Salience in Decision-Making
A Neuroeconomic Analysis

Dissertation

zur Erlangung des akademischen Grades
doctor rerum politicarum
(Dr. rer. pol.)

vorgelegt dem
Rat der Wirtschaftwissenschaftlichen Fakultät
der Friedrich-Schiller-Universität Jena
am 10.04.2019

von: Dipl.-Kffr. Katharina Klement
geboren am: 06.12.1986 in: Cottbus

Gutachter

1. Prof. Dr. Uwe Cantner, Friedrich-Schiller-Universität Jena
2. Prof. Dr. Wolfgang H. R. Miltner, Friedrich-Schiller-Universität Jena
3. Prof. Dr. Andreas Glöckner, Universität zu Köln

Datum der Verteidigung: 20. April 2023

Für meine Familie.
Danke für Eure Unterstützung und Euer Vertrauen.

Acknowledgement

I would like to thank Wolfgang H.R. Miltner for giving me the opportunity to write my dissertation under his supervision. He opened up to me the very interesting field of neuropsychology. He enabled the EEG experiment and always gave me valuable advice.

I am very grateful to Uwe Cantner who agreed to be my second supervisor. He was always very supportive, giving me access to the great doctoral program and helping me to find the necessary financial support to realize my dissertation.

I also want to thank Andreas Glöckner for having kindly agreed to review my dissertation.

Further thanks go to the Friedrich-Schiller-Universität for funding my dissertation with the Landesgraduiertenstipendium.

I am very grateful to Marcel Franz whose support was very important to finish my dissertation.

I want to thank my co-authors Hendrik Bruns, Elena Kantorowicz-Reznichenko, Marijane Luistro Jonsson, and Bilel Rahali for a fruitful and reliable teamwork.

I also want to thank Susanne Büchner, Angela Köhler-Saß, Wladislaw Mill, Elisa Hofmann, and Philip Dörr for being always friendly and supportive.

My sincere gratitude goes to my family who was always supportive. Without their help and their trust I would not have been able to finish my dissertation.

Deutsche Zusammenfassung

Entscheidungen bestimmen unser Leben, angefangen bei Entscheidungen mit minimalen Auswirkungen hin zu Entscheidungen mit weitreichenden Konsequenzen für uns und andere. Viele unserer Entscheidungen sind nicht rational. Dies kann einerseits dazu führen, dass sie negative Auswirkungen für den Entscheidungsträger sowie für die soziale Wohlfahrt haben. Andererseits sind Entscheidungen dadurch schwer berechenbar. Daher versuchen Wissenschaftler verschiedener Disziplinen seit langem, das menschliche Entscheidungsverhalten zu verstehen. In den letzten 20 Jahren wurden durch die verstärkte Integration von Ergebnissen und Methoden aus Ökonomie, Psychologie und Neurowissenschaften wertvolle neue Erkenntnisse und Modelle zur Erklärung von menschlichem Entscheidungsverhalten gefunden. Die Erwartungswerttheorie, welche über eine lange Zeit als Standardtheorie zur Erklärung des Entscheidungsverhaltens unter Risiko innerhalb der Wirtschaftswissenschaften diente, wurde abgelöst von Kahneman und Tversky's (1992) *Cumulative Prospect Theory*. Durch Einbeziehung verschiedener (psychologischer) Erkenntnisse über das menschliche Entscheidungsverhalten schafft diese es, einige bekannte Phänomene des Entscheidungsverhaltens zu erklären. Darüber hinaus zeigen *Dual-Process* Theorien (z.B. Evans und Stanovich, 2013; Kahnemann, 2011; Loewnstein und O'Donoghue, 2004) die Komplexität des Entscheidungsverhaltens sowie den bedeutenden Einfluss vieler Faktoren auf. Diese umfassen u.a. Emotionen, kognitive Ressourcen, Zeitdruck und Stress, das Alter, das soziale Umfeld sowie die Aufmerksamkeit.

Der Einfluss der Aufmerksamkeit steht im Fokus dieser Arbeit. Innerhalb von komplexen Umgebungen richtet sich unsere Aufmerksamkeit auf relevante Aspekte. Die Relevanz steht in engem Zusammenhang mit der Salienz der Reize. Ein sehr salienter Reiz wird von unserem Gehirn vorrangig verarbeitet, da die Salienz als Maß für die Wahrscheinlichkeit gilt, dass unser Organismus schnell eine wichtige Handlung umsetzen muss (Cooper und Knutson, 2008). Somit beeinflussen saliente Reize auch unsere Entscheidungen, in denen wir sie unbewusst überbewerten (Taylor und Thompson,1982). Die Salienz wird einerseits von sensorischen Prozessen bestimmt: ein Reiz, der sich stark von der Umgebung unterscheidet, ist sehr salient. Andererseits spielen sog. *Top-down* Prozesse eine wichtige Rolle. Dabei beeinflussen Faktoren wie Ängste und Ziele unsere Wahrnehmung der Salienz. Vorallem diese *Top-down* Prozesse begründen den engen Zusammenhang zwischen der Salienz und anderen Faktoren wie Emotionen, Stress oder dem sozialen Umfeld, die unser Entscheidungsverhalten beeinflussen. Das bedeutet, dass sich der Einfluss vieler Aspekte in der individuell wahrgenommenen Salienz widerspiegelt. Dies macht die Salienz zu einem interessanten Einflussfaktor in einer Entscheidungstheorie.

Die *Salience Theory* von Bordalo und Kollegen (2012a) versucht, menschliches Entscheidungsverhalten unter Risiko auf Basis der Salienz der möglichen Resultate zu beschreiben. Dafür entwickelten Bordalo und Kollegen eine Formel zur Messung der Salienz. Auf Grundlage der

Salienzunterschiede der möglichen Resultate zeigen sie auf, wie die wahrgenommenen Wahrscheinlichkeiten zugunsten des salientesten Resultats verzerrt werden. So kann die Theorie verschiedenste Phänomene des menschlichen Entscheidungsverhaltens erklären.

Die Berechnung der Salienz gemäß der *Salience Theory* beruht auf einer konzeptionellen Herleitung und wurde meines Wissens nach nicht experimentell überprüft. Die erste Studie dieser Arbeit, welche im zweiten Teil präsentiert wird, zielte darauf ab, diese Lücke zu schließen. Dafür wurde ein Experiment durchgeführt, in dem sich die Teilnehmer zwischen Risikolotterien entscheiden mussten. Um eine Aufmerksamkeitsverzerrung, wie sie von der *Salience Theory* vorgeschlagen wird, zu messen, wurden die Risikolotterien mit einer *Dot-Probe Task* kombiniert. Diese erlaubt anhand von Reaktionszeiten einen Rückschluss auf die Aufmerksamkeitsverzerrungen innerhalb der Lotterien. Die Daten zeigten einen signifikanten Salienzeffekt nachdem die Lotterien für 150 ms präsentiert wurden. Dieser Effekt deutet auf eine sehr zeitige Orientierung zu den salienten Auszahlungsmöglichkeiten hin. Dies ist ein potentieller Nachweis für Salienzunterschiede innerhalb von Risikolotterien gemäß der *Salience Theory*. Hingegen zeigte sich nach einer langen Lotteriepräsentation von 4000 ms kein signifikanter Salienzeffekt. Das deutet darauf hin, dass spätere strategische Aufmerksamkeitsprozesse nicht von den Salienzunterschieden beeinflusst werden. Weiterhin sind die Ergebnisse im Einklang mit dem zeitlichem Verlauf unserer Aufmerksamkeitsprozesse, in denen anfänglich die salientesten Reize verarbeitet werden, um dann schrittweise die weniger salienten Reize zu erfassen (Itti und Koch, 2000).

Um die Aufmerksamkeitsprozesse innerhalb von Risikoentscheidungen noch präziser zu erfassen und aufspalten zu können, wurden in dieser Studie auch EEG-Daten erhoben. Diese dienen als ein kontiniuierliches Maß mit einer hohen Auflösung im Millisekundenbereich. Zeitpunkte, Verläufe und Position von ERP-Komponenten können Hinweise auf Aufmerksamkeitsprozesse geben. Die zeitigen Komponenten P1 und die N1 spiegeln vorrangig sensorische Einflüsse wider und bilden Aufmerksamkeitsverzerrungen anhand der Amplitudenhöhe ab. Die P3a und P3b hingegen sind späte Komponenten, welche von *Top-down* Prozessen beeinflusst werden und auch mit der Allokation von Aufmerksamkeitsressourcen in Verbindung gebracht werden. Für alle untersuchten Komponenten zeigte die Datenanalyse keine signifikanten Salienzeffekte. Das kann einerseits einen Widerspruch zu den Ergebissen der Verhaltensdaten darstellen und deren Validität in Frage stellen. Andererseits könnten die fehlenden Effekte auch auf ein zu komplexes und für eine EEG-Studie ungeeignetes Paradigma zurückzuführen sein.

Der dritte Teil der Arbeit beschäftigt sich mit *Nudges*. Diese umfassen Interventionen von öffentlichen und privaten Insitutionen, die darauf abzielen, Entscheidungen im Sinne der Entscheidungsträger oder auch hinsichtlich der sozialen Wohlfahrt zu verbessern ohne dabei die Entscheidungsfreiheit einzuschränken. Dafür machen sie sich die verschiedenen Einflüsse auf das Entscheidungsverhalten zu Nutze. Oft verändern sie die Salienz innerhalb von Entscheidungsproblemen oder lenken die Aufmerksamkeit zum Entscheidungsproblem selbst. Damit zielen sie darauf ab, Entscheidungen im Sinne der Entscheidungsträger oder auch hinsichtlich der sozialen Wohlfahrt zu verbessern. Da sie dies meist auf unbewusster Ebene der Entscheidungsträger erreichen, stehen sie in der Kritik, diese zu manipulieren. In einer weiteren experimentellen Studie wurden die Effekte expliziter Hinweise bzgl. des Einsatzes solcher *Nudges* auf ihre Wirkung untersucht. Die Ergebnisse bestätigen bisherige Studien, in dem sie aufzeigen, dass auch transparente *Nudges* (also solche, die von diesen Hinweisen begleitet werden), die gewünschte Wirkung erzielen können. Die weiteren Ergebnisse eines fehlenden Zusammenhangs zwischen verschiedenen Transparenzhinweisen und der psychologischen Reaktanz der Teilnehmer unterstreichen diesen Befund.

Zusammenfassend betont diese Arbeit den Nutzen der Einbeziehung von Salienz bei der Betrachtung menschlichen Entscheidungsverhaltens. Die Ergebnisse zeigen auf, dass die Rolle von Salienz innerhalb von Risikoentscheidungen noch genauer untersucht werden muss, um zunächst die Gültigkeit der Salienzberechnung gemäß der *Salience Theory* und in einem weiteren Schritt auch den Nutzen der *Salience Theory* zu erfassen. Dies sollte anhand von anderen geeigneteren Paradigmen geschehen. Desweiteren könnte eine Weiterentwicklung der Salienzformel unter Einbeziehung von *Top-down* Faktoren wie Motivation oder Emotionen als bessere Grundlage zur Beschreibung von Aufmerksamkeitsverzerrungen in Risikoentscheidungen dienen. Schließlich zeigt die Arbeit auf, dass ein libertärer Paternalismus[1] beispielsweise in Form einer Salienzveränderung innerhalb des Entscheidungsproblems auch funktionieren kann, wenn die Aufmerksamkeit zusätzlich auf das Instrument (den *Nudge*) selbst gelenkt und dieses transparent eingesetzt wird.

[1] Diese Begriff wurde von Thaler und Sunstein (2003) geprägt, um Interventionen gemäß dem Prinzip der *Nudges* zu beschreiben.

Contents

List of Figures

List of Tables

List of Abbreviations

ACC Anterior cingulate cortex

EEG Electroencephalogram

EOG Electrooculogram

ERP Event-related potentials

EU ETS European Union Emissions Trading Scheme

EV Expected value

FIR Finite impulse response

fMRI Functional magnetic resonance imaging

NGO Non-governmental organization

OLS Ordinary least squares

PFC Prefrontal cortex

SOA Stimulus onset asynchrony

vmPFC Ventromedial prefrontal cortex

Part I.

Decision-Making

Despite the divergent worldviews of contemporary psychologists and economists, the two disciplines are essentially siblings separated at birth. Both have a fundamental interest in understanding human behavior. Psychology chose early on to focus on empirical questions, largely deferring attempts to formalize the resulting insights until there were sufficient data to constrain theory. By contrast, economics chose to build a foundation of formal theory, at the expense of adopting highly simplified and, ultimately, unrealistic assumptions about the processes governing human behavior.

Loewenstein, Rick, and Cohen (2008)

1. Introduction

Every day, during all our live we have to make decisions. Often small ones, sometimes big ones. We make decisions often very fast. Sometimes it is also very hard to decide. For some decisions the prospects are clear, but often we do not know with certainty what prospects will follow our choices and we have to choose under conditions of risk. Decisions may have a minimal impact, but also far-reaching consequences for us and for many others. Sometimes the reasons why we made the decisions are clear. Sometimes we do not really know why we chose an option. Maybe it just felt right.

Decision-making is a very relevant, interesting and complex research field. Various disciplines aim to understand the process behind human decision-making, to predict our decisions and finally to improve them. Even if we think we make a decision completely consciously many factors do influence our decisions without us noticing. Dual-process theories of reasoning (e.g., Evans & Stanovich, 2013; Kahneman, 2011; Loewenstein & O'Donoghue, 2004) explain these partly unconscious decision processes by two systems that guide our behavior: the affective and the deliberative system. The affective system operates very fast, autonomous and largely unconsciously. It does not require working memory. In contrast, the deliberative system is relatively slow and demanding of cognitive capacity, it operates controlled and consciously. The affective system is generally in control and can be influenced by the deliberative system through costly cognitive effort.[2] That means that our decisions are guided by fast and largely unconscious intuitions (the affective system) and (only up to a certain degree) also by conscious, deliberate and rationalized decision processes (the deliberative system) (Frith & Singer, 2008).

[2] Besides this default-interventionist view there also exist parallel-competitive forms of dual-process theories (e.g., Sloman, 1996) which suggest that both processes are always performed in parallel.

2. Brain Regions Involved in Decision-Making

When looking closer at decision-making processes, it is very useful to include findings from neuroscientific research. Since the late 1990s, when behavioral economists started to make use of neuroscientific methods in order to get new insights in decision-making processes, the new research field of neuroeconomics developed. Via important insights from neuroscience, for instance, that the human brain does not work like a homogeneous processor but instead integrates many specialised processes in different ways depending on the decision problem, neuroeconomics has inspired lots of change in economics. New economic models, like the dual-process models (see chapter 1), have been developed (Loewenstein et al., 2008).

Neuroscientific research has shown that it is a complex neural network that is involved in our decision-making processes. It is assumed that each region serves a specific function. However, regions and functions are highly interconnected (Starcke & Brand, 2012). Based on a review from Ernst and Paulus (2005), the table 2.1 provides an overview of brain regions involved in decision-making and the functions they are assumed to serve within the decision-making process.

Table 2.1.: Brain regions involved in decision-making.

Brain Region	Functions	References
Forming preferences		
Parietal cortex	Assessment of probability	Ernst et al. (2004), Platt and Glimcher (1999), Shadlen and Newsome (2001)
	Computation of probability	Dehaene, Spelke, Pinel, Stanescu, and Tsivkin (1999)
	Estimation and integration of gain/loss magnitudes	Labudda et al. (2008)
	Integration of successes and errors over time	Carter, Botvinick, and Cohen (1999)
- Superior posterior parietel cortex	Representation of value	Kahnt, Park, Haynes, and Tobler (2014)
- Inferior posterior parietel cortex	Representation of salience	Kahnt, Park, Haynes, and Tobler (2014)
Dorsolateral prefrontal cortex	Comparing contingencies of alternatives	Jonides et al. (1997)
Right dorsolateral cortex	Editing options (e.g., ignoring least attractive options, pairing options of similar values)	Cummings (1995), Dias, Robbins, and Roberts (1997)
Ventromedial prefrontal cortex	Representation of subjective value	Chib, Rangel, Shimojo, and O'Doherty (2009), Rangel, Camerer, and Montague (2008)
	Secondary emotional integration	Bechara, Damasio, and Damasio (2000), Sanfey, Hastie, Colvin, and Grafman (2003)
- Orbitofrontal cortex	Editing options (e.g., ignoring least attractive options, pairing options of similar values)	Cummings (1995), Dias, Robbins, and Roberts (1997)
Left middle and inferior frontal gyri	Reasoning	Goel, Gold, Kapur, and Houle (1998)
Limbic regions	Identification of emotional significance of a stimulus	Phillips, Drevets, Rauch, and Lane (2003)
	Generation of affective response	Phillips, Drevets, Rauch, and Lane (2003)
	Regulation of affective state	Phillips, Drevets, Rauch, and Lane (2003)
- Amygdala	Generation of somatic markers (autonomic changes such as skin conductance, blood pressure, heart rate)	Bechara (2004)
- Anterior cingulate cortex	Processes of uncertainty	Critchley, Mathias, and Dolan (2001), Elliott, Rees, and Dolan (1999)
	Representation of salience	Kahnt, Park, Haynes, and Tobler (2014)
Insula	Generation of somatic markers (autonomic changes such as skin conductance, blood pressure, heart rate)	Bechara (2004)
Ventral striatum	Emotional intensity (salience) of stimuli	Zink, Pagnoni, Martin-Skurski, Chappelow, and Berns (2004)
- Nucleus accumbens	Representation of subjective value	Rangel, Camerer, and Montague (2008), Montague and Berns (2002)
	Representation of salience	Cooper and Knutson (2008)
Execution of actions		
Anterior cingulate cortex	Error monitoring	Mathalon, Whitfield, and Ford (2003)
	Conflict detection	Gehring and Knight (2000)
Lateral prefrontal cortex	Monitoring of an action	Mathalon, Whitfield, and Ford (2003)
	Guiding compensatory actions	Gehring and Knight (2000)
Dorsolateral prefrontal cortex	Monitoring goal-directed behavior	Jonides et al. (1997)
Nucleus accumbens	Modulation of motivational aspects of an action* **	Ernst et al. (2004), Knutson, Fong, Adams, Varner, and Hommer (2001), Salamone and Correa (2002)
Amygdala	Modulation of motivational aspects of an action	Breiter and Rosen (1999)
Ventrolateral prefrontal cortex	Modulation of motivational aspects of an action	Taylor et al. (2004)
Insula	Rejection of unfair offers	Sanfey, Hastie, Colvin, and Grafman (2003)
Experiencing the outcome		
Ventral striatum	Generation of difference signal between expected and actual outcome value	Pagnoni, Zink, Montague, and Berns (2002), McClure, Berns, and Montague (2003)
Orbitofrontal cortex	Generation of difference signal between expected and actual outcome value	O'Doherty, Dayan, Friston, Critchley, and Dolan (2003), McClure, Berns, and Montague (2003)
Amygdala	Associative learning	Baxter and Murray (2002), Cardinal, Parkinson, Hall, and Everitt (2002), Gabriel, Burhans, and Kashef (2003), Salamone and Correa (2002), Schoenbaum and Setlow (2003)
Nucleus accumbens	Associative learning	Baxter and Murray (2002), Cardinal, Parkinson, Hall, and Everitt (2002), Gabriel, Burhans, and Kashef (2003), Salamone and Correa (2002), Schoenbaum and Setlow 2003
Medial prefrontal cortex	Feedback processes	Knutson, Fong, Bennett, Adams, and Hommer (2003)
Ventral medial prefrontal cortex		
-Orbitofrontal cortex	Assessment of pleasurability	Mitterschiffthaler et al. (2003)
	Tracking of rewarding outcomes	Knutson, Fong, Bennett, Adams, and Hommer (2003)

Notes: *Here, motivation is defined "as the determinant of the direction and energy of an action" (Ernst & Paulus, 2005, p. 599). **Ernst and Paulus (2005) mention that the differentiation between motivation and arousal here is still not clear.

