engage in meat eating, the information needs to be recalled imminently before the meat is to be eaten. To fulfill these requirements, the present experiment employed a pre-survey to make sure only non-vegetarians are invited into the lab. Furthermore, participants were provided with the information that was to be recalled later in the first part of the experiment. Recall was tested with a delay of about 20 minutes within the same session and imminently before a meat sample is to be consumed.

In this experiment, I find no statistically significant effects of being asked to eat meat on the recall of meat related statements relative to being asked to eat fruit in a sample of 194 participants. Recent papers found self serving recall effects (Saucet and Villeval, 2019, Carlson et al., 2018; Kunda, 1990; Zimmermann, 2020; Chew et al., 2020) but their designs test recalling past behavior or performance. Thus there are three possible reasons for why there is no effect in this experiment. First, it is possible that there is no strategic, forward looking biased recall. Second, the current experiment may have been underpowered. Third, the design of the experiment may have been insufficient for testing the hypothesis: perhaps the design did not allow for (motivated) inattention to occur, the treatment of eating meat was not salient enough, or maybe the time span between learning and recall was too short. Follow up studies should investigate whether alternative designs could isolate the strategic motivated recall more effectively.

The paper proceeds as follows: Section 4.2 lays out the experimental design, hypotheses, and procedures. Section 4.3 describes the results, which are discussed in Section 4.4. Section 4.5 concludes.

4.2. Experimental Design

4.2.1. Background

A bias in the probability to accurately recall a given piece of information is a key modelling assumption in the family of models on motivated belief by Bénabou and Tirole (2016). In their model, an agent at the initial time period decides whether to accurately pass on a signal to his future self or to obfuscate the signal by engaging in self deception by biasing her accuracy of correct recall of the signal. Though, Bénabou and Tirole (2016) note that those memory biases need not be literally detectable for their mechanism to exist[2], anectodally it is not uncommon to observe individuals inaccurately remembering facts in a self serving way: for example a job applicant might like to misremember his GPA to be a little better than it actually was or a nefarious researcher might misremember how many specifications were run before finding a significant treatment effect.

This interpretation of self serving recall of facts would also be in line with the recent new theory on the social nature of reasoning (Mercier, 2016, Mercier and Sperber, 2017). According to the argumentative theory of reasoning, reasoning's main function is to find arguments to convince others. The theory predicts that people systematically exhibit a self-serving bias for finding reasons to argue for their goal. While this bias can be corrected by a vigilant listener, in isolation it may manifest in self-serving recall of facts or events.

[2] Bénabou and Tirole (2016) describe "as if" explanations of memory bias such as biased attention or information avoidance which would result in the same mechanism that could be modelled identically

Similarly, Von Hippel and Trivers (2011) observed that self-deception in its multiple forms may have evolved to better convince others. Motivated recall of facts to better suit one's goals would be in line with the argumentative theory as the goal of convincing others may be in conflict with the goal of the most accurate memory or set of beliefs. Several recent laboratory results support the argumentative theory of reasoning, such as the finding that overconfidence helps in deceiving others (Schwardmann and van der Weele, 2019) or that arguing a randomly determined position leads to motivated beliefs supporting that position (Schwardmann, Tripodi, et al., 2019).

Given these self-signalling (Bénabou and Tirole, 2016) and others signalling (Von Hippel and Trivers, 2011) motives for self-deception, I predict that participants in a lab experiment will have motive to bias the accuracy of their recall of factual statements to signal (to oneself or others) that a behavior they are about to engage in is not that bad. In other words, I predict that people inaccurately recall facts as an excuse for why engaging in a future behavior should not lead to cognitive dissonance. I hypothesize that there might be two likely ways in how people might bias their recall accuracy: first, people might exhibit a directional bias in their recall. Such overoptimistic beliefs about the negative externalities of personal actions have been identified as important barriers to environmental action (Wade-Benzoni et al., 2007). Alternatively, it is possible that people remember dissonant facts more inaccurately overall. In other words, it is conceivable that participants' hypothesized motivated recall results in a tendency to forget or generally misremember information by exhibiting more noisy recall.

4.2.2. Conditions

I test these questions in a lab experiment applied to the topic of meat consumption. Meat consumption is a demonstrated source of self-deception (Hestermann et al., 2020, Bastian, Loughnan, et al., 2012, Bastian and Loughnan, 2017) and policy relevant with the respect to climate change (Willett et al., 2019) because eating a plant based diet has a high potential to reduce one's personal greenhouse gas emissions Wynes and Nicholas (2017). I operationalized the research question in terms of the provision and recall of numerical information on the consequences of eating meat. After a learning phase participants are asked to recall the numerical information accurately while anticipating to eat meat in the treatment condition or fruit in the control condition.