3. Factors Influencing our Decisions

The following chapter explores factors that guide our decisions. This is important in order to understand the irrational biases of our decisions (common examples of decision biases are given in chapter 5).

3.1. Emotions

When thinking of emotions influencing decisions the common idea is that they bias our decisions in an unfavorable manner. While this is certainly true for some decision situations it does not capture the extensive influence of emotions on human decisions. Neuroscientific evidence demonstrates that sound and rational choices depend on prior accurate emotional processing. Bechara and Damasio (2005) propose emotions[3] as major factor moderating the interaction between environmental conditions and decision processes. In their somatic marker hypothesis they describe the emotional system as influencing decisions via marker signals that arise in bioregulatory processes (e.g., endocrine release, heart rate modification, smooth muscle contractions or the release of neurotransmitters like dopamine, serotonin or noradrenaline). These may result in emotions and feelings and influence our choices on a conscious as well as on an unconscious level. Thus, our emotional system provides important implicit and explicit knowledge that supports us with making fast and advantageous decisions. The somatic marker hypothesis is based on the observation that patients with lesions of the ventromedial prefrontal cortex (vmPFC) develop difficulties in making advantageous real-life decisions, such as planning their work day or choosing friends or activities. Yet their intellect remains on a normal level. The patients often choose against their best interest resulting in financial losses or losses of family and friends. They are also unable to learn from previous mistakes. In addition to these problems in decision-making they show difficulties in expressing emotions and experiencing feelings (Bechara & Damasio, 2005). According to the somatic marker hypothesis "a defect in an emotional mechanism that rapidly signals the prospective consequences of an action, and accordingly assists in the selection of an advantageous response option" (Bechara & Damasio, 2005, p. 339) leads to the patients' inability to decide advantageously in real-life.

[3] According to Damasio (2003) an emotion may be defined as the ensemble of changes in body and brain states that are triggered by a brain system that responds to particular contents of our perception, which may be actual or recalled, relative to a specific object or event. The modifications of body states may affect internal factors like endocrine release, heart rate, and smooth muscle contraction, as well as external factors such as posture, facial expression, or freezing (Bechara & Damasio, 2005).

3.2. Cognitive Resource Scarcity

When facing complex decisions we need to be attentive, process lots of information, and evaluate the trade-offs between alternatives. This deliberative decision process requires cognitive resources which are naturally limited. In addition, we often have to pay attention simultaneously to multiple tasks or choices. Also having too many options for one choice may deplete cognitive resources. A growing body of literature suggests that low cognitive capacity[4] leads to poorer reasoning and math performance, poorer probability judgements, more risk-aversion, more random choices, a higher susceptibility to anchoring and defaults, and less sophisticated strategies in games (Burks, Carpenter, Goette, & Rustichini, 2009; Sprenger et al., 2011; Deck & Jahedi, 2015; Bergman, Ellingsen, Johannesson, & Svensson, 2010; Altmann, Grunewald, & Radbruch, 2018; Franco-Watkins, Rickard, & Pashler, 2010).

Cognitive resource scarcity may also be increased by scarcity of other resources, namely fincancial resources, hunger, or other facets of poverty (Schilbach, Schofield, & Mullainathan, 2016). This may partiallly explain the link between a low household income and a high propensity to stick with defaults (Brown, Farrell, & Weisbenner, 2012; Bhargava, Loewenstein, & Sydnor, 2017).

3.3. Time Pressure and Stress

Natural or artificial time constraints are common to decisions. Decision situations under high time pressure include, e.g., bargaining, auctions and negotiations, as well as medical emergencies or human conflicts. In other situations without direct time limits the opportunity costs of the decision time may constitute indirect constraints. Time pressure thus represents a very relevant influence on decision making. Research suggests that time pressure significantly reduces overall decision quality (de Paola & Gioia, 2016; Dambacher & Hübner, 2015; Kocher & Sutter, 2006). Reasons are an adjustment of individual decision criteria where boundaries are lowered under time pressure, but also a reduced integration of perceptual evidence[5] (Dambacher & Hübner, 2015). Another interesting effect of time pressure are altered risk preferences. Kocher, Pahlke, and Trautmann (2013) experimentally show that time pressure may have differential effects on risk attitudes in the gain, loss, and mixed domain: in the gain domain, i.e., in risky prospects with only positive outcomes, time pressure has no effect. In contrast, in the loss domain, i.e., in risky prospects with only negative outcomes, time pressure seems to elicite risk-aversion[6]. In mixed prospects effects hold depending on the framing, with increased loss aversion and gain seeking. However, the influence seems to be complex and results are somewhat inconclusive. Young, Goodie, Hall, and Wu (2012), for instance, find heightened risk seeking under time pressure in the gain and loss domain.

[4]Cognitive capacity may be low due to idiographic differences, or due to a high cognitive load.

[5]The integration of perceptual evidence is reduced under time pressure because the ability to focus perceptional processing on response-relevant stimuli is low after stimulus onset, i.e.,the presentation of the decision problem, and increases with time (Dambacher & Hübner, 2015).

[6]Kocher et al. (2013) point out that the effect in the loss domain only holds for decision-makers that are risk seeking in the loss domain without time pressure.

To investigate the influence of stress in general on risky decision making Buckert, Schwieren, Kudielka, and Fiebach (2014) induced mild psychosocial stress using the Trier Social Stress Test for groups (von Dawans, Kirschbaum, & Heinrichs, 2011). To measure the individual response to acute stress the cortisol level was used. They found heightened risk seeking in the gain domain only for subjects with a robust cortisol response to acute stress. They conclude that cortisol is an important mediator of this effect. If the concentration is high under acute stress cortisol (the primary glucocorticoid in humans) binds to glucocorticoid receptors (de Kloet, 2004). Since these exist everywhere in the brain, cortisol can have modulating influences on emotional and cognitive processes (Buckert et al., 2014). In addition, the catecholamines dopamine, adrenaline and noradrenaline also lead to metabolic changes in the brain during acute stress exposure. Neuroimaging studies show a decreased activity in the prefrontal cortex (PFC) (Ossewaarde et al., 2011; Schwabe, Tegenthoff, Höffken, & Wolf, 2012) which is involved in comparing contingencies of alternatives and monitoring goal-directed behavior. Moreover, stress induces changes in the ventral striatum, limbic regions and the orbitofrontal cortex. These could lead to increased reward sensitivity, a disrupted balance between automated emotional responses and deliberate calculative responses as well as reduced feedback-processing abilities (Starcke & Brand, 2012).

3.4. Age

Heightened risk-taking in adolescence is a well-known phenomenon. Due to the sometimes fatal consequences it is of deep interest to understand the reasons behind these decisions. Yet, the correlation between age and risk-taking behavior is still not clear. Based on a meta-analysis on age differences in risky decision-making Defoe, Dubas, Figner, and van Aken (2015) suggest that risk-taking levels are highest in childhood and decline thereafter. Neurodevelopmental imbalance models (Albert & Steinberg, 2011; Somerville, Jones, & Casey, 2010) explain this correlation with the development of brain regions involved in the affective and the deliberative system. Whereas the prefrontal cortex, which is responsible for cognitive control, develops linearly with age and begins to stabilize by adolescence, subcortical brain regions governing the affective system develop relatively faster. Thus, especially in emotionally charged situations the affective system is in control. These models also hypothesize that the affective system is hypersensitive in adolescence and risk-taking therefore is highest during that stage of life. But neurological evidence on that is not clear and findings of risky behavior in children compared to adolescents in the laboratory contradict this hypothesis (Defoe et al., 2015). Defoe et al. (2015) argue that situational factors may provide a better explanation for the heightened risk-taking in adolescents compared to children in the real world. That is, adolescents face much more opportunities to engage in risky behavior than children since these are monitored more closely and are not allowed, e.g., to drive a car or to buy alcohol.

3.5. Social Influence

Many studies have shown that our choices and attitudes are influenced by others (e.g., Carter, Bowling, Reeck, & Huettel, 2012; Chung, Christopoulos, King-Casas, Ball, & Chiu, 2015; Zaki, Schirmer, & Mitchell, 2011). Since social information is frequent, salient and a good indicator of necessary future behavior (Carter et al., 2012), using social information is very beneficial in order to learn about environmental contingencies and to adapt to a dynamic environment (Rendell et al., 2010; Lee, 2008; Danchin, Giraldeau, Valone, & Wagner, 2004). Therefore, our ability to predict others' intentions and beliefs, also referred to as theory of mind or mental state attribution, is central (Singer & Tusche, 2014; Olsson & Ochsner, 2008). In addition, identifying relevant actors (Kameda, Ohtsubo, & Takezawa, 1997; Stasser, Stewart, & Wittenbaum, 1995), comparing ourselves to more similar others (Festinger, 1954) and distinguishing strategic actions from random behavior (Yoshida, Seymour, Friston, & Dolan, 2010) are important social capacities. In line with that, Chung et al. (2015) found that participants' risky decisions were distorted towards others' choices, choices were not influenced by random (computer-generated) choices and choices were more likely to be influenced if the risk preferences of others' choices corresponded to participants' preferences. Functional magnetic resonance imaging (fMRI) results showed a greater sensitivity in the ventromedial prefrontal cortex (vmPFC)[7] in subjects that were more likely to conform with others' choices. They interpret that as higher value that participants assign to the options that were selected by others. Zaki et al. (2011) showed that participants were influenced by others' choices when rating the attractiveness of faces. Based on significantly greater activations in bilateral nucleus accumbens[8] and orbitofrontal cortex[9] for faces that were rated as more attractive than without social influence, they similarly conclude that the assigned subjective value changed through social influence.

Not only others' choices and attitudes but already their mere presence influences our decisions. This so-called "social facilitation" (Allport, 1924) enhances the performance of well-learned behavior and impedes the acquisition of new behavior (Zajonc, 1965). Reasons behind this effect are still unclear. Some theories suggest that social facilitation increases the psychological arousal and the motivation to perform well (Zajonc, 1965; Sanders & Baron, 1975; Harkins, 2006). Other theories posit that others' presence affects the attention by restricting the focus to relevant stimuli in well-learned tasks but impeding the recognition of all relevant stimuli in new tasks (Baron, 1986; Sharma, Booth, Brown, & Huguet, 2010; Muller, Atzeni, & Butera, 2004).

In order to find evidence for either motivational or attentional theories Monfardini et al. (2015) investigated the activation of brain networks regulating motivation and attention in monkeys. They found an increased activation in the frontoparietal attention network, more precisely in the prefrontal cortex. This area is believed to be responsible for filtering distractors (Suzuki & Gottlieb, 2013). Monfardini et al. (2015) argue that these findings prove the enhanced

[7]The vmPFC is a brain region that seems to be responsible for encoding subjective value during value-based decision-making (Chib, Rangel, Shimojo, & O'Doherty, 2009; Rangel, Camerer, & Montague, 2008).

[8]The nucleus accumbens, a region in the ventral striatum, also belongs to a network that represents subjective values (Montague & Berns, 2002; Rangel et al., 2008).

[9]The orbitofrontal cortex is a region within the vmPFC.

attention through others' presence. However, motivational effects still may be possible within other contexts.

The influence of others seems to be especially relevant in adolescence. Research suggests that adolescents behave more risky in the presence of peers (Gardner & Steinberg, 2005) (see section 3.4). In their study on the peer effect on adolescent risk taking, Chein, Albert, O'Brien, Uckert, and Steinberg (2011) found an increased activation of adolescents' brain regions associated with the prediction and valuation of reward, including the ventral striatum and the orbitofrontal cortex, in the presence of peers. These results indicate a heightened sensitivity to rewards among adolescents in the presence of others.

3.6. Attention and Salience

Every day we face complex visual scenes. Within these we do not perceive every detail with the same intensity. Instead our mind focuses on relevant aspects that are processed preferentially (Itti and Koch, 2000). The intensity with which we perceive the diverse visual signals is determined by their salience.[10] The salience increases in the signal's magnitude relative to it's context. Stimuli that are very different or unusual from it's context are highly salient and will capture our attention. This is important since salient stimuli "increase the chance an organism will need to make an important behavioral response in the near future" (Cooper & Knutson, 2008, p. 539). In addition to this sensory information that determines the salience via the *bottom-up process* attentional modulation also affects salience via the *top-down process* (Treue, 2003; Itti & Koch, 2000). This modulation embraces factors like goals and drives, experiences and memories as well as the current psychological state (Uddin, 2015). This can be easily shown by the different focus that will be paid to a little spider within a nice garden view by spider phobics compared to healthy people. Whereas the salience based on the sensory infomation - the spider is not very different from it's context - is rather low, the salience is highly increased by the fear of spiders. Due to the integration of multiple factors through these bottom-up and top-down processes we are not always aware of what we perceive as salient (Critchley & Harrison, 2013).

Importantly, salient attributes will be disproportionately weighted in subsequent judgements (Taylor & Thompson, 1982). This has been suggested in psychological models of decision-making (e.g., Busemeyer & Townsend, 1993; Roe, Busemeyer, & Townsend, 2001; Rieskamp, Busemeyer, & Mellers, 2006) as well as in behavioral studies (e.g., Shimojo, Simion, Shimojo, & Scheier, 2003; Armel, Beaumel, & Rangel, 2008; Krajbich, Armel, & Rangel, 2010). Further studies investigated the influence of salience on risky decisions. Weber and Kirsner (1997) manipulated perceptual and motivational salience and found both to affect the evaluation of risky lotteries. Schwager and Rothermund (2013) induced attentional biases by manipulating affective processing. They suggest that subsequent risky decisions were affected via increased salience of win- and loss-colors in a gambling task with ambivalent color stimuli. Also, recent fMRI studies show the importance of visuospatial attention in (risky) decision-making (Studer, Cen, & Walsh, 2014) and

[10] The concept of salience is not restricted to visual stimuli. Here, salience is explained exemplarily with respect to the sense of sight. However, in many decision situations the visual sense is very relevant.

hint on a causal role of the neural salience network[11] in decision-making (Chand and Dhamala, 2016). Further evidence even suggests that some neural signals that were previously interpreted as reflecting value instead reflect salience (Kahnt, Park, Haynes, & Tobler, 2014; Kahnt & Tobler, 2013; Cooper & Knutson, 2008).

3.7. Salience as Special Influence Factor

This chapter has made clear that human decision-making is strongly influenced by multiple factors. Models of decision-making aim at capturing human decisions as best as possible. At the same time they need to be straightforward and not too complex in order to apply them properly. Therefore, integrating only a few influence factors which, up to a certain degree, can reflect the influence of diverse other factors is very useful in order to build a good model of human decision-making.

Interestingly, all influence factors presented in this chapter are closely interrelated. Especially salience and emotions are interrelated with each other, and with other influence factors. The following list captures some of these interrelations based on research findings provided in this chapter.

* Different *emotions* may arise due to different *salient* aspects within a decision problem.
* Via the top-down process *emotions* may influence the *salience* of certain stimuli.
* *Cognitive resource scarcity* as well as *time pressure* and *stress* increase the effect of *salience* on subsequent decisions and may also modulate *emotions*.
* Due to the differential development of brain regions with *age*, decision situations may evoke different *emotions* over age.
* *Social presence* might affect decisions by restricting attention to *salient* stimuli, thus enhancing the performance in well-learned tasks but diminishing it in new ones.

The interrelation of salience and many other influence factors - which may be directly interrelated but also indirectly via emotions - provides an interesting argument for the consideration of salience when investigating decision-making. Here, salience may serve as special influence factor that captures the influence of many other factors.

[11] The salience network embraces brain regions having a central role in detecting behaviorally relevant stimuli and coordinating neural resources (Uddin, 2015).

4. Salience Theory of Risky Choice

Until recently, attention has only played a minor role in most theories on decision-making.

Expected Utility Theory, first proposed by Daniel Bernoulli in 1738, has been the standard theory of risky decisions in economics for a long time. According to Expected Utility Theory, the decision-maker compares the available alternatives by their expected utility, i.e., the sum of all possible outcomes weighted by their probabilities. Although the maximization of the expected utility seems to be rational, it does not capture any other influence factors like those presented in the last chapter. Consequently, Expected Utility Theory cannot explain many decision biases that commonly occur in risky decisions. Famous examples are the Allais paradoxes and preference reversals. Allais (1953) proposed that adding a common consequence[12] to a risky choice may change preferences from risk averse to risk seeking. This has been experimentally confirmed many times. For example, Kahneman and Tversky (1979) presented the following choices (1, 2) between two options (A, B) to experimental subjects:

$$(1) \quad \text{A=}\begin{cases} \$2500 & p^{*}= 0.33 \\ \$0 & p = 0.01 \\ \$2400 & p = 0.66 \end{cases} \qquad \text{B=}\begin{cases} \$2400 & p = 1.00 \end{cases}$$

$$(2) \quad \text{A=}\begin{cases} \$2500 & p = 0.33 \\ \$0 & p = 0.67 \end{cases} \qquad \text{B=}\begin{cases} \$2400 & p = 0.34 \\ \$0 & p = 0.66 \end{cases}$$

* p = probability

Both choice problems differ only in the common consequence z with z=2400 in choice 1 and z=0 in choice 2:

$$A(z) = \begin{cases} \$2500 & p = 0.33 \\ \$0 & p = 0.01 \\ \$z & p = 0.66 \end{cases} \qquad \text{B(z)=}\begin{cases} \$2400 & p = 0.34 \\ \$z & p = 0.66 \end{cases}$$

For choice 1 (z=2400), most subjects choose B over A, thus acting risk averse. In contrast, with the common consequence z=0 (choice 2) the majority prefers A over B, acting risk seeking. This change of preferences violates the independence axiom of Expected Utility Theory, which states that preferences are stable irrespective of the presence of irrelevant alternatives.

In 1979, Kahneman and Tversky (1979) introduced Prospect Theory, which was later further developed to Cumulative Prospect Theory (Tversky & Kahneman, 1992). Today, it is still

[12] A common consequence is a potential outcome occuring with a certain probability that is identical for all options.

the most prominent theory of choice under risk. It is based on Expected Utility Theory and incorporates several typical decision biases, namely different risk preferences for gains (risk aversion) and losses (risk seeking), loss aversion, payoff values that are relative to a reference point, and the overweighting of unlikely extreme events. Tversky and Kahneman (1992) explain the latter by increased perceptual salience. Thereby Cumulatvie Prospect Theory provides an explanation for the changing risk preferences in the Allais paradox. In choice 1, the unlikely (p=0.01) extreme outcome of 0 is overweighted and elicits risk aversion (choice B). In choice 2, the extra probability of 0.01 is included in the average probability of 0.67 and is therefore not overweighted. This induces risk seeking (choice A).

In 2012, Bordalo, Gennaioli and Shleifer proposed their Salience Theory of choice under risk. The theory is based on the idea that decision-makers do not fully consider all information that is available, but overemphasize the information their minds focus on. More precisely, Salience Theory explains the natural systematic instability of risky choices by the salience of different payoffs, i.e., the different amount of the decision-makers attention drawn by the different payoffs. Here, the salience is determined by the contrasting features of the alternatives. Bordalo and colleagues assume that the decision-maker focuses on salient payoffs. Salient payoffs are those which differ most from the payoffs of the alternative option. The decision-maker unconsciously inflates their weights when making a choice. The true probabilities are replaced by these decision weights and thus distorted in favor of salient payoffs. According to Bordalo and colleagues the salience of payoffs is based on all feasable payoff combinations across the options (referred to as states). The decision-makers perceive the differences in payoffs across these states as signals with different intensities. That means, the intensity or salience is always the same for a state, consisting of one payoff of each option. Bordalo and colleagues propose three conditions to measure this intensity (i.e., the salience). The *ordering* property captures the magnitude of the signal: The salience increases in the distance between the best payoff of one option[13] and the worst payoff of the alternative option. The influence of the context is captured by two further properties, *diminishing sensitivity* and *reflection*: The perceived intensity of the payoffs decreases with increasing average payoffs[14], and is determined rather by the magnitude of payoffs than the gain or loss domain.[15]

Salience Theory may explain changing risk preferences in the Allais paradox. Looking at the choice problem above, in choice 1 (z=2400) the outcome of 0 is most salient within option A and feels much lower than the sure payoff of 2400. This causes risk aversion (choice B). In choice 2 (z=0), both options have the same downside of 0. Here, the upside of 2500 in the riskier option A and the downside of 0 in option B are most salient. This induces risk seeking (choice A).