The experimental treatment required participants to be willing and able to consume meat and fruit. Since no data on food preferences of potential participants existed, inviting all participants to the experiment directly could have resulted in differential selection effects. People abstaining from eating meat in general would have not consented to the meat treatment but would presumably not have had a problem eating dried fruit introducing systematic differences in participant population of the two treatments. To deal with this potential differential selection problem I employed an incentivized pre-survey that was sent out to all potential participants through e-mail. The pre-survey invited potential participants to sign up to become a potential participant in a lab experiment in which "food will be served". To be able to accomodate food preferences in this lab experiment, responders were asked to provide their food preferences and allergies. Every respondent was entered into a lottery to win one of three Amazon gift cards each worth €20. Participants were not told any details about the lab study and that meat or fruit will be served.

Participants who said to eat meat, fruit, and were not allergic to sulphites[3] were invited to the lab experiment. On arrival, participants were randomized within the session into either the meat or the fruit treatment by choosing a randomly determined seat in the laboratory during arrival. Participants were asked to put away their phones and all note taking devices as these could be used to affect recall abilities. All other instructions were given via the computer screen.

Participants signed an informed consent form before beginning the study. The study was designed in four parts. In the first part, the learning part, participants were told that they were to learn 16 statements that they would be incentivized to recall later in the experiment. These 16 statements were made up of 8 statements concerning the negative externalities of eating meat (e.g. 'Meat and dairy uses the vast majority of farmland and produces 60% of agriculture's greenhouse gas emissions.') and 8 statements concerning problems regarding plastic waste management (e.g. 'Plastic packaging comprises more than 62% of all items (including non-plastics) collected in international coastal clean-up operations.'). The latter topic was chosen to prevent experimenter demand effects by increasing the range and topics of the statements. Statements were shown one-by-one in a random order and for at least 10 seconds to ensure that the participants read all statements.

After all participants in a session completed this stage, the experimenter brought out individual samples of meat and fruit and put them on the participants' desks while they already worked on the next part of the experiment. Participants were unable to see other participants' samples due to seat dividers. Food samples were dealt out after all participants completed part one to make sure that the food does not impact the learning stage of the experiment, as the goal of the study was to investigate a recall and not a learning or attention bias. Participants were not allowed to eat the food samples until the end of the experiment.

In the second part, participants filled out three incentivized measures that fulfilled two purposes: one, some of the measures are used as controls for the analysis, and two, the measures took at least 20 minutes to fill out which increases the time between learning

[3] The dried fruit contained sulphites

and recall[4]. The following measures were filled out: a numeracy test (comprised of the Berlin numeracy questions and the CRT, (Weller et al., 2013, Frederick, 2005) with each correct answer being incentivized by a € 0.25 payment; a free recall test (Murdock, 1962) in which participants had to learn 16 words which were shown for 4 seconds each, wait for 90 seconds, and then freely recall as many words as possible (repeated three times and incentivized by choosing one of the three 16 word lists at random and paying € 0.25 for each correctly recalled word); and an anagram test where participants had to solve up to 16 word scrambles, being paid € 0.25 for each correctly solved scramble.

After the second part, participants were asked to make sure that they received a food item on their desk and to raise their hands if they did not. This was done to make sure participants notice and anticipate that they will have to eat the sample in the course of the experiment.

In the third part, participants were asked to recall the numerical information from the statements learned in part one in a cued recall format (e.g. 'What is agriculture's percentage of greenhouse gases that can be attributed to producing meat and dairy?'). Questions based on the 16 original statements were shown one-by-one in random order. All questions were asking for a single number, a percentage from 0 to 100. Accuracy was incentivized through the quadratic scoring rule, paying € 0.25 for an exactly correct answer. After the recall stage, participants were asked to indicate how much they believe each statement using a slider from 0 to 100 without incentives. Additionally, participants reported whether they would vote for an increase in taxes on meat by 10% and whether they believe that their individual actions contribute to climate change on a five-point Likert scale. Lastly, participants report age, gender, education, and income.

In the last part, participants in the meat treatment were asked to taste a small sample of Bresaola (a slice of dried beef) and participants in the fruit treatment tasted a small sample of dried apricot. After the tasting, they filled out a brief survey about the taste, appearance, texture, and general quality of the item on a five-point Likert-scale. The results of the tasting was not recorded as it only served to create a pretense for why the participants were asked to try food samples. Compliance with the tasting protocol was very high, only two participants did not taste their respective samples. Since the prespecified analysis employed an intention to treat analysis, compliance was not recorded and thus those participants could not be identified or excluded from data analysis.

[4]The minimum time between learning and recall required for memory effects is still under debate. In economics, studies use between 8 minutes (Saucet and Villeval, 2019) and multiple weeks (Zimmermann, 2020 to generate memory effects.)

4.2.3. Hypotheses

The experimental design, standard error corrections, hypotheses, excluded observations, significance levels, sample size, and analyses were pre-registered on OSF[5] prior to data collection. I mention any deviations from the pre-analysis plan. I test two possible ways of memory distortion. First, I test whether participants have a lower accuracy in recalling facts when those facts might result in cognitive dissonance. Zimmermann (2020) finds that motivated recall results in a tendency to forget or generally misremember dissonant facts by exhibiting noisy recall. I operationalize this hypothesis by testing whether there is a statistically significant difference in the absolute deviation between the given and the true answers between the meat and the fruit treatments, measured by the accuracy index.

The accuracy index is calculated from the absolute deviation between the given answer and the true answer for each of the 8 recalled facts and then taking the mean of the 8 deviations.

$$\text{accuracy index}_i = \frac{1}{8}\sum_{p=1}^{q}\left|\text{given answer}_{i,p} - \text{true answer}_{i,p}\right|$$

where q stands for each question $q \in \{1, \ldots, 8\}$ and given and true answers are integers between 0 and 100.

Hypothesis 1:

Participants assigned to the "Meat" treatment exhibit motivated recall of information relative to participants in the "Fruit" treatment as measured by the accuracy index of the recalled information: the accuracy index of participants in the "meat" treatment is higher than the accuracy index of participants in the "Fruit" treatment.

I test this hypothesis using a two-sided t-test of the accuracy index between the participants in the "meat" vs the "fruit" treatment. I also prespecified running a controlled regression to test for robustness of the results to inclusion of covariates. I regress the accuracy index on treatment using heteroskedasticity robust standard errors and the following controls: sex, age, income, level of education, numeracy score, free recall score, and time of experiment. I test the hypothesis using a t-test on the coefficient of treatment.

Second, I test whether there is a directional bias in recall accuracy. I aggregate the numerical answers given in the recall stage of the experiment into a single index, called the estimation index. The estimation index measures the average deviation between the given answer and the true answer.

[5] see https://osf.io/ehac3

$$\text{estimation index}_i = \frac{1}{8}\sum_{p=1}^{q} \text{given answer}_{i,q} - \text{true answer}_{i,q}$$

All numerical facts given in the experiment were chosen such that a smaller given answer should result in a lower level of cognitive dissonance for meat eaters: by giving a lower answer the participant can downplay the negative consequences of eating meat. Thus, using this index I try to test whether there are overly optimistic beliefs about the externalities of personal actions. Overly optimistic beliefs have been hypothesized to be important barriers to environmental action (Shu and Bazerman, 2010, Chew et al., 2020).

Hypothesis 2:

Participants assigned to the "Meat" treatment exhibit motivated recall of information relative to participants in the "Fruit" treatment as measured by the estimation index of the recalled information: the estimation index of participants in the meat treatment is lower than the estimation index of participants in the control treatment.

I test this hypothesis using a two-sided t-test of the estimation index between the participants in the "meat" vs the "fruit" treatment. For robustness, I regress the accuracy index on treatment and the following controls: sex, age, income, level of education, numeracy score, free recall score, and time of experiment. I test the hypothesis using a t-test on the coefficient of treatment.

In addition to the main hypotheses, I prespecified related secondary and exploratory hypotheses regarding the effect of having to eat meat on responses to policy relevant questions regarding meat taxes, personal responsibility for climate change, as well as whether the average belief in the veracity of the statements is lower in the meat treatment participants than in the fruit treatment participants.

Hypothesis 3 (secondary):

Participants assigned to the "Meat" treatment show different support for increasing the taxes on meat relative to participants in the "Fruit" treatment

I test this hypothesis using a Mann-Whitney U test on the Likert scale answers on the question of whether one supports increasing the taxes on meat between the "meat" vs the "fruit" treatment.

Hypothesis 4 (secondary):

Participants assigned to the "Meat" treatment show different beliefs in whether their individual actions contribute to climate change relative to participants in the "Fruit" treatment

I test this hypothesis using a Mann-Whitney U test for beliefs in whether individual actions contribute to climate change between the participants in the "meat" vs the "fruit" treatment.

Hypothesis 5 (secondary):

Participants assigned to the "Meat" treatment have a lower average belief in the statements relative to participants in the "Fruit" treatment

I test this hypothesis using a two-sided t-test of the average belief in the statements between the participants in the "meat" vs the "fruit" treatment. For robustness, I additionally regress average belief in the statements on treatment and controls discussed above and perform a t-test on the coefficient of treatment.

4.2.4. Procedures

The study was conducted at the WULABS, the experimental economics laboratory at the Vienna University of Economics and Business in Vienna, Austria between December 16th 2019 and January 16th 2020. Participants were recruited in two waves, one in December and one in January. Recruitment was performed in two stages. First, I sent invitations to a brief[6] pre-survey containing one question about food allergies and preferences to the participant pool of about 1,600 participants using ORSEE (Greiner 2004). Participants were incentivized to fill out the survey by being entered into a drawing of three €20 Amazon gift cards, per wave. 376 people filled out the pre survey in the December wave, and 243 in the January wave. To rule out selection effects, the pre-survey contained no information regarding the requirements to eat meat but did mention that it is necessary to fill out the pre-survey to be potentially invited into a lab experiment in which food will be served. In the second step of participant recruitment, we invited all participants who answered the pre-survey and reported that they ate beef, pork, and fruits, and were not allergic to sulfites (because the fruit treatment food item contained sulfites). I invited 248 potential participants in Wave 1, and 166 in Wave 2. 194 participants responded to the invitation and completed the experiment. Sample size was based on the research budget: potential participants were invited until the budget limit was reached. After participants finished

[6] Mean completion time: 72 seconds

all parts, they were paid and left the laboratory individually. Participants' payments were rounded up to the nearest 10 cents. The average payment was € 13.73.