Furthermore, Salience Theory can explain preference reversals and context effects (such as

[13]Bordalo and colleagues refer to each option as *lottery*. Here the term *option* is used instead and the decision-tasks between options are referred to as *lotteries*.

[14]Just as in Weber's law of diminishing sensitivity of a perceived change relative to the actual change in a physical stimulus where for example "a change in luminosity is perceived less intensely if it occurs at a higher luminosity level, the [...][decision-maker] perceives less intensely payoff differences occurring at high (absolute) payoff levels" (Bordalo, Gennaioli, & Shleifer, 2012a, p. 1254).

[15]A more detailed explanation of salience in lotteries according to Salience Theory is provided in chapter 8, section 8.1.

decoy effects, see chapter 5), decision biases that Cumulative Prospect Theory cannot rationalize (for a detailed explanation see Bordalo et al. (2012a, 1265ff.)).[16]

5. Decision Biases

There are many well known decision biases, that is, decisions commonly deviating from rational choice, e.g., through information that is irrelevant to the decision problem. These biases arise, among others, due to the influence factors outlined in chapter 3.

An important bias is the *framing effect*, which was termed by Tversky and Kahneman (1981) and is a key aspect of Prospect Theory. In an experiment, Tversky and Kahneman (1981) asked subjects which program they favor to combat a disease that is expected to kill 600 people. When framing the options as follows, the majority chose the risk averse Program A:

> If Program A is adopted 200 people will be saved.
> If Program B is adopted there is 1/3 probability that 600 people will be saved,
> and 2/3 probability that no people will be saved.

In contrast, the following framing of the same programs elicited risk taking (Program D) for the majority of subjects:

> If Program C is adopted 400 people will die.
> If Program D is adopted there is 1/3 probability that no one will die,
> and 2/3 probability that 600 people will die.

Prospect Theory explains the framing effect with the differing risk preferences for gains and losses. Programs A and B are framed as gains (in terms of saved lives), eliciting risk averse choices. Programs C and D are framed as losses (people that will die), eliciting risk seeking. Salience Theory provides a different explanation for the framing effect. According to Bordalo, Gennaioli, and Shleifer (2015) the different framing changes the outcome that is perceived as salient. In the gain frame the most salient outcome is the one where no one will be saved, causing risk aversion. In the loss frame the most salient outcome is the one where no one will die, causing risk seeking.[17] This is in line with research findings. Framing changes the salient information of

[16] For a detailed description and estimation of decision weights, the integration in and assumptions about value functions, etc., proposed by Salience Theory, please see Bordalo et al. (2012a), since this is out of the scope of this dissertation.

[17] Bordalo et al. (2012a) point out that this reflection of risk attitudes only holds when the decision-maker's value function is linear. In line with that, Laury and Holt (2005) found that the reflection effect diminishes with payoff magnitude.

the decision problem (see, e.g., Shafir, 1993). Consequently, different emotions may be evoked by different framing of choices. In an fMRI study, de Martino, Kumaran, Seymour, and Dolan (2006) found an increased activation in the amygdala when subjects responded in accordance with the framing effect. The amygdala is associated with rapid orientation towards biologically salient stimuli and related emotional responses (LeDoux, 2000). de Martino et al. (2006) interprete the increased activation as processing of the emotional type of contextual information.

Another interesting bias is the *endowment effect.* This effect describes the willingness to trade an endowed good against something else. Usually, subjects refrain from trading a present against a good with similar value. In other experiments, subjects also state selling prices for endowments that are much higher than their buying prices for the same good. According to Prospect Theory, this behavior may be explained by loss aversion. Bordalo, Gennaioli, and Shleifer (2012c) propose an explanation based on the differing salience of information on the goods. The decision-maker compares the endowment with his status-quo of having nothing, focusing on the endowments best attributes. The overweighting of these results in an overvaluation of the endowed good. However, when the decision-maker is asked to trade the endowment against a new good, now the evaluation of the new good is determined by the contrast between the endowment and the new good. This contrast may render the disadvantages of the new good salient, resulting in an underevaluation of it.

According to Bordalo et al. (2015) a similar comparative evaluation may explain the *decoy effect.* Here, a third option is introduced to a pairwise choice. If this option is clearly worse than one of the original options, the good attributes of the dominating option become salient and the demand for it increases.

Our decisions are also strongly influenced by information that we think of prior to the decision. Thinking of a specific number that is completely irrelevant to a subsequent decision problem subconsciously serves as *anchor* for our decision. In an experiment Bergman et al. (2010) let subjects state their willingness to pay for several products. At first, subjects had to write down the last two digits of their Social Security Number. Then they had to decide whether they would buy the products at an equal price to that number. Bergman et al. (2010) found that subjects willingness to pay typically increases with this irrelevant number. Their findings also suggest that the anchoring effect is reduced but still positive for subjects with high cognitive ability[18].

In decisions where risk probabilities are not given and decision-makers have to assess risk probabilities themselves, the *availability bias*, which is closely related to accessibility and salience, often distorts risk assessments. That is, when people recently experienced, for instance, a flood, or they know other people who did, they are more likely to purchase insurance. As memories fade, insurance purchases decline (Thaler & Sunstein, 2008).

Another well known phenomenon is the *status quo bias.* It describes the general tendency of people to stick with their current situation. A good example here are diverse subscriptions that are automatically renewed each year unless they are actively cancelled. People often tend to keep those even they do not need them anymore. Reasons behind the status quo bias are inertia and

[18] Cognitive ability was measured by a standard psychometric test of general intelligence developed by the Swedish psychometric company Assessio, which they took from Sjöberg, Sjöberg, and Forssén (2006).

also a lack of attention (Thaler & Sunstein, 2008).

Often these decision biases may lead to disadvantagous outcomes. Therefore, politics aim at improving decisions. Here, nudges are a form of intervention that has become popular. The next chapter provides further information on nudges.

6. Nudges - Improving Decisions

Nudges are a form of interventions by public and private institutions that aim at improving people's decisions while maintaining their freedom of choice and without changing their economic incentives (Thaler & Sunstein, 2008). Thaler and Sunstein (2003) describe such interventions as libertarian paternalism. Nudges can be pro-self, nudging people towards making better decisions for themselves, and pro-social, aiming at individual behavior that increases social welfare (Hagman, Andersson, Västfjäll, & Tinghög, 2015).

Nudges focus on automatic decision processes and make use of decision biases (like those described in the last chapter) that often arise through small and irrelevant details. They embrace a diverse set of instruments. Often these instruments work by directing the attention of the decision-makers in a particular direction, thus altering the salience of information regarding the decision problem.

A simple example of a nudge is to place healthy food at eye level. This enhances its salience and thus increases the probability of people to buy it. Also, giving a specific anchor prior to a decision serves as a nudge. When charities ask for a donation, they typically offer a range of options. These serve as anchor and will influence how much people give. Regarding the availability bias, simple nudges are to remind people of dangerous incidents in order to increase their willingness to take precautionary measures (Thaler & Sunstein, 2008). Commonly used nudges direct the attention to specific decisions, e.g., by providing additional information, (Kling, Mullainathan, Shafir, Vermeulen, & Wrobel, 2012; Karlan, McConnell, Mullainathan, & Zinman, 2016) or force people to choose actively (Carroll, Choi, Laibson, Madrian, & Metrick, 2009; Stutzer, Goette, & Zehnder, 2011). Very powerful nudges are default options. They make use of the status quo bias. People who do not want to pay more attention to all possible options of the required decision will stick to the default. Furthermore, some people do understand default options as implicit recommendation (Thaler & Sunstein, 2008). However, even if decision-makers do not choose the default, they may also work as anchor, thus still biasing the decisions towards the default option. Regarding salience differences within a decision problem, a default option changes these by increasing the salience of the respective option.

Altmann et al. (2018) experimentally investigated how directing the attention to a decision and forcing to choose actively affect the decision quality. They found both nudges to improve the quality. However, they point out that these interventions may reduce decision qualities of other tasks when cognitive capacity is scarce. Therefore, different decision domains that compete for cognitive resources should be considered for the evaluation of nudges. Other studies also point at possible risks of nudges by showing that they can even have the opposite effect of the intended one resulting in a decrease of people's and social welfare (Murooka & Schwarz, 2018). They may aswell inhibit individual (Caplin & Martin, 2012; de Haan & Linde, 2018) and social learning (Carlin, Gervais, & Manso, 2013) about the decision tasks. Finally, nudges are criticized to be manipulative due to there subtle nature.

7. Subject of Own Research

This work aims at investigating the role of salience within decision-making. Therefore, part I presented a broad overview of human decision-making and important influence factors. Here the possible role of salience as variable that reflects the influence of many other factors was conceptually developed. Salience Theory (Bordalo et al., 2012a) describes human decision-making on the basis of salience differences within a decision problem. Looking at choices under risk Bordalo et al. (2012a) present a formula identifying those salience differences among payoffs. Salience Theory calculates decision weights based on these salience differences that distort decisions in favor of salient payoffs. To my knowdledge, this formula has been theoretically developed and has not been experimentally validated before. Since Salience Theory builds on this salience formula it is important to experimentally investigate its validity. Part II presents an experiment that was designed for this purpose. Here, reaction times from a dot-probe task were used in order to identify attentional biases in risky lotteries on the basis of behavioral data. Furthermore, EEG data was used in order to more closely examine attentional processes underlying this possible salience effect. Part III is devoted to nudges. These interventions often work by altering the salience of specific options of decisions or by directing the attention to the decision task itself. Thus, they influence decisions at least partly on an unconscious level. Consequently, they are subject to criticism. If people do not notice the interventions they may feel betrayed. Therefore, we experimentally investigated how providing additional information about the nudge itself affects the efficacy of the nudge.

Part II.

Salience in Risky Decisions

8. The first glance in risky decisions - Can Salience Theory predict it? A Behavioral Validation of Salience in Risky Lotteries as Defined by Salience Theory

8.1. Introduction

Salient stimuli draw our attention. Subconsciously they also influence our decisions (Taylor and Thompson, 1982). Therefore, studies from diverse disciplines suggest that salience should be taken into account when investigating decision-making in general as well as decision-making under risk (see chapter 3.6). In 2012, Bordalo and colleagues proposed their Salience Theory (see chapter 4) which incorporates salience as crucial factor in risky decisions. Looking at choices among two risky options they suggest that, within each option, one payoff is most salient and therefore perceived as having a higher probability (compared to the given probability). Consequently, decisions are distorted in favor of these salient payoffs. According to Bordalo and colleagues this may explain many deviations from rational choice including well known phenomena like the Allais paradoxes and preference reversals (whereas the latter cannot be rationalized by Prospect Theory (Bordalo et al., 2012a)).

Bordalo et al. (2012a) suggest three conditions to measure the salience of payoffs, namely, ordering, diminishing sensitivity and reflection. The salience is based on the properties of all feasable payoff combinations across the options, referred to as states. The *ordering* property captures the magnitude of the signal: The salience increases in the distance between the best payoff of one option[19] and the worst payoff of the alternative option. The influence of the context is captured by two further properties, *diminishing sensitivity* and *reflection*: The perceived intensity of the payoffs decreases with increasing average payoffs[20], and is determined rather by the magnitude of payoffs than the gain or loss domain. These conditions are summarized by the formula:

$$\sigma\left(x_s^i, x_s^{-i}\right) = \frac{\left|x_s^i - x_s^{-i}\right|}{\left|x_s^i\right| + \left|x_s^{-i}\right| + \theta'}$$

[19] Bordalo and colleagues refer to each option as *lottery*. Here the term *option* is used instead and the decision-tasks between options are referred to as *lotteries*.

[20] Just as in Weber's law of diminishing sensitivity of a perceived change relative to the actual change in a physical stimulus where for example "a change in luminosity is perceived less intensely if it occurs at a higher luminosity level, the [...][decision-maker] perceives less intensely payoff differences occurring at high (absolute) payoff levels" (Bordalo et al., 2012a, p. 1254).

The salience σ is always determined for each state s, i.e., for all feasable payoff combinations (x_s^i, x_s^{-i}) across the options $O_i, i = 1, 2$, resulting in payoffs across options that are equally salient. θ captures the "cognitive limit to the resolution of payoff magnitude when a payoff approaches zero" (Bordalo, Gennaioli, & Shleifer, 2012b, p. 28). Bordalo et al. (2012b) propose the value of $\theta \sim 0.1$ as useful reference.

Importantly, due to the proposed determination of the salience of payoffs based on properties of payoff combinations across options, Salience Theory suggests that there are salience differences *within* each option, not *across* options. That means, payoffs that are most salient are always present in all options. Therefore, salience differences affect the assessment of probabilities within an option. Depending on the quality of the salient payoff, it may result in a higher evaluation[21] (if the upside is salient) or a lower evaluation (if the downside is salient) of the option (for an illustration, see figure 8.2 where most salient payoffs are highlighted in red within the left and the right option, or see again the explanation of the Allais paradox accoding to Salience Theory presented in chapter 4).

To my knowledge, this theoretical approach to identify salient payoffs has not been empirically validated before and opens up the question whether the decision-maker's attention in risky lotteries is biased towards the payoffs that Salience Theory suggests as being salient. Therefore, the aim of this study was to experimentally examine the validity of salient payoffs in lotteries as proposed by Salience Theory.

To approach this question the dot-probe task was used. It is a common psychological measure to assess selective-attentional processing (Frewen, Dozois, Joanisse, and Neufeld, 2008). In the standard dot-probe task, originally developed by MacLeod, Mathews, and Tata (1986), two stimuli (e.g., words, pictures), of which one is neutral and one is valent[22] are simultaneously presented for a short time at different locations of a screen. Immediately afterwards a neutral object (the 'probe') appears either at the location of the valent or at the location of the neutral stimulus. Participants are instructed to categorize the probe (according to an unrelated dimension) as fast as possible. Response times are recorded. Usually, participants respond faster to probes that appeared at the location they were looking at at probe onset (Posner, Snyder, & Davidson, 1980; Bradley, Mogg, & Millar, 2000). Thus, reaction times that are faster following probes that appeared at the location of valent stimuli compared to reaction times following probes at the location of neutral stimuli suggest an attentional bias towards these valent stimuli.

A modified version of the dot-probe task similar to Schwager and Rothermund (2013) was used in combination with a risky lottery paradigm. Each lottery presented comprised two options with two payoffs each (along with probabilities) that together served as four spatially separated stimuli. Immediately afterwards the probe appeared at one of these four locations. Participants were instructed to respond as fast as possible to the probe. Afterwards they had to make their lottery choice.

Typically, an exposure duration of 500 ms is used in dot-probe tasks (Staugaard, 2009).

[21] Compared to the expected value of the option.

[22] Both valent and salient stimuli capture our attention. Whereas valence affects our attention through the intensity of an emotion which may be positive or negative, salience affects our attention without distinguishing between positive or negative.

However, exposure durations are critical since attention allocation is realized in a serial fashion. After initial attention allocation to the most salient stimuli, attention is shifted sequentially to less salient ones (Itti and Koch, 2000). Therefore, the stimulus onset asynchrony (SOA)[23] is a key variable in dot-probe studies. Using different SOAs different stages of processing can be assessed. Whereas SOAs shorter than 500 ms reveal a rapid orienting towards salient stimuli, longer SOAs reflect later strategic and controlled attentional processes (Grimshaw, Foster, & Corballis, 2014; Cisler & Koster, 2010). This is demonstrated by results of dot-probe studies on social anxiety (e.g., Stevens, Rist, & Gerlach, 2009; Cooper & Langton, 2006; Mogg, Philippot, & Bradley, 2004). They found that vigilance for threatening social cues was prevalent at exposure durations smaller (or equal) than 500 ms, whereas attentional bias seemed to change at durations around 500 ms and longer. Similar changes in attentional bias between different lottery exposure durations were expected in this study, especially since participants had to try to embrace all information as fast as possible. To examine the time-course of attentional bias exposure durations of 150 ms as well as 4000 ms were used.[24]

8.2. Method

8.2.1. Participants

34 healthy subjects participated in the experiment.[25] Number of subjects was limited to approx. 30 due to time and money constraints since EEG was recorded.[26] Data collection was terminated when data of at least 25 subjects of which half are male, half female, was collected and usable. Recruitment was realized via ORSEE (Greiner, 2015). Payment was 6 € (approx. 6.74 USD) per hour (or course credits) plus a possible gain ranging from 2 € to 12 € (approx. 2.25 USD to 13.48 USD). The average additional gain was 9.96 € (approx. 11.19 USD). Eligible subjects were right-handed and did not miss any of their previous experiments they subscribed to. The study followed the Helsinki Declaration of 1975 and was approved by the ethics committee of the Faculty of Social and Behavioral Sciences of the Friedrich Schiller University of Jena.

The first three subjects were excluded from analyses since there were technical problems during the experiments. A further subject was excluded since she obviously had not understood the task correctly and pushed the buttons in more than half of the trials already when the lotteries were shown so that no probe could appear afterwards. The remaining sample comprised 30 subjects (56.7 % female; 20-37 years old, mean age = 25.0 years, sd = 4.12).

[23] The stimulus onset asynchrony describes the time between onset of the cue and onset of the probe. If the probe onset immeadiately follows the offset of the cue, the SOA is equal to the exposure duration of the cue.

[24] In pre-tests exposure durations of 150 ms, 500 ms, 2000 ms and 4000 ms were used. Results suggested the strongest salience effect after 150 ms exposure and an opposing effect after 4000 ms.

[25] The experiment was conducted in July 2016 in Jena.

[26] Such a small sample size is standard practice in neurosciences due to the much higher time and money expenses in EEG (and similar) experiments compared to behavioral experiments.

8.2.2. Experimental Design

A 2 (salience condition: probe at salient vs. non-salient position) x 2 (lottery exposure duration: 150 ms, 4000 ms) factorial within-subject repeated measures design was used. Trials started with a centered, white fixation cross on a black background (1000 ms - 1500 ms). Lotteries were presented in the form of two options with two payoffs each. Payoffs on the left-hand side belonged to one option, payoffs on the right-hand side belonged to the other option. The exposure durations of the lotteries varied in a pre-specified random order between 150 ms and 4000 ms. Immediately after disappearing and before participants could make their lottery choice the modified dot-probe task was presented. Either a square or a circle appeared for 150 ms at one of the four locations where the payoffs just had been presented. In 50 percent of the trials the probe appeared at locations of salient payoffs, randomly distributed over the four possible locations. While the screen was blank participants had to categorize the probe as circle or square. No performance feedback was given. Afterwards they could immediately give their lottery choice. Also, no lottery feedback was given in order not to influence the attentional focus in the next lottery. The next trial commenced after 1000 ms of blank screen. All choices were given by pressing buttons (placed side by side) with the right hand. Each participant completed 240 trials divided into two blocks. At the end of the experiment the lottery payoff from ten randomly chosen trials was shown. The design is presented in figure 8.1.