Table 4.1. Dietary preferences

Food restriction	Reported
no pork	28.11%
no beef	20.36%
no gluten or wheat	2.42%
no eggs	5.01%
no peanuts or nuts	3.72%
no soy	5.49%
no dairy or lactose	10.50%
no fruits	1.94%
no sulphites	8.08%

Notes: Distribution of food allergies in pre-survey sample.

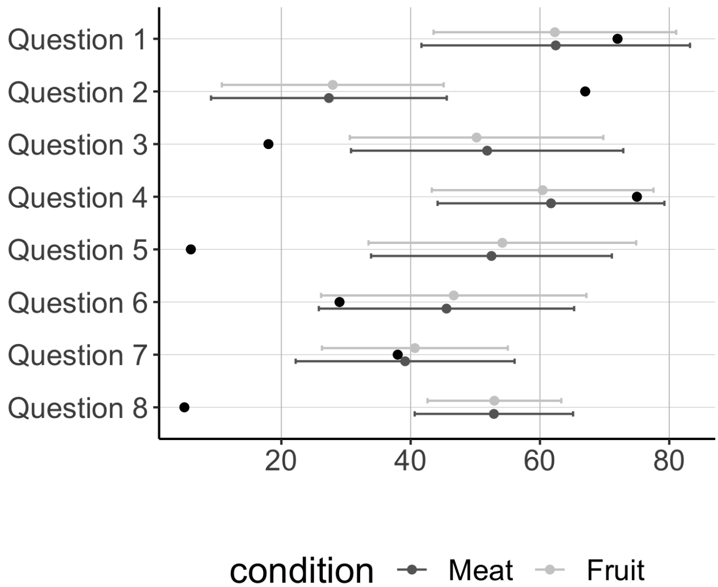
4.3. Results

For the main analysis, I follow the pre-registered analysis plan. The pre-analysis plan lays out all calculated measures, hypotheses, data exclusion criteria, and tests and regressions. All t-tests in the analysis are two-sided, independent samples tests assuming unequal variances. The prespecified statistical significance threshold is 5%. No data was removed before the final analysis.

Contrary to the hypotheses, I find no statistically significant difference between the meat and the fruit treatment group for any of the outcome indices: the accuracy index (mean difference = -0.19, $t(191.91) = -0.22$, $p = 0.82$, independent samples t-test) as well as on the estimation index (mean difference = 0.24, $t(191.01) = 0.22$, $p = 0.83$, independent samples t-test). Graphing the average given answers with one standard deviation one can see that the answers between conditions are very similar, see Figure 4.1.

Turning towards the robustness analysis of the main hypotheses, a regression of the outcome indices on the treatment, controlling for gender, age, income, level of education, numeracy score, free recall score, and time of experiment fixed effects does not show a statistically significant difference in the treatment dummy (see Table 4.2). The sign of the treatment coefficient in the accuracy index is in the opposite direction than expected. For the estimation index, the coefficient is negative as expected but not statistically significant.

Figure 4.1. Answers by condition



Note: Average answers by condition with error bars of one standard deviation. Black points show the true answers per question.

In the pre-specified secondary analyses, I perform three tests and find no statistically significant difference between any of them, contrary to my hypotheses: there is no statistically significant difference in support for increasing taxes on meat (Mann Whitney U test, $p = 0.99$), belief in climate responsibility (Mann Whitney U test, $p = 0.66$), and differences in average belief in the statements (mean difference = 3.87, $t(190.68) = 1.58$, $p = 0.12$, independent samples t-test). Column 3 in Table 4.2 reports also no statistically significant covariates of the regression of reported belief of whether the statements are true on treatment and covariates.

Table 4.2. Main and secondary regressions

	Dependent variable:		
	Accuracy index	Estimation index	Reported beliefs
	(1)	(2)	(3)
Constant	34.619***	8.025	93.132***
	(3.190)	(4.109)	(10.714)
Condition = meat	−0.707	−0.276	−2.361
	(0.928)	(1.192)	(2.604)
Gender = male	−0.355	1.649	1.152
	(0.916)	(1.201)	(2.779)
Gender = prefer not to say	1.136	2.339	−8.701
	(2.920)	(2.849)	(8.898)
Age	0.075	0.067	−0.196
	(0.074)	(0.102)	(0.333)
Numeracy score	−0.311	0.207	0.870
	(0.267)	(0.316)	(0.755)
Free recall score	−0.054	0.254**	−0.065
	(0.087)	(0.096)	(0.235)
Fixed effects	Yes	Yes	Yes
Observations	194	194	194
R^2	0.132	0.161	0.127
Adjusted R^2	0.009	0.042	0.003
Residual Std. Error (df =169)	5.897	7.470	17.132
F Statistic (df = 24; 169)	1.074	1.354	1.024

Notes: Heteroskedasticity robust standard errors in parentheses. Fixed effects include education, income, and time of day at start of the experiment in hour categories.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

4.4. Discussion

In this experiment on the effect of imminent meat consumption on biased recall of statements, I find no statistically significant effects of the treatment on recall or beliefs about the statements, or relevant policy support questions. I discuss a few possible reasons for the absence of an effect of the treatment: first, there may not be an effect of a significant size, second, the experiment may be underpowered to detect a true but small effect, or third, the study design might not have been conducive to finding an effect, i.e. the treatment used in the study was not effective.

One possible interpretation of the results is that there is no true effect of the treatment in the experiment. Still, a number of studies has found and described motivated recall errors (Saucet and Villeval, 2019, Carlson et al., 2018; Kunda, 1990; Zimmermann, 2020; Chew et al., 2020). There are two main differences between the studies that did find an effect and this present study. First, in all previous studies, participants perform an action or receive information about their abilities or actions that is potentially dissonant with their prior attitudes, self-concept, or morality. Subsequent to such a dissonance inducing treatment, motivated memory effects could be found. In this study, the dissonance inducing behavior was scheduled after the recall test, as the study wanted to test whether people strategically reduce the accuracy of their recall to justify meat eating. It is possible that the mechanisms or effect sizes are different in a backward looking dissonance reduction paradigm than in a forward looking justification paradigm as employed in this study. This seems to be inconsistent with the model by Bénabou and Tirole (2016), in which motivated beliefs arise due to biased recall or memory. However, the authors state that this bias could also arise through mechanisms unrelated to recall, such as information avoidance or motivated attention, which would be modelled equivalently to biased memory. In this study design, I tried to curtail any information avoidance and inattention by informing participants that the facts are going to be recalled later on and that the recall will be incentivized. It is possible that treatment effects would have arisen without these precautions because then these alternative mechanisms may have resulted in motivated beliefs. Future research could investigate this question further. Second, the operationalization of the recall effect is different in this study compared to earlier studies. This study operationalized accurate recall as a precise numerical recall of a cued number. Inaccurate recall was measured as the distance to the true numerical answer. Prior studies operationalized recall either as free recall (e.g.Chew et al., 2020) or as multiple choice (e.g. Saucet and Villeval, 2019), oftentimes including excuses that are potentially easier to have motivated recall about such as selecting 'I do not recall' on a multiple choice question. Given these differing operationalizations, a zero effect in this study and a robust effect in prior studies does not have to constitute a contradiction.

Second, it is possible that this study was underpowered to find a true but small effect. The estimated minimum detectable effect size at a 5% level is small: for the accuracy index it is 2.57 percentage points and for the estimation index it is 3.25 percentage points. Participants would have had to recall the facts with a bias of around 3 percentage points on average in the meat vs the fruit treatment to make detection of a bias possible. The standard deviation of the accuracy index is 5.9 percentage points and the standard deviation of the estimation index is 7.6 percentage points. Previous experiments that did find an effect did not feature substantially bigger sample sizes but reported effect sizes are potentially overestimated (Gelman and Carlin, 2014). It is possible that a higher powered experiment could have found statistically significant results but it is unknown until further research.

Third, the experimental design could have had flaws that made it more difficult to detect effects. The experiment required participants to notice and consider a small food item that was placed on their desk during the experiment. The participants' attention was drawn to the food item during placement and also within the experiment text but it is possible that participants ignored (potentially even tried to actively ignore) the item and the consequences of having to taste the item. Also, it was necessary that participants experience the different parts of the study as a single experiment. Potentially, participants mentally compartmentalized the study into two separate studies: a recall study and a tasting study. This mental division may have made it easier for participants to consider the implications of the information they received in isolation from the tasting of the food item which would imply no or only tiny effects on recall. However, participants were given ample information that signaled that the study is to be considered a single study. First, participants invited into experiments were told that 'this experiment will contain a food tasting'. Second, both the fact that the food items were dealt out during the experiment and that participants were made aware of the food item during the experiment make it unlikely that this separation explains the absence of an effect. Lastly, it is possible that participants did not experience cognitive dissonance because they considered eating meat as part of the study instead of a free choice. Thus, participants may have blamed the experimenter for the negative externalities that he caused instead of feeling cognitive dissonance. Thus, it is possible that future research could use an updated study design that takes care of these identified issues to better isolate the effect.