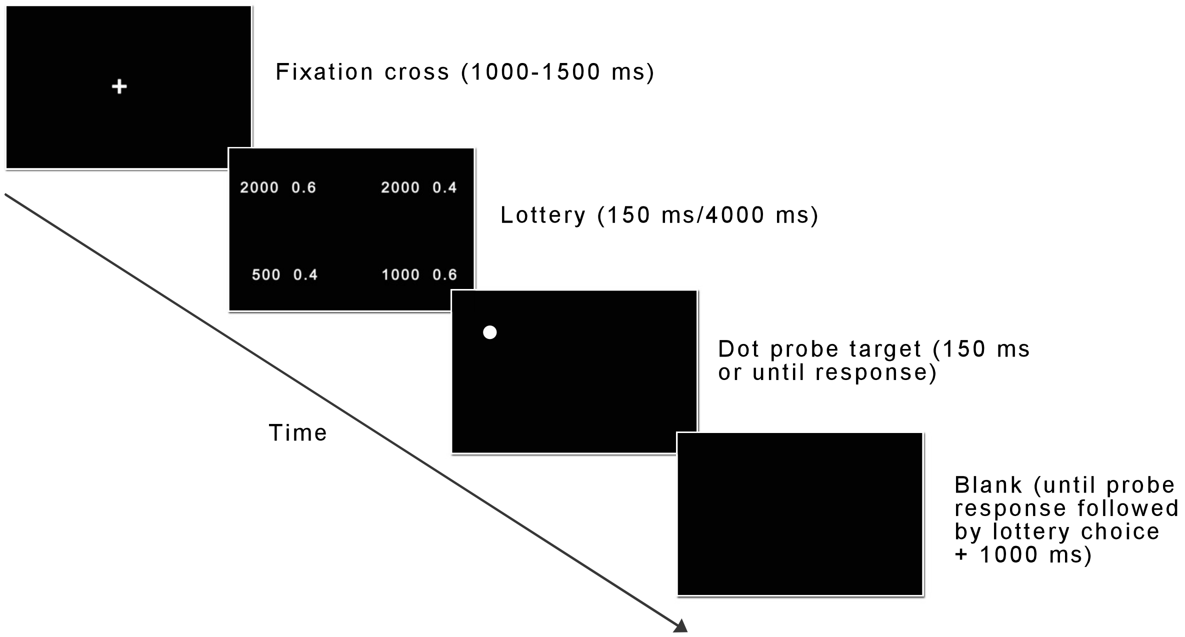

Figure 8.1.: Schematic display of the paradigm.

Lottery: Left payoffs belong to option 1, right payoffs belong to option 2.

8.2.3. Procedure and Material

Prior to participating in the experiment subjects completed an online questionnaire measuring, among others, the achievement motivation and proneness to optimism or pessimism.[27] At the experiment location participants were instructed that lotteries would be presented for a very short time and they should always make a choice even if they could not read all given option information. They were also told to give their answer to the dot-probe task as fast as possible even though this would not influence their gain. After instructions and the possibility to ask questions participants gave written informed consent. They then received a flexible cap with electrodes that recorded EEG during the experiment. Therefore, participants were seated on a comfortable chair, approximately 120 cm away from the screen, in a dimly lit and electrically shielded room. The dot-probe stimuli occupied horizontal and vertical visual angles of approximately 0.3°x 0.3°. EEG results are presented in chapter 9.3. To get used to the fast presentation of the lotteries participants played four training trials comprising two lotteries that appeared for 4000 ms and two lotteries that appeared for 150 ms. Afterwards they had again the possibility to ask questions. In between the blocks participants could make a short break and were asked whether everything was okay.

Lotteries were adapted from Pachur, Hertwig, Gigerenzer, and Brandstätter (2013) (experiment 1) and Fiedler and Glöckner (2012) (study 1). All options had similar expected values (EV). 30 different lotteries were used, 10 of them in the negative domain. Four versions of each of the 30 lotteries were used to test all four possible salience combinations within similar lotteries with the same salience conditions. The payoffs of these four versions were multiples[28] of the original lotteries in order to keep all proportions constant, probabilities were the same (see table 8.1 and figure 8.2 for an example). The resulting 120 lotteries were used once in 150 ms and once in 4000 ms trials. Lotteries were randomized.[29] The experiment was performed using Presentation software (Version 17.1, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com).

Table 8.1.: Multiples of a lottery.

	Factor	Option 1	Option 2	EV_1	EV_2
A	1	400 (0.2) / 200 (0.8)	300 (0.7) / 100 (0.3)	240	240
B	0.5	200 (0.2) / 100 (0.8)	150 (0.7) / 50 (0.3)	120	120
C	3	1200 (0.2) / 600 (0.8)	900 (0.7) / 300 (0.3)	720	720
D	5	2000 (0.2) /1000 (0.8)	1500 (0.7) / 500 (0.3)	1200	1200

Payoffs in Eurocent, probabilities in brackets.

[27] These questionnaires were used in order to take into account top-down processes of attention. Unfortunately, data could not be used for the intended purpose since the number of subjects was to small in order to investigate the influence of motivation variables on salience on subject level. Therefore, this remains a topic for future research.

[28] The same factors were used for all lotteries.

[29] Randomization was realized under the condition that at least 10 trials were between the two same lotteries represented for 150 ms and 4000 ms, respectively.

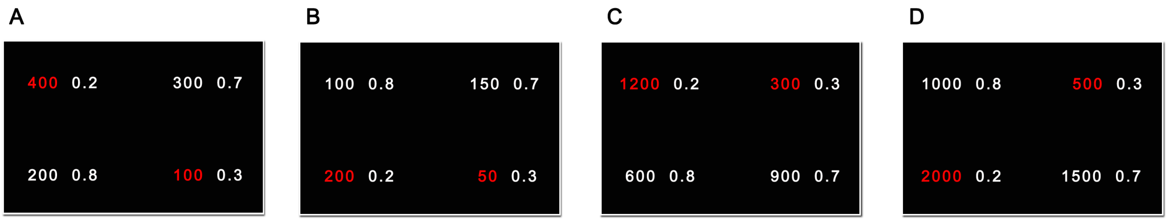

Figure 8.2.: Example of multiples of a lottery and distribution of most salient payoffs. Here, most salient payoffs are highlighted in red.

8.3. Statistical Analysis and Results

The dot-probe reaction time was measured as the time it took subjects to detect the probe and respond by pressing a button following the disappearance of the probe.[30] Trials with latencies smaller than 200 ms and larger than three interquartile ranges above the third quartile of the response time distribution (3262 ms) ("far out values", Tukey (1977)) were excluded. Two subjects met this criterion in more than 10 % of the trials and were therefore excluded from further analysis. These trials represented 1.4 % of the remaining data. Furthermore, incorrect trials (17.2 % of remaining data) of the dot-probe task where excluded since these cannot give any hint on previous attention allocation.

Since pre-tests indicated differing salience effects for different lottery exposure times data for lotteries presented for 150 ms and those presented for 4000 ms were examined separately. To explore reaction time data a multilevel generalized linear mixed-effects model was used. Thus, the effect of salience[31] could be explored allowing for variation across subjects in a varying slope model.[32] Furthermore, these models allow to account for the nested group effects of subjects and blocks as well as for longitudinal effects (such as boredom) (Baayen, Davidson, and Bates, 2008).[33] The trial was included as random effect.[34] [35]

The following explanatory variables were included as fixed effects in a stepwise variable selection procedure to correct for experimental influences on salience and reaction time, respectively: ProbeLeft and ProbeTop are one if the probe appeared on the left-hand and on the upper side, respectively, and zero otherwise; reactionTime$_{-1}$ is the reaction time to the probe in the preceding trial[36] [37]; Loss accounts for the lotteries in the loss domain; Male is one for male and zero for female subjects; OptionRight describes the lottery choice given right after the dot-probe task and is one if the right option was chosen, zero otherwise.[38]

[30] Data was analyzed with RStudio Version 1.0.136 - 2009-2016 RStudio, Inc.
[31] Salient is one if the probe appeared at a position of salient payoffs as defined by Salience Theory, zero otherwise.
[32] This leaves room for (bottom-up) salience differently influencing subjects.
[33] For further advantages of mixed-effects models see, Baayen et al. (2008).
[34] Only trials of 150 ms or 4000 ms lottery exposure were counted in the respective analysis.
[35] Trial was centered on 1 in order to interpret the intercept more easily.
[36] Baayen et al. (2008) highlight the reaction time in the preceding trial as important predictors.
[37] ReactionTime$_{-1}$ was centered on the mean in order to interpret the intercept more easily.
[38] The position of the probe at the better payoff of a option was also tested but the effects were insignificant and the model quality was reduced.

Is salience significantly influencing reaction times? Estimation results of reaction times after lottery exposure of 150 ms show a highly significant effect of salience. Reaction times following probes at salient payoff positions are significantly smaller. This effect is strongly enhanced in the loss domain. According to the Attentional Model (Yechiam and Hochman, 2013) losses lead to a higher physiological arousal and a heightend on-task attention, thus increasing the sensitivity to task incentives. Therefore, the interaction effect of salience and losses can be interpreted as emphasis of the general salience effect. Results can be seen in figures 8.3, 8.4 and table 8.2. After 4000 ms lottery exposure there is no significant effect of salience. This is in line with theory of attention allocation (see, e.g., Itti & Koch, 2000). Especially in our design where subjects knew that lotteries were presented for a very short time they probably tried to embrace all information as fast as possible, moving their attention away from salient cues as fast as possible. Our results thus show an early attentional orienting towards the salient payoffs (which is revealed at stimulus onset asynchronies (SOAs) shorter than 500 ms (Grimshaw et al., 2014; Cisler & Koster, 2010). Within later strategic and controlled attentional processes (revealed at longer SOAs (Grimshaw et al., 2014; Cisler & Koster, 2010)) no salience effect seems to exist.

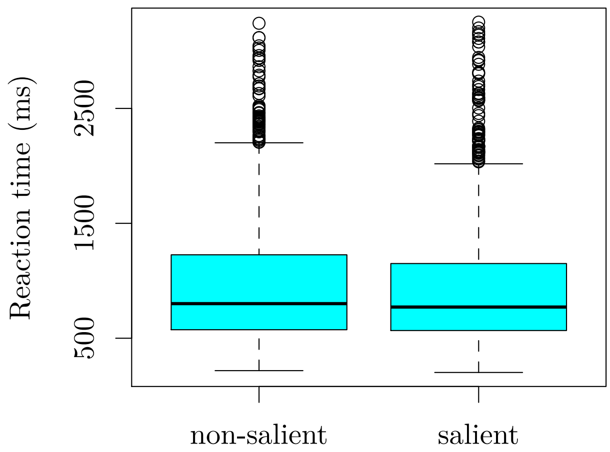

Figure 8.3.: Reaction times after probes at salient vs. non-salient payoff positions (following a lottery exposure of 150 ms).

Non-salient: Mean=967 ms, SD=541 ms. Salient: Mean=936 ms, SD=528 ms.

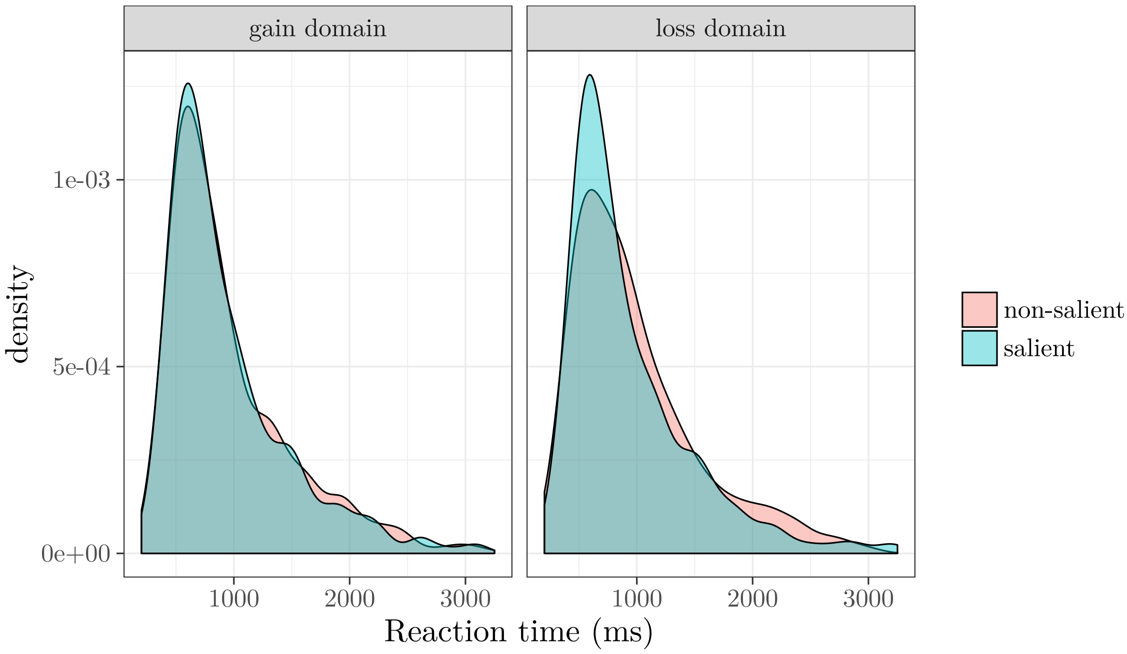

Figure 8.4.: Reaction times following probes at salient vs. non-salient payoff positions (following a lottery exposure of 150 ms) in the gain and the loss domain.

Gain domain. Non-salient: Mean=957 ms, SD=536 ms. Salient: Mean=939 ms, SD=526 ms.
Loss Domain. Non-salient: Mean=994 ms, SD=554 ms. Salient: Mean=929 ms, SD=535 ms.

Table 8.2.: Multilevel generalized linear mixed-effects model estimation of reaction times following probes after lottery exposure of 150 ms.

	Reaction time in ms			
	(1)	(2)	(3)	(4)
Salient	−38.973**	−25.751***	−26.173***	−24.226***
	(18.222)	(7.149)	(6.324)	(7.139)
$\text{Trial}_C 1^{\text{a}}$	−3.884***	−3.425***	−3.418***	−3.393***
	(0.388)	(0.453)	(0.459)	(0.467)
ProbeLeft	−21.982	−17.792**	−17.718***	−39.055***
	(16.428)	(7.021)	(5.597)	(6.488)
ProbeTop	23.396	30.229***	30.237***	31.102***
	(17.603)	(7.587)	(6.376)	(10.232)
$\text{ReationTime}_{\text{-1}}\text{\_MC}^{\text{c}}$		0.131***	0.132***	0.132***
		(0.015)	(0.015)	(0.015)
Loss		30.957***	31.118***	33.615***
		(6.373)	(6.267)	(7.621)
Male			−141.529***	−148.942***
			(6.851)	(9.238)
OptionRight				−29.903***
				(7.624)
ProbeLeft*ProbeTop	9.201	7.014	6.945	6.061
	(23.607)	(8.230)	(7.005)	(9.706)
Loss*Salient		−47.938***	−48.013***	−51.450***
		(10.443)	(7.217)	(11.045)
ProbeLeft*OptionRight				35.155***
				(7.666)
Constant	1,259.040***	1,213.691***	1,267.549***	1,285.530***
	(39.146)	(10.654)	(7.961)	(10.162)
N	2,783	2,783	2,783	2,783
Akaike Inf. Crit.	39,782.390	39,676.640	39,677.040	39,677.530
Bayesian Inf. Crit.	39,877.300	39,789.330	39,795.670	39,808.020

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.
[a]Only trials of 150 ms exposure included. Trial centered on 1.
[b]Reaction time in previous trial. Centered on mean.

8.4. Discussion

Using a dot-probe task in combination with a risky lottery paradigm, it could be shown that attentional biases towards single payoffs within lotteries seem to exist. The results suggest that these salient payoffs can be identified by the salience definition given by Bordalo and colleagues' Salience Theory.

Salience effects after 150 ms exposure might even be stronger than the results reveal. As Salience Theory suggests there are always two salient payoffs (one of each option) within a risky choice. With the dot-probe task only the attentional bias towards one payoff at once could be measured. Therefore, the other salient payoff probably served as distractor drawing the initial attention in a similar fashion. Eyetracking or similar might be better methods to assess the effect size of salience within risky lotteries as applied here. As there also exists criticism on the reliability of the dot-probe task (see, e.g., Schmukle, 2005; Staugaard, 2009) future research should validate the results using alternative methods.

With their salience formula Bordalo and colleagues only account for bottom-up salience processes, i.e., salience determined by the sensory information of the stimulus. Attentional bias is also modulated by top-down processes (like goals, drives, the current psychological state). These would have to be incorporated in the salience determination to better understand and predict risky choices.

It was not the aim to control for predictions of risky choice made by Salience Theory. We assumed that our paradigm would not be suitable, due to the dot-probe task in between lottery presentation and decision, as well as the time pressure induced by the short lottery exposure times. As suggested by several studies (e.g., Young et al., 2012; Kocher et al., 2013; Ben Zur & Breznitz, 1981) time pressure modulates risky choices. To investigate the influence of salience on risky choices under consideration of time as important moderator would be an interesting topic for future research.

9. A Neural Investigation of Attentional Processes Involved in Salience Biases within Risky Lotteries

This chapter is co-authored by Wolfgang H.R. Miltner and Marcel Franz.

9.1. Introduction

The last chapter presented the experiment investigating the salience differences in lottery payoffs proposed by Salience Theory. Therefore, risky lotteries were used in combination with the dot-probe task (see chapter 8 for a detailed description of the paradigm). This task was developed by MacLeod et al. (1986) and has become a widely used psychological measure to assess selective-attentional processing (Frewen et al., 2008). Attention allocation is realized in a serial fashion. After initial attention allocation to the most salient stimuli, attention is shifted sequentially to less salient ones (Itti & Koch, 2000). In order to examine the time-course of attentional bias we presented lotteries for 150 ms as well as 4000 ms[39]. Often stimulus exposure durations of 500 ms are used in dot-probe tasks (Staugaard, 2009). Results of dot-probe studies on social anxiety (e.g., Stevens et al., 2009; Cooper & Langton, 2006; Mogg et al., 2004) suggest that an exposure duration of 500 ms seems to be critical. They found that vigilance for threatening social cues was prevalent at exposure durations smaller (or equal) than 500 ms, whereas attentional bias seemed to change at durations around 500 ms and longer.

Behavioral data showed a significant effect of salience on reaction time after a lottery exposure of 150 ms indicating that early attention is biased towards salient payoffs as defined by Salience Theory. After a lottery exposure of 4000 ms no significant results were found. This is in line with our hypothesis that our paradigm would make subjects rapidly shift attention in order to embrace the lottery information as fast as possible. Detailed behavioral results are presented in chapter 8.3.

Whereas behavioral data can give hints on early attentional orienting versus later strategic attentional processes via the manipulation of the stimulus onset asynchrony (SOA) (Grimshaw et al., 2014) it still reflects the overall result of many distinct neural processes. Event-related potentials (ERPs) provide a continuous measure of neural processes. Thus, the time course of attentional bias, which is still barely understood (Torrence & Troup, 2017), can be examined over the course of a trial with milisecond resolution (Kappenman, Farrens, Luck, & Proudfit,

[39] In pre-tests we used exposure durations of 150 ms, 500 ms, 2000 ms and 4000 ms. Results suggested the strongest salience effect after 150 ms exposure and an opposing effect after 4000 ms.

2014). In order to gain a deeper understanding of processes involved in the attentional bias within risky lotteries caused by salience differences as proposed by Salience Theory, we recorded EEG in addition to behavioral data. ERPs are based on EEG recordings. They are time-locked and averaged across events of interest. There are typical waveforms within the ERPs. These waveforms may vary in amplitude and latency and also with scalp location. In the dot-probe task two events may be investigated regarding their ERP components: the cue (in our experiment the lottery) and the probe.

Looking at early ERP components time-locked to the cue even allows to investigate attentional bias that already shifts before probe onset. According to Müller and Rabbitt (1989) covert attention can shift between locations in as little as 100 ms. Unfortunately, investigating those ERPs was not possible in our study. Since the salience in lotteries according to Salience Theory is always the same for one payoff in both options, salient payoffs were always present on both the left-hand and the right-hand side. Therefore, ERPs that are typically larger contralateral to salient stimuli compared to their ipsilateral amplitude were not suitable here.