4.5. Conclusion

This study tests a theory of forward looking memory distortions, applied to the situation of rememebering uncomfortable facts when deciding to eat meat. I do not find any significant effects in the propensity to remember dissonant information about eating meat in the treatment group. It is possible that the non-signficant findings are due to study design, low power or the lack of a (sizeable) treatment effect in this experimental paradigm. This begs the question of whether the anectodally well known effect of favorably misremebering facts to fit one's story is a motivated memory effect that is detectable in regular laboratory experiments, or whether alternative, non-memory based theories such as motivated attention or information avoidance explain these stories. For example, maybe people do not misremember facts but rather simply lie about them to fit their agenda. Future research should study such alternative mechanisms that could also explain the appearance of motivated memory, such as conscious misrepresentation of factual statements for argumentative purposes to shed light on these questions.

4.A Appendix

4.A.1. Pre-analysis plan uploaded on OSF

Recall project: Pre-Registration Document for Experiment Data Analysis

Author: Benjamin Mandl

Introduction:

This paper studies whether the recall accuracy of decision relevant information is affected by directional goals. The existence of biased memory search and belief construction due to directional goals has been theorized for some time (e.g. Kunda 1990). This paper aims provide the evidence that directional goals have a causal effect on whether numerical factual information can be recalled accurately.

Participants are invited to participate in an experiment in which food will be served at the Vienna University of Economics and Business. In the invitation to signing up, participants are asked to report their food preferences (most importantly, whether they eat a vegetarian/vegan diet). Those subjects that do not report to be vegetarian/vegan, are invited to participate in the experiment. When the participants enter the laboratory, they are randomly assigned to two different treatments (between subject design, within session randomization):

1. Meat: Participants are asked to sample and rate a snack made of meat (air dried beef called Bresaola) at the end of the experiment.
2. Fruit: Participants are asked to sample and rate a fruit snack (dried apricots) at the end of the experiment.

Experimental procedure:

1 week before running the experiment, invites to the entire lab population of the Vienna University of Economics and Business was sent out that include a short survey on food preferences. Filling out the survey is prerequisite to be invited to the experiment. Only potential participants that eat beef are invited. This survey was not used for any other purposes than inviting participants and no data of this survey will be used in the final analysis.

The experiment consists of multiple parts.

1. Participants are told that they will read 16 statements and that there will be an incentivized test about these statements in a later part of the experiment. Participants are given 8 statements that imply negative externalities of eating meat and 8 statements that are related to plastic waste. The order of the statements is randomized. To make sure that all statements are read all participants see only one statement at a time and can only proceed after a minimum time of 10 seconds has passed.
2. Participants take a test to assess their numeracy score (Weller et al., 2012), a verbal free recall test in which 16 words are displayed for four seconds and then freely recalled after a 30 seconds waiting period. (e.g., Murdock Jr (1962); Tulving et al. (1972)) The wordlist learning and recall procedure is repeated with three word lists in total and one is chosen randomly to be scored in the end of the experiment. Lastly, participants also have four minutes to solve up to 16 anagrams (creating one word out of randomly ordered letters).

3. Participants are given food items to be sampled while they complete the tasks described in point 2 above. They are asked to sample the items at the end of the experiment so that the experimenter can prepare the payments while the participants eat the food.
4. Participants answer 16 numerical quiz questions (incentivized using the quadratic scoring rule), 8 for each topic (meat and plastic).
5. Participants report gender, age, income, level of education, whether they agree with raising the value added taxes on meat from 10% to 20% on a 5 point Likert scale, and whether they believe that personal actions contribute to climate change on a 7 point Likert scale. Lastly they also report their degree of belief in each of the 16 statements on a slider from 0 to 100.
6. Participants are asked to eat and rate them on taste, perceived quality, appearance, and texture on 5 point Likert scales.
7. Participants are paid and exit the laboratory.

All study material was uploaded to OSF with the final pre analysis plan before the data collection started.

Calculated Measures:

The following measures are calculated from the raw collected data:

1. The "**accuracy_index**" of recalled information in the quiz:
    a. Each question $q \in Q = \{1, \ldots, 8\}$ in the quiz on the topic $t \in T = \{meat, plastic\}$ has a true answer, called $true_{i,q,t} \in \{1, \ldots, 100\}$. Call participant i's answers for each question q in topic t $given_{i,q,t} \in \{1, \ldots, 100\}$.
    b. In the first step, the following is calculated:
        i. $abs\ answer\ deviation_{i,q,t} = \sqrt{\left(given_{i,q,t} - true_{i,q,t}\right)^2}$
    c. All absolute answer deviations given by a participant are averaged to create a single, average numerical accuracy_index per participant, i.e.:
        i. $accuracy\ index_{i,t} = \frac{1}{8}\sum_{p=1}^{q} abs\ answer\ deviation_{i,q,t}$
    d. The accuracy index is calculated because it is conceivable that subjects' hypothesized motivated recall results in a tendency to forget or generally misremember information by exhibiting noisy recall (Zimmermann, 2019).
2. The "**estimation_index**" of recalled information in the quiz:
    a. All normalized answers given by a participant are averaged to create a single, average numerical estimation_index per participant, i.e.:
        i. $estimation\ index_{i,t} = \frac{1}{8}\sum_{p=1}^{q}\left(given_{i,q,t} - true_{i,q,t}\right)$
    b. The estimation_index is calculated because there seem to be indications of overly optimistic beliefs about the negative externalities of personal actions which could be an important barrier to environmental action (Shu and Bazerman, 2010; Chew, Huang and Zhao, 2019). Using the estimation_index, this project aims to test whether such overoptimism is detectable in the recall of numerical information.
3. Numeracy score:
    a. In the numeracy scale, each question has a single true answer.
    b. The numeracy score is the sum of correctly answered questions.