To examine EEG data on the attentional bias caused by salient payoffs as suggested by Salience Theory we focused on ERPs time-locked to the probe. Neural correlates of attentional bias can be found both in more frontal and more posterior scalp locations. Research suggests that the former represent engagement/disengagement processes while the latter rather represent the effects of attention on sensory processing (Torrence & Troup, 2017). Two ERP components, the P1 and the N1, are early sensory responses that are not influenced by top-down factors (Luck & Kappenman, 2012b).[40] Since payoff salience proposed by Bordalo et al. (2012a) is merely based on sensory information[41] these components are well suited to validate the suggested attentional bias. Both components are sensitive to spatial attention. That is, amplitudes are larger when subjects already attended the location where the stimulus appears (see review by Hillyard, Vogel, & Luck, 1998; Mangun, Hopfinger, Kussmaul, Fletcher, & Heinze, 1997). Therefore, both components may reveal an early orienting to salient stimuli. The P1 component has a positive peak around 100-130 ms poststimulus onset. It likely arises in extrastriate visual cortex and is typically strongest in the lateral occipital lobe, contralateral to the stimulus (Luck & Kappenman, 2012b). The N1 is the first negative component following the P1. There are several visual subcomponents of the N1. The first can be seen in anterior electrodes and peaks around 100-150 ms. Two further subcomponents can be seen in posterior electrode sites, peaking around 150-200 ms poststimulus. One arises from parietal cortex and the other from lateral occipital cortex (Luck, 2005). We expected probes appearing at positions of salient outcomes to elicit a larger P1 and a larger N1 both contralateral to the probe compared to probes at non-salient positions, reflecting an early orienting towards salient stimuli.

We also wanted to test for later attentional processes indicating a salience bias within risky lotteries. Therefore, we looked at the P3 component which is mediated by top-down control.

[40] Luck and Kappenman (2012b) point out though, that several generators seem to contribute to the overall P1 and N1 voltage deflections and it is not clear whether all P1 and N1 subcomponents are purely exogenous. Therefore, P1 and N1 should be interpreted as reflecting a simple top-down modulation of the initial feedforward sensory activity.

[41] Bordalo and colleagues propose that their salience formula may be expanded by top-down factors.

It appears around 250 ms and thereafter and is associated with the allocation of attentional resources. There are two subcomponents: the P3a has a maximum amplitude over frontal areas and the P3b over central/parietal areas. The P3a is more directly related to the orienting response and appears when focal attention for standard stimuli is disrupted by the detection of a probe (Polich, 2007). It is probably generated by the anterior cingulate cortex (ACC) and related structures (Zhang, Liu, Wang, Ai, & Luo, 2017; Polich, 2007). The ACC has reciprocal connections to the amygdala, which supports quick orienting towards salient stimuli (Williams et al., 2006). The P3b was localized to the parietal lobe, which might be involved in visual-motor coordination (Zhang et al., 2017). Whereas the P3a appears to be independend of early attentional orienting, the P3b is probably enhanced only when early attention was directed to another location. It may thus reflect the cognitive resources needed to redirect attention to task-relevant locations through voluntary control (Zhang et al., 2017; Liu, Zhang, & Luo, 2015). We expected probes appearing at positions of salient outcomes to elicit an enhanced P3a. Since it was task-relevant to embrace all information as fast as possible to give a lottery decision, after initial attention allocation to the most salient payoffs subjects had to redirect their focus to non-salient payoffs. Thus, we expected probes appearing at positions of non-salient outcomes to elicit an enhanced P3b compared to probes at salient positions, reflecting the cognitive resources needed to redirect attention to task-relevant locations.

An illustration of all investigated ERP-based mechanisms for attentional modulation can be seen in figure 9.1.

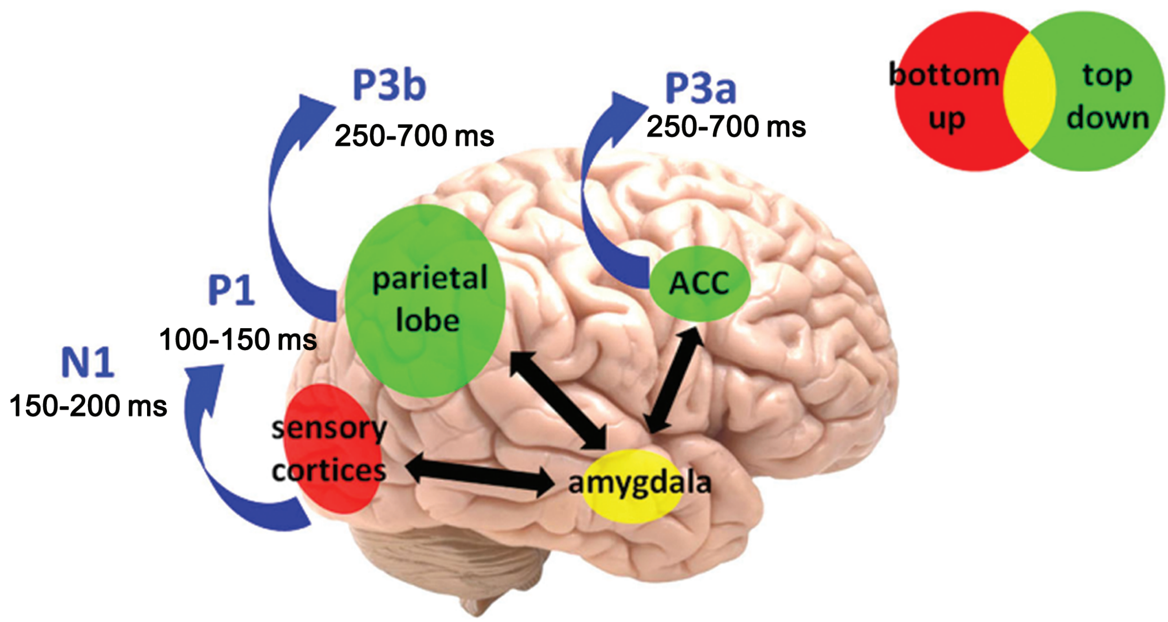

Figure 9.1.: Illustration of ERP-based mechanisms for attentional modulation.

Figure adapted from Zhang, Liu, Wang, Ai, and Luo (2017).

9.2. Method

For information on participants, experimental design and procedure see chapter 8.2.

In addition to participants that were excluded from behavioral analysis as described in chapter 8.2, three further subjects had to be excluded from EEG analysis due to problems during EEG recording.

9.2.1. EEG Recording

Continuous EEG was recorded from 64-sintered Ag/AgCl electrodes (EASYCAP GmBH, Herrsching-Breitbrunn, Germany) placed on the scalp with 21 electrodes localized according to the 10-20 system on participants' heads. The remaining electrodes were interspaced equally between these 21 sites. To record the vertical electrooculogram (EOG) a further electrode was placed underneath the lower eyelid of the left eye. All channels were online referenced to FCz. The impedance was kept below 10 kOhm. EEG and EOG signals were amplified using two BrainAmp amplifiers and recorded with the BrainVision Recorder software (both Brain Products, Gilching, Germany). Following analogue band-pass filtering (0.05–500 Hz), continuous EEG signals were digitized with a sampling rate of 1000 Hz and stored to hard drive for offline analysis.

9.2.2. Preprocessing

Data preprocessing was realized with EEGLAB software (Delorme & Makeig, 2004, Version 13.6.5b). Data was down-sampled to 250 Hz and re-referenced to linked mastoids, i.e., the average over channels TP9 and TP10, for further processing. Independent component analysis was used in order to remove artifacts related to eye-blinks or ocular movements. Therefore, a duplicate of the re-referenced data was offline filtered from 1- 40 Hz with a transition bandwidth of 1 Hz (highpass) and 10 Hz (lowpass), respectively, using a Hamming windowed sinc finite impulse response (FIR) bandpass filter. Data was then segmented into continuous intervals of 1 s. Unique, nonstereotyped artifacts indicated by unlikely EEG values (>3 SD)were removed by applying higher order statistic functions (pop_jointprob, pop_rejkurt) to electrode channels (Delorme, Sejnowski, & Makeig, 2007). Extended infomax independent component analysis, as implemented in EEGLAB, was then applied to the data. Subsequently, the resulting demixing matrix was applied to the original re-referenced EEG data. Components representing eye-blinks or ocular movements were identified using the fully automated method EyeCatch (Bigdely-Shamlo, Kreutz-Delgado, Kothe, & Makeig, 2013) and subtracted from data (Jung et al., 2000). Using a Hamming windowed sinc FIR bandpass filter (pop_eegfiltnew) from 0.1- 40 Hz with a transition bandwidth of 0.1 Hz (highpass) and 10 Hz (lowpass), respectively, data was offline filtered. EEG epochs were selected for each trial beginning from 200 ms prior to the onset of the stimulus and continuing for 1200 ms. Afterwards, baseline correction was performed using the pre-stimulus interval from -200 ms to 0 ms for each electrode. Finally, non-stereotyped artifacts (pop_jointprob, pop_rejkurt) were removed by discarding epochs with amplitude values >3 SD.

After artifact rejection (as well as exclusion of far out values and incorrect trials as described in chapter 8.2), 67.8% (Min: 41.7%, Max: 80.9%) of all trials were retained on average. The mean, minimum and maximum number of trials per condition are provided in table 9.1.

Table 9.1.: Number of trials per condition after artifact rejection, exlusion of far out values and incorrect trials.

	150 ms				4000 ms			
	Gain Domain		Loss Domain		Gain Domain		Loss Domain	
	Non-Salient	Salient	Non-Salient	Salient	Non-Salient	Salient	Non-Salient	Salient
Original	40.0	40.0	20.0	20.0	40.0	40.0	20.0	20.0
Mean	30.4	30.1	10.8	10.1	29.8	30.3	10.6	10.8
Min.	21.0	20.0	7.0	5.0	18.0	16.0	7.0	3.0
Max.	38.0	39.0	15.0	15.0	39.0	40.0	14.0	14.0

In order to obtain event-related brain potentials (ERPs) for each participant in response to stimuli, epochs were averaged for each participant and stimulus type (salient, non-salient). Preprocessed EEG datasets were then imported into the Statistical Parametric Mapping (SPM12, version: v7219; http://www.fil.ion.ucl.ac.uk/spm) software package for MEG/EEG data (Litvak et al., 2011), and used for subsequent EEG source reconstruction.

9.3. Statistical Analysis and Results

Statistical analysis was realized with Matlab (Version R2018a) using the toolbox Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). Cluster-based permutation tests (Maris & Oostenveld, 2007) were used to control for alpha-error accumulation. Therefore, t-values were calculated for all pairs of data. All t-values surpassing a critical threshold corresponding to an alpha level of 0.01 were clustered over adjacent time bins. T-values of the clusters were summed up in order to calculate the cluster-level statistics. These were then tested against the distribution of maximum cluster-level statistics obtained from Monte Carlo simulations with 1,000 permutations. A maximum cluster-level distribution was constructed under the null hypothesis of no experimental effect. If the original cluster was greater than the 95th percentile of the maximum cluster-level distribution it was considered to be significant at an alpha level of 0.05.

Analyses were performed seperately for 150 ms and 4000 ms. In addition to analysis of overall lottery data, further analyses for the gain and the loss domain were performed. The early sensory ERP components P1 and N1 were assessed first at the electrode POz and Oz and to control for laterality effects also at PO3 and PO4. For all datasets, ERP waves showed no significant salience effects for P1 and N1 time windows. Interestingly, in 150 ms data only the P1 appeared whereas no N1 could be identified. Furthermore, the P1 peaked very early at around 60 ms after dot onset.[42] In contrast, in 4000 ms data only the N1 appeared whereas no P1 could be identified. The peak around 150 ms corresponds to other research findings. To examine salience effects on P3a and P3b, ERP waves were assessed at electrodes Fz and Pz, respectively. Here,

[42]In their review of dot-probe tasks investigating attentional biases towards emotional faces Torrence and Troup (2017) report a P1 onset approximatly around 70-75 ms after dot onset.

also no significant salience effects were found for all datasets. Furthermore, for both 150 ms and 4000 ms data no P3a was found. At this late time window activity was apparently concentrated in parietal regions where a P3b was elicited.

ERP waves and topography maps are provided in figures 9.2 and 9.3.

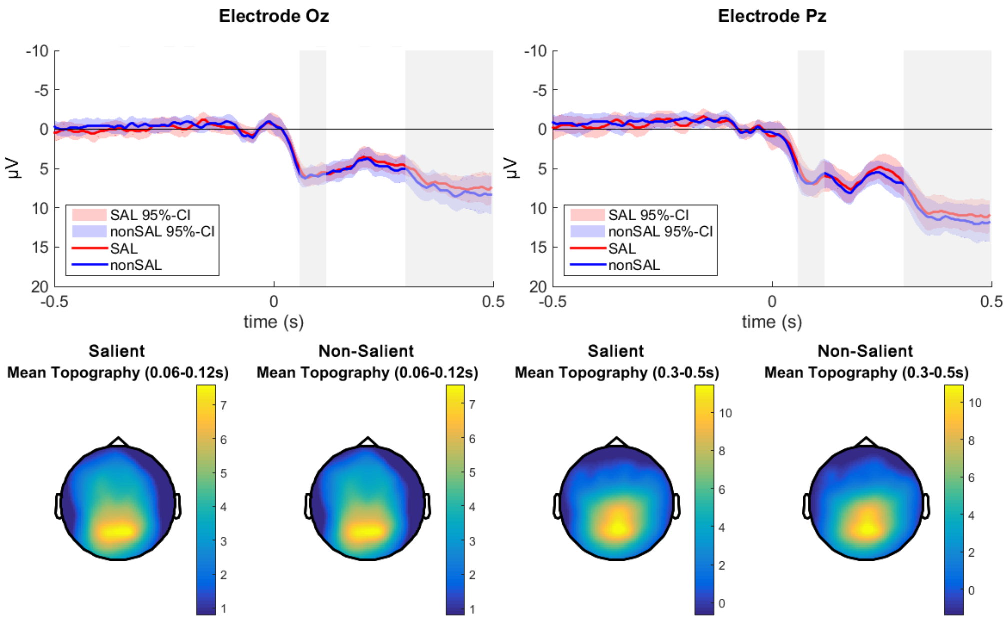

Figure 9.2.: Grand average ERPs and mean topography showing a P1 at electrode Oz and a P3b at electrode Pz after 150 ms exposure time. Grey areas indicate time windows for P1 and P3b.

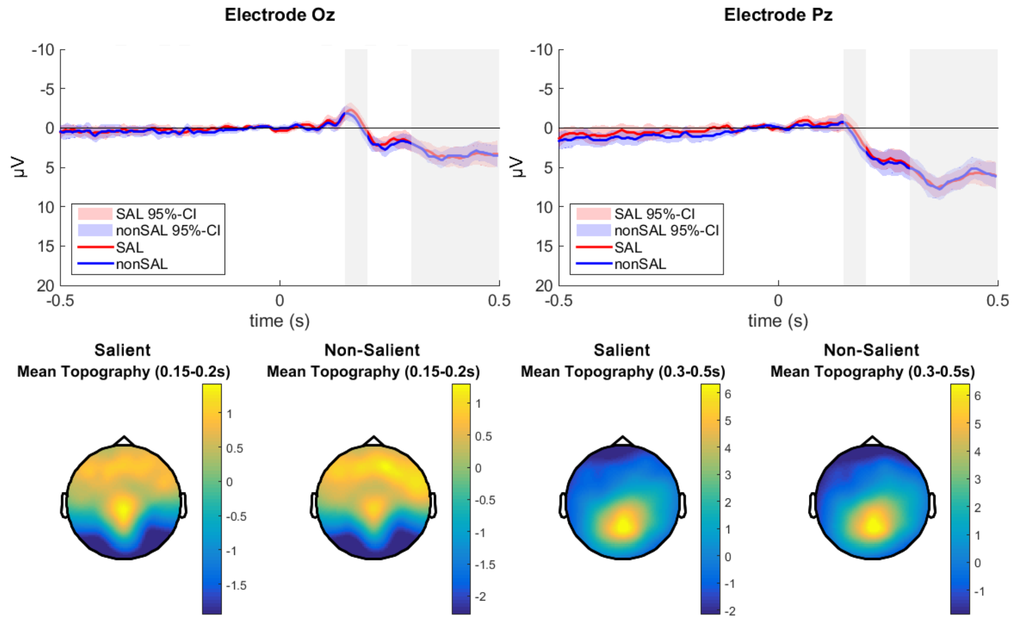

Figure 9.3.: Grand average ERPs and mean topography showing an N1 at electrode Oz and a P3b at electrode Pz after 4000 ms exposure time. Grey areas indicate time windows for N1 and P3b.

9.4. Discussion

In their Salience Theory, Bordalo et al. (2012a) propose salience differences in lottery payoffs leading to biases in decisions under risk. To calculate the salience of payoffs they suggest a formula (see chapter 8.1) that captures the properties of ordering, dimishing sensitivity and reflection. The aim of this study was to validate the salience of payoffs proposed by Bordalo et al. (2012a) and furthermore to identify attentional processes that are involved in the salience effect. Therefore, behavioral data and EEG data were used. Behavioral data indicated a significant effect within lotteries exposed for 150 ms, suggesting that early attention is biased towards salient payoffs. After a lottery exposure of 4000 ms, no significant salience effect could be found. This is in line with our hypothesis that salience would only affect early attention whereas later strategic processes would direct attention to other task-relevant information.

However, EEG data could not confirm any salience effects proposed by Salience Theory. On the one hand, this may put in question the validity of behavioral results. On the other, further reasons might be responsible for EEG results that we will explore in the following.

Before, we want to present possible explanations for our differential findings of P1 and N1 components in 150 ms and 4000 ms data. These are supported by other research (e.g., Luck, Fan, & Hillyard, 1993; Vogel & Luck, 2000) that also found P1 attention effects without N1 attention effects and vice versa, indicating a clear differentiation between both attention effects. Luck and colleagues (see, e.g., Vogel & Luck, 2000; Hillyard et al., 1998) proposed that the P1 component may reflect a suppression of sensory processing at unattended locations whereas the

N1 component may reflect a limited-capacity discrimination process at the attended location. When items were presented in a rapid succession at the attented location, the N1 effect but not the P1 effect was eliminated (Luck, Heinze, Mangun, & Hillyard, 1990; Heinze, Luck, Mangun, & Hillyard, 1990), confirming the hypothesis of the mechanisms behind the N1 component being highly limited in capacity and not being able to operate repeatedly in a short period of time (Luck & Kappenman, 2012a). This may explain our finding of an N1 component after an exposure time of 4000 ms which did not appear after an exposure time of 150 ms. In the latter, cognitive capacity may have been overloaded by the complex lottery stimuli which were rapidly followed by the probe. In another study, Luck et al. (1993) also found a P1 attention effect only after a short SOA of 250 ms whereas the effect was eliminated after a longer SOA of 400 ms, which corresponds to our findings. They propose that the early attention effect indicated by the P1 component might already have been completed before the onset of the probe after the long SOA.

Looking at our paradigm (see chapter 8.2.2) it is very complex compared to other EEG studies investigating attentional biases based on the dot-probe task. Often, stimuli like pictures of faces or words are used that may be processed more easily. In our paradigm, four different numbers served as stimuli, and in addition the respective probabilities were given. This paradigm was chosen in order to validate the salience within lotteries that are very similar to those described by Bordalo et al. (2012a). It may have been too complex in order to validate salience effects based on EEG data. Also, the fact that, to our knowledge, no comparable EEG study exists that investigates the attentional bias based on a dot-probe task following risky lotteries makes it difficult to interprete our results.

Salience effects in behavioral data were very small. Here, the complexity of the paradigm might also be a possible reason. Also, as already pointed out in the discussion of behavioral results (chapter 8.4), payoffs that are equally salient always appear in both options (since the salience is based on comparisons across options, see chapter 8.1), that is, on both the left-hand and the right-hand side. In the dot-probe task only the attentional bias towards a single stimulus can be assessed. Thus, the chance to capture an attentional bias towards salient payoffs was reduced by 50%.

Future research should address these problems by reducing the complexity of the lotteries presented. Furthermore, another paradigm than the dot-probe task, which was designed to measure reaction times, that may be more suitalbe to assess differences within attentional processes using EEG data (cf. Torrence & Troup, 2017) and that may also allow to account for salient stimuli on both sides, should be used.