4. Free recall score:
    a. In the free recall task, participants are shown 16 random words one by one for 4 seconds each. After a 10 second pause, they are asked to freely recall as many words as possible in 90 seconds. This procedure is done with 3 lists in total.
    b. The sum of freely recalled words of the word lists. Only exactly correct spellings will be counted.
5. Average belief in statements per topic:
    a. Arithmetic mean of the reported belief in each statement, $average\ belief_{i,t} = \frac{1}{8}\sum_{p=1}^{q}\left(belief_{i,q,t}\right)$

Regression controls

For all prespecified regressions below, the following control variables will be used:

- Female indicator, where 1 indicates a female participant.
- Age.
- Income.
- Level of education.
- Numeracy score
- Free recall score
- Time of experiment (categorical variable of the full hour at start of the experiment, e.g. experiment that started at 10:05 am is coded as categorical 10.).

Standard error corrections

T-tests will assume unequal variances and all regressions will use heteroskedasticity robust standard errors. T-tests are always two sided.

Primary Hypothesis and Tests:

Hypothesis 1:
Participants assigned to the "Meat" treatment exhibit motivated recall of information relative to participants in the "Fruit" treatment as measured by the accuracy index of the recalled information: the accuracy index of participants in the "meat" treatment is ***higher*** *than the accuracy index of participants in the "Fruit" treatment.*
Test of Hypothesis:

1. Main test: Two-sided t-test of the $accuracy\ index_{i,meat}$ between the participants in the "meat" vs the "fruit" treatment.
2. Robustness test: Regressing $accuracy\ index_{i,meat}$ on treatment and controls discussed above and performing a t-test on the coefficient of treatment.

Hypothesis 2:
Participants assigned to the "Meat" treatment exhibit motivated recall of information relative to participants in the "Fruit" treatment as measured by the estimation index of the recalled information: the estimation index of participants in the meat treatment is ***lower*** *than the estimation index of participants in the control treatment.*
Test of Hypothesis

1. Main test: Two-sided t-test of the $estimation\ index_{i,meat}$ between the participants in the "meat" vs the "fruit" treatment.

2. Robustness test: Regressing $estimation\ index_{i,meat}$ on treatment and controls discussed above and performing a t-test on the coefficient of treatment.

Secondary Hypotheses and Tests:

Hypothesis 3 (secondary):
Participants assigned to the "Meat" treatment show different support for increasing the taxes on meat relative to participants in the "Fruit" treatment
Test of Hypothesis:

1. Main test: Mann-Whitney U test for increasing the taxes on meat between the participants in the "meat" vs the "fruit" treatment.

Hypothesis 4 (secondary):
Participants assigned to the "Meat" treatment show different beliefs in whether their individual actions contribute to climate change relative to participants in the "Fruit" treatment
Test of Hypothesis:

1. Main test: Mann-Whitney U test for beliefs in whether individual actions contribute to climate change between the participants in the "meat" vs the "fruit" treatment.

Hypothesis 5 (secondary):
Participants assigned to the "Meat" treatment have a ***lower*** *average belief in the statements relative to participants in the "Fruit" treatment*
Test of Hypothesis:

1. Main test: Two-sided t-test of the $average\ belief_{i,meat}$ in the statements between the participants in the "meat" vs the "fruit" treatment.
2. Robustness test: Regressing $average\ belief_{i,meat}$ in the statements on treatment and controls discussed above and performing a t-test on the coefficient of treatment.

Missing observations:

If there are missing observations due to e.g. computer errors, the affected observations will be removed from analysis. Subjects that do not end up eating the food items will be included in the final analysis.

Significance levels

All tests will be performed at a 95% confidence level.

Power

Sample size will be fixed by expenses relative to a predetermined grant amount and the lab population available for the study. Data will be collected until either the expenses equal the predetermined grant amount or the lab population is exhausted.

The minimum detectable effect (MDE) sizes are provided for finding an effect at a 95% confidence level with a power of 80% at different sample sizes.

Potentially Achieved Samples, 1:1 allocation b/w treatments, 80% power.	MDE
150	.46
200	.40
250	.36
300	.32
350	.30
400	.28

In the paper, the MDE will also be reported in natural units, based on the observed standard error. For this, I will transform the minimum detectable effect size (MDE) by multiplying the standard error by a transformation constant of 2.8 for testing at $p<0.05$.

Final Questions

I attest that no data have been collected at the time of writing.

4.B References

Alloy, L. B. and L. Y. Abramson (1979). "Judgment of Contingency in Depressed and Nondepressed Students: Sadder but Wiser?" In: *Journal of Experimental Psychology: General* 108.4, pp. 441–485.