Part III.

Salience and Nudges

10. Can Nudges Be Transparent and Yet Effective?

The following chapter is based on the publication

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

10.1. Introduction

Nudges are evolving into a popular form of soft regulation in various fields such as health, finance, and environmental protection (Alemanno & Sibony, 2015; Lourenço, Ciriolo, Rafael Almeida, & Troussard, 2016; Sunstein, 2014; World Bank, 2015). Despite its growing popularity, the use of behavioral insights in policy-making is subject to criticism (e.g., Hausman & Welch, 2010; Rebonato, 2014). One remarkable and often criticized aspect of nudges is that they often influence individual behavior without being noticed by the affected subject (Dhingra, Gorn, Kener, & Dana, 2012; Hansen & Jespersen, 2013; Sunstein, 2016). This raises the concern that nudges covertly violate individual autonomy and are therefore unethical (Bovens, 2009; House of Lords Report, 2011). Such regulation thus lacks the transparency that characterizes other regulatory instruments. For instance, when the government imposes a tax to reduce consumption of a product (e.g., cigarettes, or carbon dioxide), people are aware of this tax and can compel the government to justify it (Sunstein, 2014). On the other hand, when the government sets an opt-out system instead of an opt-in system to promote certain behavior (e.g., organ donation) it exploits several psychological biases, often without people's awareness (Hansen and Jespersen, 2013). Felsen, Castelo, and Reiner (2013) demonstrated in a vignette study that a significant proportion of individuals have reservations towards nudges they perceive as covert. Additionally, another recent research stream provides evidence of the intrinsic value of decision rights and autonomy (Bartling, Fehr, & Herz, 2014; Fehr, Herz, & Wilkening, 2013; Owens, Grossman, & Fackler, 2014). To address this criticism we investigated whether nudges can be made transparent without reducing their effectiveness. In this context, we take into account that the covert nature of nudges is often said to be essential for their effectiveness (Bovens, 2009; House of Lords Report, 2011). Also, we acknowledge that telling people that the nudge is used to influence their decision potentially evokes a perceived threat to their freedom, leading them to experience psychological reactance. The latter can be defined as "the motivational state that is hypothesized to occur when a freedom is eliminated or threatened with elimination" (Brehm and Brehm, 2013, p. 37). This could not only inhibit the effect of the nudge but could even lead to the opposite effect than the one intended. We presumed that experiencing reactance is mitigated when information on its purpose substitutes or complements the nudge. According to Salience Theory (Bordalo et al., 2012a), providing the purpose increases the degree to which the ultimate goal of the nudge, relative to its means of behavioral influence, is taken into account during the decision process. This hypothetically reduces the propensity to elicit a state of psychological reactance. Therefore, this phenomenon is important when investigating the influence of different types of transparency on the effectiveness of nudges. We report evidence from a laboratory experiment where subjects can contribute to real climate protection. The nudge is a default value that intends to increase contributions. Such a default in a public goods context, unlike nudges aiming to improve individual outcomes, attempts to increase positive external effects that only benefit the individual in the aggregate, but affords them to forfeit immediate personal economic gains. Thus, this context is more likely to produce a state of psychological reactance, and is thus suitable for testing it.

A default value alters the salience differences within a decision problem with increasing the salience of the default value and decreasing the salience of all other options. In addition, the presentation of a default value may result in different mechanisms through which it potentially influences behavior, e.g., as a reference value and anchor (for construction of preferences), through provision of social norms or information, or through inertia (by imposing pecuniary or cognitive costs on deviating from the default). Sunstein and Reisch (2016) provide a review on default-mechanisms. Note that Cappelletti, Mittone, and Ploner (2014) provided evidence from a public good game that defaults do not work as recommendations, i.e., as information provision in such a context. We expected the default value to increase contributions through two possible ways. First, it can increase the fraction of people picking the default value. Second, it can induce people to increase their contribution towards this value. We discuss our possible mechanisms in the second section and relate them to our findings in the last section.
The type of transparency that accompanies the default varies across treatments and consists of either informing decision-makers about its potential behavioral influence and/or informing them about its purpose to increase contributions to climate protection. After the experiment, we assessed two different measures of psychological reactance. Thus, we tested whether the influence of transparency is limited to a sub-group of participants distinct in their proneness to show psychological reactance (trait reactance). Additionally, we tested whether transparency influences the perception of a nudge as a threat for freedom of choice, and whether it functions as a source of anger (state reactance).

Recent findings from Arad and Rubinstein (2017) illustrate why our investigation of transparency and psychological reactance in the context of nudges is important. Their findings suggest that some subjects may consciously act contrary to the encouraged action, presumably in order to protest against the intervention of the government. The authors argue that full transparency of nudges, thus, may even lead to the opposite outcome than the one intended (as opposed to simply eliminating the effectiveness of a nudge). Some people behave in a completely different way simply out of protest against being manipulated. Contrary to this argument, findings by Sunstein (2016) from a nationally representative survey in the USA show that there is widespread support for nudges, and that transparency concerning the nudge will not diminish its effectiveness. Reisch and Sunstein (2016) show that there is also a general support of nudges in six European countries.

To the best of our knowledge, there are only three empirical studies directly relevant to our research question. Loewenstein, Bryce, Hagmann, and Rajpal (2015), in a laboratory experiment, found no evidence that informing subjects that they were presented with a pro-self default option influences their effectiveness. Similarly, Kroese, Marchiori, and de Ridder (2016), in a field experiment, found no evidence that making subjects aware of the purpose behind a pro-self default has any effect. Steffel, Williams, and Pogacar (2016), in several hypothetical and marginally incentivized consumer-related experiments, found no evidence that stressing the potential behavioral influence of a pro-self, as well as a pro-social default impacts their effectiveness, although it affects perception by the consumer.

While existing evidence unanimously suggests the impact of transparency on effectiveness of nudges is absent, our research augments the extant literature in various ways. First, subjects in our experiment face a trade-off between real monetary payoffs and real contributions to a (global) public good. By contrast, two of the previous studies employed relatively abstract and stylized environments, and did not demand subjects to make (substantial) financial tradeoffs. Although Kroese et al. (2016) investigated behavior in the field, they did neither study pro-social nudges, nor did they incorporate both types of transparency. Second, we investigated the distinct, as well as combined effect of two types of transparency on the default effect. Previous research focused exclusively on either of these two categories. However, there are reasons to expect that informing decision-makers about the potential behavioral influence of a nudge has different consequences than informing them about its purpose. Third, we enriched our analysis with the concept of psychological reactance, allowing for a deeper understanding of potential channels through which transparency influences default effects. Recent research on nudges, although focusing conceptually on the role of reactance (Arad & Rubinstein, 2017; Hedlin & Sunstein, 2016), did not investigate its interaction with transparency.

Consequently, we contribute to the topic of transparency of nudges in various ways. First, we enable a more nuanced view by investigating two types of transparency, thus contributing to a better understanding on how transparency works and whether policy-makers can make nudges more transparent without diminishing effectiveness. Second, our experimental setup, albeit controlled, sets up a realistic context, enabling us to make more valid inferences about the impact of transparency on nudges in ”the real world”. Third, we widen the discussion on transparency by investigating its connection to the concept of psychological reactance.

To preview our results, defaulted contributions are significantly higher than in the control group, even when accompanied by either type of transparency, including both types. In addition, contributions in the treatment groups (with or without transparency) do not significantly differ from each other. Thus, we replicate the lack of an effect of transparency, indicated by evidence from the studies outlined above. Finally, we neither find evidence that trait reactance interacts with transparency, nor that transparency changes the perception of nudges as freedom threatening or sources of anger. Therefore, our findings advocate that nudges (in the form of defaults) can be transparent and effective.

The remainder of this chapter is structured as follows. In section 10.2 we discuss psychological reactance as a conceptual background to covert nudges, followed by derivation of behavioral predictions. We lay out the experimental design in section 10.3. In section we present and analyze the results. Section 10.5 concludes.

10.2. Conceptual Framework and Behavioral Predictions

Since Brehm (1966) introduced the theory of psychological reactance, many studies have explored this phenomenon. Social influence attempts (such as nudges) that are detected by an individual may be perceived as a threat to freedom of choice (Brehm, 1966). The elicited state of psychological reactance may result in behavioral and cognitive efforts to reestablish freedom as well as uncomfortable, hostile, aggressive, and angry feelings (Dillard and Shen, 2005). Consequently, people may try to restore their freedom by exhibiting exactly the restricted behavior, thus, in our case, strongly deviating from the default value. In addition, they may devaluate the source of threat (the initiator of the nudge), increase their liking for the restricted freedom, or counter-argue against the imposed option (Brehm, 1966; Dillard & Shen, 2005). People react in such a manner not only to obvious and direct, but also to subtle and subliminal threats (Chartrand, Dalton, and Fitzsimons, 2007).

In order to investigate whether transparency influences the effectiveness of pro-social nudges, specifically defaults, we chose the context of climate protection. With climate change being one of the major challenges faced by society on a global scale today, information-based instruments and nudges are becoming increasingly important to increase individual contributions to climate and environmental protection (Allcott & Mullainathan, 2010; Araña & León, 2013; World Bank, 2015).

One way to contribute to climate protection is to offset (parts of) one's own yearly CO_2 emissions by donating to specific charitable organizations (in the experiment, referred to as 'climate protection fund'). These organizations use donations to purchase and delete carbon emission licenses from the European Union Emissions Trading Scheme (EU ETS).[43] Buying carbon licenses is an effective way for individuals to contribute to climate protection, when compared to, e.g., electricity-saving (Perino, 2015). Therefore, individual payment for carbon license retirement is a relevant context in which the influence of transparency on the effectiveness of a pro-social nudge can be investigated.

Based on psychological reactance theory we expected that mentioning the potential influence of a default will evoke the most reactance and thus reduce its effectiveness. In contrast, the sole provision of the purpose, i.e., climate protection, should evoke little reactance since this induces perspective taking. In addition, it renders the positive goal of the contribution more salient. According to Salience Theory, more salient attributes will be over-weighted in the decision process. Based on this argument, providing the

[43] The EU ETS is a European market that ultimately prices carbon emissions and allows regulated industries to trade their emission rights. Buying licenses off the market increases the scarcity of emission rights, resulting in higher prices and thus increasing the incentives for regulated firms to invest in emission-reducing technology.

purpose will work as an additional nudge and thus increase the default effect. Finally, accompanying the default with both types of information will be the most transparent form of the nudge. Due to combining the hypothesized ”downside” effect of reactance and ”upside” effect of the salience of the purpose of the nudge we expected the contribution level to be in between the other treatments. In sum, hypotheses concerning people's contribution decisions in the presence of the default are as follows:

H 1. If participants are confronted with a default, contributions will be higher compared to when there is no default.

H 2. If participants are informed that the default may have an influence on their decision, contributions will be lower compared to when they are not informed.

H 3. If participants are informed of the purpose of the default, contributions will be higher compared to when they are not informed.

H 4. If participants are informed of the potential influence of a default and of its purpose, contributions will be higher than with information solely on influence and lower than with information solely on purpose.

Although it was not the purpose of this experiment to identify the mechanism underlying the potential default effect, hypothesizing about a transparency-effect relies on certain assumptions regarding this mechanism. Transparency can only exert an effect if subjects are aware of the transparency and consequently of the default. This necessity rules out default effects that rely on unawareness (Madrian and Shea, 2001). If defaults work via costs of opting out (Johnson and Goldstein, 2003), providing a reference point (Dinner, Johnson, Goldstein, & Liu, 2011; Samuelson & Zeckhauser, 1988) or an anchor (Dhingra et al., 2012), transparency could have an impact.[44] More precisely, information regarding the potential influence of the default then increases the awareness of decision-makers to the manipulated structure of the decision. This in turn then may cause reactance. Mentioning the purpose of the default and thus justifying its use has the potential to mitigate reactance. However, note that Wilson, Houston, Etling, and Brekke (1996) observed anchoring effects despite forewarning, suggesting an unintentional and subconscious working mechanism that could also apply to defaults working as anchors. If defaults work as an implicit recommendation (McKenzie, Liersch, and Finkelstein, 2006), a persuasion attempt (Brown and Krishna, 2004), or a coordination device (Cappelletti et al., 2014) it is less clear whether transparency has an effect. Informing decision-makers on the potential influence

[44]Note that the potential impact can vary considerably between these mechanisms, and that it can also be close to zero. The point is that here, as opposed to the case of unawareness, transparency could logically influence the default effect.

given their interpretation of the default as a recommendation, persuasion attempt, or coordination device would provide no additional information, because decision-makers would already be aware of this potential influence. Mentioning the purpose would increase the salience of the climate protection goal, causing a similar effect as when any of the previous mechanisms is at play. When analyzing findings with respect to psychological reactance, we hypothesized that trait reactance interacts with the type of transparency accompanying the default value. Specifically, we expected that:

H 5. If participants are informed that the default may have an influence on their decision, the default effect for participants with higher trait reactance will be lower than for participants with lower trait reactance.

We further hypothesize that the evaluation of a default as freedom-threatening, autonomy-decreasing, manipulative, and pressuring (perceived threat to freedom), as well as its potential to elicit negative emotions (anger) differs with respect to the types of transparency accompanying the default value. Specifically, we expect that:

H 6: If participants are informed that the default may have an influence on their decision, experience of state reactance will be higher compared to when they are not informed.

We deduced hypotheses H5 and H6 exclusively with respect to a default accompanied by information on its potential influence, because we expected this type of transparency to increase the salience of the potentially manipulative and autonomy-threatening default-characteristic. For the purpose of the default, the conceptual link to reactance is less clear. We therefore abstained from formulating specific hypotheses.

10.3. Experimental Design

The laboratory experiment consisted of five experimental groups, of which one was the control group.[45] We conducted 11 sessions in the Econ-lab of the Erasmus School of Economics at the Erasmus University Rotterdam, the Netherlands, recruited with ORSEE (Greiner, 2015) in June 2016, and additional 15 sessions in July 2017 in the WiSo-lab of the University of Hamburg, Germany, recruited with hroot (Bock, Baetge, and Nicklisch, 2014). A total of 498 students participated in the experiment using the z-tree software (Fischbacher, 2007). Of these, 53.21% were female, the average age was 23.74 years (median: 23 years), and about half (53.01%) studied economics. More information on the differences between samples from both locations, as well as a disaggregated analysis of

[45]Prior to the experiment, pilot sessions were conducted in Germany, Sweden, France and the Netherlands. The pilot session in Germany focused on developing the design, which was further improved on and tested among Master students in the Netherlands, Sweden, and Bachelor students in France. The experimental design was not identical in all these pilots. Therefore, findings from the pilot sessions are not included in data analysis.

effect differences are provided in Appendix C.2. All participants were randomly assigned to separate computer terminals and were instructed not to communicate. They were given instruction sheets that were read aloud (see Appendix B.1). All participants received an endowment of 10 Euro and were asked to indicate how much (if any) of their endowment they would like to contribute to the 'climate protection fund'. The remaining amount was their private payoff. After the experiment, they were paid according to their decisions, and contributions were used to retire real carbon licenses from the EU ETS, through donations to 'TheCompensators*'.[46] In the control group, participants were presented with a text box where they could enter their contribution in any integer amount between 0 and 10 Euro. Neither a preselected default value for the contribution, nor any additional information were presented. In the other experimental groups, subjects encountered an 8 Euro default contribution in form of a button (see Figs. B.1 and B.2 in Appendix B). They could either press this button or choose another one that stated 'Different amount'. In the latter case they were referred to another screen that contained exactly the same information but with the addition of a text box where they could insert any amount between 0 and 10 Euro. In three of four default treatments, the default was complemented by a sentence that induced transparency, respectively on the default's potential influence, its purpose, or both. Table 10.1 shows the exact wording used to provide each type of transparency in the respective treatment group. The Default + Info transparency message informs subjects about the fact that they may be (subconsciously) affected by the default value. It resembles the wording by Steffel et al. (2016) which they use in order to deploy a default ethically. We expect that this wording stimulates the participants defensive systems against the threat to their behavioral autonomy, potentially motivating reactant behavior. The Default + Purpose transparency message informs subjects about the purpose of the default, i.e. increasing contributions to the climate protection fund. The wording implies the existence of a default effect, increases the salience of the purpose and, contrary to Default + Info, causes subjects to focus on the goal instead of the fact that it potentially threatens their behavioral autonomy. The Default + Info + Purpose combines both messages. Once subjects made their decision, they received information regarding their contribution, their private payoff and the amount of CO_2 that would be retired with the contributed amount.[47]

After making their decision, participants answered a questionnaire measuring, among others, their attributed importance to climate protection, and their belief in the effectiveness of retiring emission rights as a measure to protect the climate. In order to find out whether

[46] 'TheCompensators*' is a non-profit association founded in 2006 by researchers from the Potsdam Institute for Climate Impact Research. They offer a way for individuals and firms to compensate for their emissions. With donations, they buy and retire emission rights from the EU ETS. At the end of the experiment, all participants received an email with a confirmation and a certificate of aggregate experimental donations to 'TheCompensators*'.

[47] At that time, "TheCompensators*" offered to retire licenses at a price of 5.53 Euro. Note that this price can be different from the actual spot-price at the time we conducted the experiment, since "TheCompensators*" buy batches of licenses at a specific price and then retire them based on the donations they receive, irrespective of price-changes that appear in the meantime.

reactions to the different types of transparency can be explained by psychological reactance, we had two approaches. First, we assessed participants' perception of the default value as freedom threatening, autonomy-decreasing, manipulative, and pressuring, as well as its tendency to evoke negative emotional reactions, such as irritation, anger, annoyance, and aggravation. We refer to this as state reactance (Dillard and Shen, 2005). Second, we measured subjects' proneness to psychological reactance, referred to as trait reactance, with Hong's Psychological Reactance Scale (Hong and Faedda, 1996). Both measures were assessed after subjects made their decision of how much to contribute.[48] Relevant questions are in Appendix C.3. After conducting the sessions in Rotterdam, we calculated observed power for the most important tests. For H1, simulated post hoc observed power analyses produced power coefficients of 0.72, 0.26, 0.51, and 0.46, respectively for Control vs. Default, Control vs. Default + Information, Control vs. Default + Purpose, and Control vs. Default + Info + Purpose. Concerning Findings 2-4, post hoc observed power analyses for the estimates in model (1) produced power coefficients of 0.22, 0.87, 0.95, respectively for Default vs. Default + Information, Default vs. Default + Purpose, and Default + Info + Purpose vs. Default + Information vs. Default + Purpose. In order to further substantiate Finding 2, we conducted additional sessions for the Control group, Default, and Default + Information groups. The number of additional observations based is on an a priori power analysis. The simulation suggested that pooling data from all sessions allowed to detect a true difference of roughly 1.15 EUR (Cohen's d = 0.37) in mean contributions between the Default and Default + Information group in 78.81% of the time.

[48]We assume that measuring reactance items before treatments would have introduced an "additional nudge" with a potential influence on contributions. Kruskal-Wallis tests and Steel-Dwass-Critchlow-Fligner multiple comparison tests do not show any significant difference between treatments for all state and trait reactance items. This suggests there is no significant effect of treatments. However, we cannot completely exclude a potential common impact of all treatments on reactance.

Table 10.1.: Experimental design.

Experimental Group	Default Contribution	Transparency Information
Control	No	No information
Default	8 €	No information
Default+Info	8 €	"Please consider that the preselected default value might have an influence on your decision."
Default+Purpose	8 €	"Please consider that the preselected default value is meant to encourage higher contributions for the climate protection fund."
Default+Info+Purpose	8 €	"Please consider that the preselected default value might have an influence on your decision. This is meant to encourage higher contributions for the climate protection fund."

Notes: The table reports the experimental groups, the respective default value presented to participants, as well as the respective transparency information as it was shown to the subjects.

10.4. Results

We present and discuss findings in the following way: First, we demonstrate main results regarding the effectiveness of defaults and their interrelation with transparency. Second, we analyze the measures used to investigate the relevance of psychological reactance to transparency of defaults.

10.4.1. Default Effects

Overall, 498 subjects contributed 1385.5 Euro to retire carbon licenses, resulting in 2.78 Euro per subject. Of all participants, 68.27% contributed a positive amount, and 9.44% opted for the default value. Table 10.2 presents summary statistics of the variables divided by experimental groups. Fig. 10.1 presents the respective mean contributions.

Table 10.2.: Descriptive statistics of all outcome variables to assess the default effect.

	Contribution		Contributed	Picked default	n
Experimental Group	Mean	SD	Mean	Mean	
Control	1.82	2.66	51.76	0	85
Default	2.95	2.98	70.76	12.28	171
Default+Info	3.04	2.98	74.07	8.02	162
Default+Purpose	2.92	3.19	71.79	15.38	39
Default+Info+Purpose	2.85	2.95	65.85	17.07	41

Notes: The table reports summary statistics (means and standard deviations) of different outcome variables, as well as the number of subjects per experimental group. Outcome variables are: contributions to the climate protection fund, the percentage of subjects contributing a positive amount, as well as the percentage of subjects contributing the default value.

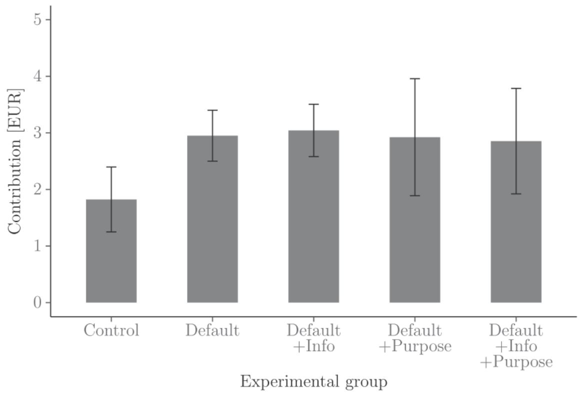

Figure 10.1.: Mean contributions per experimental group.

Notes: The figure shows mean contribution levels in the experimental groups. Error bars represent 95% confidence intervals.

Table 10.3.: Descriptive statistics of covariates.

	Age		Gender (Male)	Importance of CP	No exp. Experience	EU ETS not effective
Experimental Group	Mean	SD	Mean	Mean	Mean	Mean
Control	23.75	4.94	48.24	76.47	23.53	60
Default	24.16	4.29	43.27	82.46	29.82	60.23
Default+Info	23.92	4.53	45.06	88.27	25.93	56.79
Default+Purpose	22.28	4.65	53.85	51.28	20.51	64.1
Default+Info+Purpose	22.68	3.72	58.54	63.41	19.51	58.54

Notes: The table reports summary statistics (means and standard deviations) of different covariates per experimental group. Covariates are: age of participants, percentage of males, percentage of subjects perceiving climate protection as (very) important, percentage of subjects without prior experience with experiments, as well as the percentage of subjects judging license retirement as an ineffective mean for climate protection.

A Mann-Whitney test of H1 rejects the null hypothesis of equal contributions between Control vs. Default (W=5486, p=0.001), Control vs. Default + Info (W=4974, p=0.001), Control vs. Default + Purpose (W=1275, p=0.032), and Control vs. Default + Info + Purpose (W=1376.5, p=0.046). Overall, we find evidence for a default- and pull-effect. To check robustness of the default effect we focused on contributions as an outcome variable in Tobit regression. The Tobit model accounts for left-censored contributions and allows testing effects on the latent, unobserved contribution variable. This means we assume that at least some subjects would choose to take from instead of contribute to the public good. Thus, we interpret the dependent variable as desired contributions, and indeed even damages, to climate protection. This assumption is common in dictator games and empirically valid (Engel, 2011).

We began with a restricted model limited to the treatment variable, then added a dummy variable indicating that subjects perceive climate protection to be (very) important, and proceeded to add other relevant covariates shown in Table 10.3. The reason we added importance to protect the climate separately is that a Chi^2 -Test rejects the hypothesis that subjects are equally distributed among the treatment groups with respect to this variable

Table 10.4.: Stepwise Tobit-models with and without interaction term.

	(1) Contribution	(2) Contribution	(3) Contribution	(4) Contribution	(5) Contribution	(6) Contribution
Default	1.868***	1.718***	1.659***			
	(0.587)	(0.571)	(0.539)			
Default+Info	2.056***	1.758***	1.670**	0.165	0.152	0.00216
	(0.586)	(0.577)	(0.538)	(0.438)	(0.429)	(0.410)
Default+Purpose	1.866**	2.612***	2.528***	???0.0343	0.858	0.841
	(0.839)	(0.845)	(0.784)	(0.730)	(0.750)	(0.775)
Default+Info+Purpose	1.628*	1.921**	1.896**	???0.260	0.169	0.174
	(0.829)	(0.779)	(0.779)	(0.726)	(0.670)	(0.756)
Importance of CP		2.806***	2.350***		2.810***	2.353***
		(0.517)	(0.502)		(0.558)	(0.534)
Gender (Male)			-1.045***			-1.065***
			(0.353)			(0.391)
Age			-0.0406			-0.0200
			(0.0403)			(0.0431)
No exp. Experience			-0.577			-0.522
			(0.425)			(0.451)
EU ETS not effective			-2.512***			-2.329***
			(0.347)			(0.368)
Hamburg			-0.0494			-0.102
			(0.453)			(0.504)
React				-0.0897	-0.0977	-0.0783
				(0.106)	(0.102)	(0.0971)
Default + Info ? React				-0.108	-0.109	-0.0764
				(0.145)	(0.141)	(0.133)
Default + Purpose ? React				0.183	0.208	0.114
				(0.276)	(0.285)	(0.250)
Default + Info + Purpose ? React				0.0646	0.0316	-0.0483
				(0.224)	(0.206)	(0.190)
Constant	0.357	-1.824***	1.986*	2.259***	-0.0734	3.07***
	(0.497)	(0.644)	(1.094)	(0.314)	(0.563)	(1.100)
Sigma	3.969***	3.848***	3.591***	3.888***	3.766***	3.550***
	(0.152)	(0.153)	(0.143)	(0.153)	(0.152)	(0.147)
Observations	498	498	498	413	413	413
Log Pseudolikelihood	-1088.416	-1071.872	-1038.671	-929.4	-915.187	-890.107
F	(4, 494) = 3.33	(5, 493) = 8.64	(10, 488) = 13.19	(7, 406) = 0.76	(8, 405) = 3.98	(13, 400) = 7.36
Prob > F	0.010	< 40.001	< 0.001	0.624	< 0.001	< 0.001
Pseudo R^2	0.007	0.022	0.052	0.002	0.018	0.044

Notes: The table reports estimates of Tobit models with contributions censored at 0 as the dependent variable, with and without interaction terms. Robust standard errors are in brackets. Default + Info, Default + Purpose, and Default + Info + Purpose denote the respective treatment group, with Default as the base category. React measures subjects' proneness to experience reactance in a metric scale, and is mean centered. Def + Inf × React, Def + Pur × React, and Def + Inf + Pur × React are interaction terms of the transparency type with proneness to experience reactance. Importance of CP is a dummy that takes the value 1 if the subject perceives climate protection as (very) important. Gender takes the value 1 if the subject is male. Age denotes the age of the subject. No exp. Experience is a dummy which takes the value 1 if a subject did not participate in another experiment before. EU ETS not effective is a dummy that takes the value 1 when a subject judges license retirement as an ineffective mean for climate protection. Hamburg takes the value 1 if the subject is from the Hamburg, as opposed to the Rotterdam sample.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

($\chi^2(4) = 34.37, p < 0.001$). By controlling for this variable we ensure that estimates of treatment effects are not conditionally biased. Because the questionnaire was taken by subjects after being exposed to treatments, there is a risk of the respective manipulations being the reason for the differences in importance-ratings. Regarding Tobit models in Table 10.4, un-restricted model (3) includes all covariates, i.e., rating of the importance of climate protection, gender, age, no previous experience with experiments, judgment of buying emission licenses from the EU ETS as an ineffective tool for climate protection, and a location dummy. Model (1) predicts that a mere default, a default plus info, and a default plus its purpose lead to higher average contributions compared to no default. The effect of Default + Info + Purpose is marginally significant. When controlling for subjects' perception of the importance of climate protection in model (2), coefficients change. This results in significance for Default + Info + Purpose. Importance of CP positively predicts the latent contribution variable. A likelihood-ratio test suggests that model (2) fits the data significantly better than model (1) ($\chi^2(1) = 33.09, p = 0.001$). Controlling for additional covariates increases precision of the estimated average treatment effects. A likelihood-ratio test suggests that un-restricted model (3) fits the data significantly better than restricted model (2) ($\chi^2(5) = 66.40, p = 0.001$).

F 1. There is a default effect on contributions for a default, a default plus information, a default with added purpose, as well as for a default with both types of transparency.

10.4.2. Influence of Transparency on Default Effectiveness

A Kruskal-Wallis test for equal contribution distributions in the treatment groups is not significant ($H(3)$=0.484, p=0.922). So are respective pairwise comparisons with Dunn's test (not reported). Consequently, there is no evidence for either of H2, H3, and H4. As above, we augmented our analysis by focusing on contributions in stepwise Tobit-regression (Table 10.4). In un-restricted model (3), an omnibus Wald-test for equality of parameter estimates for Default, Default + Info, Default + Purpose, and Default + Info + Purpose does not lead us to reject the null hypothesis (F=(3,488)=0.49, p=0.692). The same holds for the restricted models. There is no evidence of unequal contributions in the treatment groups. Consequently, there is no evidence that transparency significantly reduces contributions.[49]

F 2. Informing participants that the default may have an influence on their decision does not significantly decrease contributions compared to when they are not informed.

F 3. Informing participants about the default's purpose does not significantly increase contributions compared to when they are not informed.

F 4. Informing participants that the default may have an influence on their decision, as well as of the default's purpose does not decrease or increase contributions, compared to the other types of transparency (including no transparency at all).

Of the additional covariates, Gender and EU ETS not effective are significant. Being male, as well as judging the EU ETS as not effective to protect the climate, negatively predict the latent outcome variable. The former finding is consistent with evidence from dictator games (Engel, 2011). Findings on gender differences in public good games are ambiguous, however (Croson and Gneezy, 2009). In the context of real contributions to

[49]Estimated treatment-effects of un-restricted regression models are plotted in Appendix C.1 (Figs. C.2-C.4).

climate protection, evidence by Diederich and Goeschl (2014), while suggesting that female subjects are less indifferent to climate protection, does not support a higher willingness to pay for emission certificates of women. Findings with respect to age somewhat align with those of Borghans and Golsteyn (2015) who found, in a less restricted sample, that the default effect does vary with age. However, at around 22 years (the mean of our sample) they found a relatively large default effect. This may explain why we find a default effect, but no effect of age.

10.4.3. Psychological Reactance and Transparency

To test if reactions towards the combination of a default value with different types of transparency can be explained by psychological reactance, we measured subjects' proneness to experience psychological reactance.[50]
Specifically, we tested whether subjects' reactions towards different types of transparency accompanying the default differ depending on subjects' trait reactance. Therefore, we ran regressions with an interaction term of the treatment variable and the trait reactance index. The latter is centered on the mean, so that treatment-main-effects are meaningful (Table 10.4). Note that this regression excludes observations from the control group. For reasons of brevity, we focused on the main effects of trait reactance, as well as on interaction-effects. As in previous Tobit models, model (5) fits the data better than model (4) ($\chi^2(1) = 28.42, p = 0.001$), and model (6) fits the data better than model (5) ($\chi^2(4) = 50.11, p = 0.001$). We find no significant main effect of trait reactance, nor do we find that the different types of transparency and the trait reactance index interact significantly for any of the three model-specifications. In other words, there is no evidence that the effect of different types of transparency on average contributions is conditional on subjects' trait reactance.

F 5. The influence of information on the default effect does not depend on the level of trait reactance of participants.

In order to test whether reactions to different types of transparency can be explained by psychological reactance, we created an index for each of the two state reactance-categories, i.e., for the perceived threat to freedom and the anger-category.[51] Then, we added the respective dummies in each category, to form two indexes, each ranging from zero to four. Findings are consistent for when both dependent variables are included as (un-weighted) factor-based scores in linear OLS-regression.

[50] To create an index for trait reactance, we constructed dummy variables for each of the 14 items of the scale, which are equal to 1 when the subject responded with "Agree" or "Strongly agree" to the respective question, 0 otherwise. We then added the dummies for each subject to create the index, which ranges from 0 to 14. Findings are consistent for trait reactance included as a (un-weighted) factor-based score.

[51] We constructed a dummy-variable, which is equal to 1 when the subject "agreed" or "strongly agreed", resp. replied with "to some extent" or "very" to the respective statements, for each item (see Appendix C.3).

We modeled the log odds of subjects being in a higher level of each of both ordinal indexes on all explanatory variables used above (Table 10.5). Note that this regression excludes observations from the control group since subjects in this group were not presented with the default option which they could rate. None of the coefficients modeling treatment effects are significant.[52]

F 6. Combining the default with information about its potential behavioral influence does not increase participants' experience of state reactance.

Age negatively predicts experienced anger triggered by the default value. The finding that experiencing negative emotions decreases with age is known in the literature (e.g., Charles, Reynolds, & Gatz, 2001).

Both approaches that are linking different types of transparency of a default to psychological reactance suggest that subjects neither perceive a default value differently based on the type of transparency accompanying it, nor does their inherent propensity to show psychological reactance change the way they react to these different types of transparency.

[52]This finding is consistent with non-parametric tests for differences of individual items of the scales (not reported).

Table 10.5.: Ordered logistic model of state reactance.

	(1) Threat to freedom	(2) Anger
Default + Info	-0.00294	-0.167
	(0.199)	(0.223)
Default + Purpose	-0.0297	0.0868
	(0.418)	(0.453)
Default + Info + Purpose	-0.0686	-0.560
	(0.330)	(0.470)
Importance of CP	-0.0275	-0.334
	(0.232)	(0.276)
Male	-0.0798	-0.300
	(0.190)	(0.217)
Age	-0.0594***	-0.0832**
	(0.0183)	(0.0268)
Participated	-0.0221	-0.0560
	(0.192)	(0.242)
EU ETS not effective	0.183	0.173
	(0.191)	(0.216)
Hamburg	-0.0120	-0.325
	(0.250)	(0.260)
Cut 1	-3.125***	-2.029***
	(0.528)	(0.683)
Cut 2	-2.270***	-1.126*
	(0.524)	(0.679)
Cut 3	-1.088**	-0.251
	(0.517)	(0.685)
Cut 4	0.346	0.508
	(0.525)	(0.718)
Observations	413	413
Log Pseudolikelihood	-640.583	-443.190
Wald $Chi^2(9)$	12.96	19.80
Prob>Chi^2	0.165	0.019
Pseudo R^2	0.008	0.024

Notes: The table reports estimates of ordered logit models with ratings of defaults as threatening to freedom, and anger arousing as the respective dependent variable. Robust standard errors are in brackets. Default + Info, Default + Purpose, and Default + Info + Purpose denote the respective treatment group, with Default as the base category. Importance of CP is a dummy that takes the value 1 if the subject perceives climate protection as (very) important. Gender takes the value 1 if the subject is male. Age denotes the age of the subject. No exp. Experience is a dummy which takes the value 1 if a subject did not participate in another experiment before. EU ETS not effective is a dummy that takes the value 1 when a subject judges license retirement as an ineffective mean for climate protection. Hamburg takes the value 1 if the subject is from the Hamburg, as opposed to the Rotterdam sample.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

10.5. Discussion and Conclusion

The experiment advances the discussion of nudges and transparency by providing empirical evidence on the effect of transparency on the performance of a pro-environmental default value. Despite the widespread application of nudges, many researchers and consumers are concerned of the potentially manipulative nature of behavioral interventions. In democratic societies, public authorities are expected to be transparent with regard to their actions and intentions. Therefore, covertly 'exploiting' people's psychological biases potentially inhibits perceived legitimacy, and ultimately effectiveness of such policies. The most straightforward solution to this problem is to instruct policy-makers to disclose information regarding the potential influence of the nudge, and its purpose. However, this suggestion raises the concern that nudges will no longer be effective. As expressed by Bovens (2009), nudges "work best in the dark". The results of this study suggest that this concern might be overstated.

The experiment provides evidence that defaults increase contributions to climate protection even when complemented by disclosure regarding the potential influence of the default, its purpose, or both. Furthermore, there is no evidence that information on the potential behavioral influence and/or purpose of the default triggers psychological reactance. Likewise, there is no evidence that subjects differing in their proneness to experience reactance also differ in how they react towards the default with additional information.

These findings suggest that despite the initial concern over the inhibiting influence of transparency, nudges in the form of defaults can be transparent and at the same time effective. In order to preserve the effect of defaults and increase the legitimacy of behaviorally informed policies, policy makers should be transparent about their motives, as well as the potential behavioral influence of the instrument. The motive and how it is perceived by the decision-maker has been found to matter for advice (Kuang, Weber, and Dana, 2007).

Our findings replicate and add to previous evidence on the influence of transparency. Loewenstein et al. (2015) and Kroese et al. (2016) reported that pro-self defaults were effective in health contexts even after disclosing information about them. Our study extends this conclusion to pro-social nudges, a type that is widely used in the context of public policy-making. Moreover, we extend findings of Steffel et al. (2016) by examining the influence of transparency in a more realistic setting where participants' decisions have an actual consequence for them, and for the environment. Findings are also useful for the private sector and non-governmentalorganizations (NGOs) aiming to include nudges in their inventory to increase contributions to environmental protection, and possibly other public goods, e.g., charity.

Although several recent studies link nudges to psychological reactance, they do so either indirectly, or they deal with hypothetical and attitudinal, instead of behavioral outcomes

(Arad & Rubinstein, 2017; Haggag & Paci, 2014; Hedlin & Sunstein, 2016; Loewenstein et al., 2015). By measuring both state and trait reactance, we enable a more direct way of assessing the interaction of psychological reactance with the influence of transparency on the effectiveness of a default value. To our best knowledge, Goswami and Urminsky (2016) is the only study that assessed the interaction of trait reactance with the size of a default value on behavioral outcomes, i.e., charitable giving. They find no significant interaction effect. On a more general level, our findings, in line with theirs, suggest that psychological reactance plays a minor or no role with respect to behavioral effects of defaults, and, in our case, transparency. In fact, a possible explanation of this might be the relatively high default value, which is 80% of the experimental endowment. Instead of eliciting psychological reactance, such a high default might lead subjects to ignore it altogether.

Findings suggest that the default value is an effective way of increasing individual voluntary contributions to climate protection. Increased aggregate contributions are consistent with inertia, as well as anchoring. A higher fraction of participants picking the default value instead of specifying another amount in the default, compared to the control group, supports the inertia/effort reduction explanation. However, deviation costs in the experiment are marginal (the subject had to make two mouse-clicks, as well as to type in the contribution amount, instead of just making one mouse click on the default button), and contributing the default value is also consistent with an anchoring explanation: Subjects may choose the default value not only because of inertia, but also because they consider this value first and only then employ reasons against it, conditional on what they wanted to contribute initially. This anchoring-explanation is consistent both with picking the default and moving towards the default, whereas inertia is only consistent with the former behavior (Dhingra et al., 2012).