Bastian, B. and S. Loughnan (2017). "Resolving the Meat-Paradox: A Motivational Account of Morally Troublesome Behavior and Its Maintenance." In: *Personality and Social Psychology Review* 21.3, pp. 278–299.

Bastian, B., S. Loughnan, N. Haslam, and H. R. Radke (2012). "Don't Mind Meat? The Denial of Mind to Animals Used for Human Consumption." In: *Personality and Social Psychology Bulletin* 38.2, pp. 247–256.

Bénabou, R. and J. Tirole (2016). "Mindful Economics: The Production, Consumption, and Value of Beliefs." In: *Journal of Economic Perspectives* 30.3, pp. 141–164.

Carlson, R. W. et al. (2018). "Motivated Misremembering: Selfish Decisions Are More Generous in Hindsight." In: *PsyArXiv Preprints*.

Chew, S. H., W. Huang, and X. Zhao (2020). "Motivated False Memory." In: *Journal of Political Economy* 128.10, pp. 3913–3939.

Festinger, L. (1962). "A Theory of Cognitive Dissonance (Vol. 2)." In: *Stanford university press.*

Frederick, S. (2005). "Cognitive Reflection and Decision Making." In: *Journal of Economic Perspectives* 19.4, pp. 25–42.

Gelman, A. and J. Carlin (2014). "Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors." In: *Perspectives on Psychological Science* 9.6, pp. 641–651.

Hestermann, N., Y. Le Yaouanq, and N. Treich (2020). "An Economic Model of the Meat Paradox." In: *European Economic Review* 129, p. 103569.

Ivanova, D. et al. (2016). "Environmental Impact Assessment of Household Consumption." In: *Journal of Industrial Ecology*.

Korn, C. W. et al. (2014). "Depression Is Related to an Absence of Optimistically Biased Belief Updating about Future Life Events." In: *Psychological Medicine* 44.3, pp. 579–592.

Kunda, Z. (1990). "The Case for Motivated Reasoning." In: *Psychological Bulletin* 108.3, pp. 480–498.

Mercier, H. (2016). "The Argumentative Theory: Predictions and Empirical Evidence." In: *Trends in Cognitive Sciences* 20.9, pp. 689–700.

Mercier, H. and D. Sperber (2017). *The Enigma of Reason*. Harvard University Press.

Murdock, B. B. (1962). "The Serial Position Effect of Free Recall." In: *Journal of Experimental Psychology* 64.5, pp. 482–488.

Nickerson, R. S. (1998). "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises." In: *Review of general psychology* 2.2, p. 175.

OECD and F. a. A. O. o. t. U. Nations (2020). *OECD-FAO Agricultural Outlook 2020-2029*.

Saucet, C. and M. C. Villeval (2019). "Motivated Memory in Dictator Games." In: *Games and Economic Behavior* 117, pp. 250–275.

Schwardmann, P., E. Tripodi, and J. J. van der Weele (2019). *Self-Persuasion: Evidence from Field Experiments at Two International Debating Competitions*. SSRN Scholarly Paper ID 3490410. Rochester, NY: Social Science Research Network.

Schwardmann, P. and J. van der Weele (2019). "Deception and Self-Deception." In: *Nature Human Behaviour* 3.10 (10), pp. 1055–1061.

Shu, L. L. and M. H. Bazerman (2010). "Cognitive Barriers to Environmental Action: Problems and Solutions." In: *Oxford Handbook of Business and the Environment*, pp. 1–26.

Shu, L. L. and F. Gino (2012). "Sweeping Dishonesty under the Rug: How Unethical Actions Lead to Forgetting of Moral Rules." In: *Journal of Personality and Social Psychology* 102.6, pp. 1164–1177.

Von Hippel, W. and R. Trivers (2011). "The Evolution and Psychology of Self-Deception." In: *Behavioral and Brain Sciences* 34.1, p. 1.

Wade-Benzoni, K. A., M. Li, L. L. Thompson, and M. H. Bazerman (2007). "The Malleability of Environmentalism." In: *Analyses of Social Issues and Public Policy* 7.1, pp. 163–189.

Weller, J. A. et al. (2013). "Development and Testing of an Abbreviated Numeracy Scale: A Rasch Analysis Approach: Rasch-Based Numeracy Scale." In: *Journal of Behavioral Decision Making* 26.2, pp. 198–212.

Willett, W. et al. (2019). "Food in the Anthropocene: The EAT–Lancet Commission on Healthy Diets from Sustainable Food Systems." In: *The Lancet* 393.10170, pp. 447–492.

Wynes, S. and K. A. Nicholas (2017). "The Climate Mitigation Gap: Education and Government Recommendations Miss the Most Effective Individual Actions." In: *Environmental Research Letters* 12, p. 074024.

Zimmermann, F. (2020). "The Dynamics of Motivated Beliefs." In: *American Economic Review* 110.2, pp. 337–61.