We observe that subjects who contribute a positive amount do contribute more on average when there is a default value with either type of transparency, but the differences to the control group are not significant. Additionally, we observe an increase of subjects giving a positive amount due to the default, which is consistent with the anchoring explanation. Together, our findings suggest that increased aggregate contributions in the default groups are due to an increase of the fraction of subjects contributing, as well as of an increase of the fraction of subjects choosing the default value, but not because of increased average contributions of subjects that contribute. Inertia, as well as anchoring may therefore both be reasons for why we observe default effects. Intuitively, we would expect anchoring to play a more pronounced role in real world applications of pro-environmental nudges, especially if defaults result in repeated and/or significant financial costs. For someone who highly values environmental and climate protection, deviating from a default, which may be perceived as conveying information about social norms, can incur non-financial costs, especially if he or she aims to uphold a positive self-image. Maintaining a positive self image, as well as being consistent with social norms, can be achieved by decreasing (not necessarily closing) the gap between default value and initially intended contribution. Note that our design did not allow to unambiguously identify the underlying mechanisms causing the default effect in the experiment. Anchoring is consistent with the interpretation of the default value as an implicit recommendation, a persuasion attempt, coordination device, or a reference point. If a decision-maker regards the default as an implicit recommendation, she may consequently increase/decrease her donation relative to her preferences, after seeing the default. However, we cannot identify whether she interpreted the default as a recommendation.

Furthermore, while being able to differentiate between the effects of different types of transparency is insightful for policy-makers, the difference between the information and purpose treatments is not analytically clear. Communicating the purpose of the default implicitly reveals that the default is expected to have an effect on individual decision-making, without spelling it out. Still, we think that the findings concerning this type of transparency are important for practical purposes.

Further research could evaluate the role of trait reactance on how subjects respond to different types of transparency for different types of nudges, i.e., social norms or framing. Additionally, building on the shortcoming of our experimental design, further studies should further investigate the link between transparency and the different underlying working mechanisms of defaults and other types of nudges. Since our experiment has a rather limited amount of subjects, field experiments can establish statistically more powerful findings for interaction effects. Due to a more realistic context, a field experimental approach would also increase external validity. Nevertheless, our experiment is less abstract than a 'regular' laboratory experiment due to the fact that contributions have a real effect on climate protection (Harrison & List, 2004). The current study focuses on one type of nudge, and a specific context. Further research is needed in order to determine the overall influence of transparency on the effectiveness of nudges. Moreover, results might be context-specific, thus requiring further investigation into pro-social nudges. Delving into the welfare implications of transparency can also become a promising research endeavor (Sunstein, 2015).

Overall, our findings advance the understanding of how nudges in general, and defaults specifically, affect individual behavior with social consequences, and how policy-makers can increase their transparency without limiting their effectiveness.

Part IV.

Conclusion

This work has shed light on human decision-making as complex process. It is shaped by many influence factors. Among these, attention, and the closely related concept of salience, play an important role. The salience of a stimulus is determined by bottom-up and top-down factors. Bottom-up describes the sensory information, e.g., the luminosity of a stimulus compared to its context, that shapes the salience of a stimulus. Top-down processes embrace emotions, goals, drives, etc., that also influence the salience of a stimulus. That means, the salience is influenced by factors that also directly exert influence on the decision-making process. Thus, the salience of a stimulus reflects a broad set of factors that influence our decisions. Therefore, it has been proposed that the inclusion of salience in a theory of human decision-making can be very useful in order to embrace many possible influences while the theory remains simple and practical.

In 2012, Bordalo et al. (2012a) proposed a theory on human decision-making that is based on salience. They suggest that salience differences within a decision problem may explain many decision biases. Concerning decisions under risk, Bordalo and colleagues developed a formula to calculate salience differences that are shaped by bottom-up processes. To my best knowledge, this salience approach is based on a theoretical derivation and has not been validated experimentally. This research gap has been addressed in Part II of this work. Therefore, an experiment has been conducted in order to investigate salience differences in risky lotteries as proposed by Salience Theory. Reaction times in a dot-probe task served as indicator of attentional biases. The data revealed a significant salience effect after a lottery exposure duration of 150 ms. This supports the salience concept proposed by Bordalo et al. (2012a) and suggests an early attentional orienting towards salient payoffs. In contrast, after a lottery exposure of 4000 ms, reaction times revealed no salience differences. This leads to the conclusion that later controlled strategic attentional processes are not affected by salience differences in risky lotteries.

In order to further differentiate attentional processes involved in the salience effect EEG has been recorded. Different ERP-components may indicate attentional biases at different stages of attentional processing and give a hint at more detailed reasons behind the salience effect. All investigated components, namely, P1, N1, P3a and P3b, showed no significant salience differences. On the one hand, this may put in question the validity of behavioral results after a lottery exposure of 150 ms. On the other, the experimental design may have been too complex to investigate the salience effect with EEG data. Therefore, future research should use a simpler design to examine attentional biases within risky lotteries. Furthermore, the dot-probe task was developed in order to assess attentional biases based on behavioral data. A measure that might be more suitable than the dot-probe task for assessing attentional biases based on EEG data should be used. In addition, other methods like eyetracking or similar could be applied.

Before subjects participated in the experiment they answered online questionnaires on achievement motivation and proneness to optimism or pessimism. Achievement motivation and proneness to optimism and pessimism might be relevant factors that influence the salience within risky lotteries via top-down processes. Therefore, the data was gathered in order to examine a possible top-down influence of these factors on salience. Unfortunately, not enough data could be collected to include those variables in the analysis on subject level. Incorporating top-down factors in the determination of salience differences within decision problems could be a promising topic for future research. As described previously, top-down processes play an important role in shaping the salience of a stimulus. Therefore, the calculation of salience differences within a decision problem only based on sensory (bottom-up) information may not capture the subjectively perceived salience properly. Finding important top-down influence factors on salience within risky lotteries and assessing the respective contribution to the overall perceived salience might be an important step towards the comprehension of how salience influences risky decisions. Furthermore, this could improve a theory on risky decision-making based on the influence of salience, like Salience Theory from Bordalo et al. (2012a). However, it could also be that, when describing risky decisions not on a subject level but on an aggregate level, as theories on decision-making aim for, the sensory information shaping the salience differences within decision problems might be crucial. This information is inherent in the decision problem and does not vary substantially from subject to subject. In contrast, top-down influences on the perceived salience may differ extensively on the subject level. In a diverse set of decision-makers and a neutral presentation of decision problems, these top-down influences could cancel each other out on the aggregate level and therefore be not relevant in order to describe the average risk behavior. However, when assessing attentional biases on the subject level as in the presented experiment, top-down factors could be crucial in determining the salience within the decision problem. This could be another explanation for EEG results showing no significant differences with respect to the bottom-up salience in risky lotteries.

Part III presents a further experiment that was devoted to nudges. These are interventions that often work by altering the salience within a decision problem or by directing the attention to the decision task itself. In the experiment a default option was used as a nudge. Many mechanisms are discussed to cause a default effect. Regarding the salience within a decision problem, the default increases the salience of the respective option. According to Salience Theory (Bordalo et al., 2012a) this can be an explanation for the default effect. Since these interventions influence decisions at least partly on an unconscious level, nudges are subject to criticism. Therefore, the experiment aimed at investigating the effect of transparent information accompanying the nudges on their efficacy. In line with previous research it could be shown that adding information on the nudge itself, on its purpose and the combination of both had no significant effect on the efficacy of the nudge, even though this additional information again alters salience ratios within the decision problem. Apparently, this change in our experiment was subtle enough not to interfere significantly with the default effect. This is underlinded by findings on psychological reactance which was also not significantly influenced by the additional information.

Overall, this work contributes to research in several ways. First, a special role of salience with respect to other influence factors on human decision-making has been proposed. Second, a new experimental paradigm to assess salience differences within risky decisions has been tested. From this, several improvements for follow up experiments have been proposed. Future research could realize these in order to gain a deeper understanding of attentional processes that are responsible for the salience effects within decision problems. Third, thougths that might improve a theory on human decisions based on salience were stated. Forth, it could be shown that modifying the salience within a decision problem by adding information on the influence of the nudge and its purpose is possible without significantly diminishing the performance of the nudge.

Appendices

A. Part II. Salience in Risky Decisions

A.1. Instructions

A.1.1. Original Version in German

Sehr geehrte Versuchsteilnehmerin, sehr geehrter Versuchsteilnehmer,

vielen Dank für Ihr Interesse an unserer wissenschaftlichen Studie. Bitte lesen Sie sich die folgenden Informationen zunächst sorgfältig durch und entscheiden Sie dann über Ihre Teilnahme oder auch Nichtteilnahme an dieser Studie. Beides, Ihre Teilnahme oder Nichtteilnahme stehen Ihnen frei. Falls Sie über diese Information hinaus noch weitere Fragen zur Studie haben sollten, beantworten wir Ihnen diese gern.

Sie werden in diesem Experiment ein Spiel spielen, bei dem Sie Geld gewinnen können. Das Spiel läuft folgendermaßen ab:

Sie spielen insgesamt 120 Lotterien[53]. Dabei werden Ihnen immer zwei Optionen dargeboten, die jeweils angeben, wieviel Sie mit welcher Wahrscheinlichkeit gewinnen können. Dies sieht beispielsweise wie folgt aus:

500	0,4	300	0,2
50	0,6	210	0,8

Wählen Sie die linke Option, können Sie hier mit einer Wahrscheinlichkeit von 40% 500 Cent und mit einer Wahrscheinlichkeit von 60% 50 Cent gewinnen. Wählen Sie die rechte Option, gewinnen Sie mit einer 20%-igen Wahrscheinlichkeit 300 Cent und mit einer 80%-igen Wahrscheinlichkeit 210 Cent.

Die Lotterien werden Ihnen nur sehr kurz gezeigt. Treffen Sie bitte immer eine Entscheidung, auch wenn Sie das Gefühl haben, nicht alles richtig gesehen zu haben.

Bevor Sie sich entscheiden, erfolgt noch eine kleine Zwischenaufgabe. Hierbei werden gleich nach der Lotterie an verschiedenen Positionen auf dem Bildschirm ein Punkt oder ein Quadrat gezeigt. Geben Sie bitte so schnell wie möglich an, was Sie gesehen haben. Diese Aufgabe hat keinen Einfluss auf Ihren Gewinn!

Geben Sie nun an, für welche Lotterieoption Sie sich entscheiden. Danach erfolgt der nächste Durchgang.

[53]Unfortunately, this was a mistake in the instructions. Overall, subjects completed 240 trials.

Nachdem Sie alle Lotterien gespielt haben, werden zufällig 10 davon ausgewählt. Ihr Gesamtgewinn aus diesen Lotterien wird Ihnen am Ende angezeigt und anschließend ausgezahlt. Der Gewinn kann bis zu 10 Euro betragen.

Während Sie spielen, werden wir Ihre Hirnaktivität mittels Elektroenzephalogramm (EEG) aufzeichnen. Durch das EEG erfassen wir passiv an der Kopfoberfläche kleinste Spannungsunterschiede, die durch die Aktivierung bzw. Deaktivierung größerer Nervenzellverbände bei verschiedensten psychischen Prozessen entstehen. Hierfür wird Ihnen vor dem Versuch eine EEG-Kappe aufgesetzt und ein leitfähiges Elektrodengel appliziert. Durch das Gel wird der Kontakt zwischen Kopfhaut und Elektrode verbessert, sodass die sehr kleinen Signale erfasst werden können. Die Anwendung der EEG-Messung folgt den in Medizin und Neurowissenschaften/Psychologie gängigen Standardmethoden und ist nach allgemeinem wissenschaftlichem und klinischem Wissen risikofrei.

Mit Vor- und Nachbereitung wird der Versuch etwa 1,5 Stunden dauern. Für Ihre Teilnahme erhalten Sie 6 Euro pro Stunde, zusätzlich wird Ihnen der Gewinn ausgezahlt. Sollten Sie die Studie vorzeitig abbrechen, werden Ihnen die bis dahin aufgewendete Zeit und die bis dahin erspielten Gewinne ausgezahlt.

Die Studie dient rein wissenschaftlichen Interessen. Unser Vorgehen birgt keinerlei Risiken für Sie. Ihre Daten werden nur im Rahmen unserer wissenschaftlichen Studie am Lehrstuhl für Biologische und Klinische Psychologie anonymisiert aufgezeichnet, weiterverarbeitet und ausgewertet. Ergebnisse werden nicht personenbezogen veröffentlicht. Alle Personen, die mit Ihnen und Ihren Daten in Kontakt kommen, sind zur Verschwiegenheit hierüber verpflichtet.

Ihre Teilnahme an der Studie ist freiwillig. Durch die Nichtteilnahme an der Studie entstehen Ihnen keinerlei Nachteile. Sie können Ihre freiwillige Teilnahme an der Studie jederzeit und ohne Angabe von Gründen abbrechen, ohne dass Ihnen daraus Nachteile entstehen. Auch der Studienleiter kann die Entscheidung treffen, die gesamte Studie abzubrechen oder Ihre Teilnahme vorzeitig zu beenden, wenn dies (etwa aus medizinischen Gründen) angezeigt sein sollte.

A.1.2. Translated Version in English

Dear participant,

thank you very much for your interest in our scientific study. Please read the following information carefully and decide afterwards whether you want to participate. The decision about your participation is fully up to you. In case you have further questions concerning the study that will not be answered in the following information we are happy to answer these.

During the experiment you will play a game with the possibility of winning money. The game will look like this:

Overall, you will play 120 lotteries[54]. There you will always have two options which tell you how much you can win and how probable that is. Here you can see an example:

500	0,4	300	0,2
50	0,6	210	0,8

If you choose the left option, you may win 500 Cent with a probability of 40% and 50 Cent with a probability of 60%. If you choose the right option, you may win 300 Cent with a probability of 20% and 210 Cent with a probability of 80%.

The lotteries will be presented only for a very short time. Please always choose an option, even if you feel you have not been able to look at all information.

Before you decide, there will be a little task in-between. Immedeatly after the lottery a circle or a square will be presented at different positions of the screen. Please indicate as fast as possible what you have seen. This task has no influence on your gain!

Now, please state the lottery option that you choose. Afterwards the next trial will follow.

After you completed all lotteries, 10 of them will be randomly chosen. The overall gain from these lotteries will be presented at the end and paid to you afterwards. The maximum gain is 10 Euro.

During the experiment the electrical activity of your brain will be recorded with an electroencephalogram (EEG). Thus, minimal voltage differences at your scalp are recorded that arise through activation and inactivation of larger clusters of nerve cells during different psychological processes. For this purpose an EEG cap will be placed on your head and a conductive gel will be applied before the experiment. The gel improves the contact between the scalp and the electrodes in order to record the very small signals. The EEG measurement is realized according to standard methods from medicine and neurosciences/psychology and is without any risk according to general scientific and clinical knowledge.

The experiment will take about 1.5 hours including preparation and review. You will receive 6 Euro per hour for your participation. In case you terminate the experiment prematurely, you will be paid for the time spent and the gain received until then.

The experiment serves purely scientific interests. Our approach does not carry any risk for you. Your data will be recorded, processed and analyzed anonymously as part of our scientific study at the Chair of Biological and Clinical Psychology. Published results will not be personally identifiable. All persons that have access to the data are committed to secrecy.

Your participation in the experiment is voluntary. You will not suffer any disadvantages if you decide not to participate. You can decide to terminate your voluntary participation in the experiment at any time without stating reasons and without any disadvantages for

[54]Unfortunately, this was a mistake in the instructions. Overall, subjects completed 240 trials.

you. The director of the study can also decide to terminate the whole experiment or your participation early if this should be necessary (e.g., for medical reasons).

A.2. Lotteries

Table A.1.: Lotteries.
Each lottery consists of two options with two payoffs (in Eurocent) (and probabilities).
Positions: (1) upper left, (2) lower left, (3) upper right, (4) lower right corner.

	Option 1		Option 2		Salient Payoffs (Position)		Position of Probe
Lottery	Position 1	Position 2	Position 3	Position 4	Option 1	Option 2	(s=salient)
1	200 (0.6)	50 (0.4)	200 (0.4)	100 (0.6)	50 (2)	200 (3)	2 (s)
2	25 (0.4)	100 (0.6)	100 (0.4)	50 (0.6)	25 (1)	100 (3)	2
3	600 (0.6)	150 (0.4)	300 (0.6)	600 (0.4)	150 (2)	600 (4)	1
4	250 (0.4)	1000 (0.6)	500 (0.6)	1000 (0.4)	250 (1)	1000 (4)	4 (s)
5	400 (0.2)	200 (0.8)	300 (0.7)	100 (0.3)	400 (1)	100 (4)	3
6	100 (0.8)	200 (0.2)	150 (0.7)	50 (0.3)	200 (2)	50 (4)	2 (s)
7	1200 (0.2)	600 (0.8)	300 (0.3)	900 (0.7)	1200 (1)	300 (3)	1 (s)
8	1000 (0.8)	2000 (0.8)	500 (0.3)	1500 (0.7)	2000 (2)	500 (3)	4
9	500 (0.2)	800 (0.8)	820 (0.6)	600 (0.4)	500 (1)	820 (3)	2 (s)
10	400 (0.8)	250 (0.2)	410 (0.6)	300 (0.4)	250 (2)	410 (3)	2 (s)
11	1500 (0.2)	2400 (0.8)	1800 (0.4)	2460 (0.6)	1500 (1)	2460 (4)	4 (s)
12	4000 (0.8)	2500 (0.2)	3000 (0.4)	4100 (0.6)	2500 (2)	4100 (4)	1
13	500 (0.7)	10 (0.3)	500 (0.65)	100 (0.35)	10 (2)	500 (3)	2 (s)
14	5 (0.3)	250 (0.7)	250 (0.65)	50 (0.35)	5 (1)	250 (3)	2
15	1500 (0.7)	30 (0.3)	300 (0.35)	1500 (0.65)	30 (2)	1500 (4)	1
16	50 (0.3)	2500 (0.7)	500 (0.35)	2500 (0.65)	50 (1)	2500 (4)	4 (s)
17	-200 (0.6)	-50 (0.4)	-100 (0.6)	-200 (0.4)	-50 (2)	-200 (4)	4 (s)
18	-25 (0.4)	-100 (0.6)	-50 (0.6)	-100 (0.4)	-25 (1)	-100 (4)	3
19	-600 (0.6)	-150 (0.4)	-600 (0.4)	-3 00 (0.6)	-150 (2)	-600 (3)	1
20	-250 (0.4)	-1000 (0.6)	-1000 (0.4)	-500 (0.6)	-250 (1)	-1000 (3)	3 (s)
21	-200 (0.7)	-500 (0.3)	-280 (0.9)	-480 (0.1)	-200 (1)	-480 (4)	3
22	-250 (0.3)	-100 (0.7)	-140 (0.9)	-240 (0.1)	-100 (2)	-240 (4)	2 (s)
23	-600 (0.7)	-1500 (0.3)	-1440 (0.1)	-840 (0.9)	-600 (1)	-1440 (3)	1 (s)
24	-2500 (0.3)	-1000 (0.7)	-2400 (0.1)	-1400 (0.9)	-1000 (2)	-2400 (3)	4
25	-800 (0.8)	-500 (0.2)	-600 (0.4)	-820 (0.6)	-500 (2)	-820 (4)	2 (s)
26	-250 (0.2)	-400 (0.8)	-300 (0.4)	-410 (0.6)	-250 (1)	-410 (4)	3
27	-2400 (0.8)	-1500 (0.2)	-2460 (0.6)	-1800 (0.4)	-1500 (2)	-2460 (3)	4
28	-2500 (0.2)	-4000 (0.8)	-4100 (0.6)	-3000 (0.4)	-2500 (1)	-4100 (3)	1 (s)
29	-10 (0.3)	-500 (0.7)	-100 (0.35)	-500 (0.65)	-10 (1)	-500 (4)	3
30	-250 (0.7)	-5 (0.3)	-50 (0.35)	-250 (0.65)	-5 (2)	-250 (4)	2 (s)
31	-30 (0.3)	-1500 (0.7)	-1500 (0.65)	-300 (0.35)	-30 (1)	-1500 (3)	1 (s)
32	-2500 (0.7)	-50 (0.3)	-2500 (0.65)	-500 (0.35)	-50 (2)	-2500 (3)	4
33	-50 (0.1)	-900 (0.9)	-400 (0.15)	-880 (0.85)	-50 (1)	-880 (4)	3
34	-450 (0.9)	-25 (0.1)	-200 (0.15)	-440 (0.85)	-25 (2)	-440 (4)	2 (s)

..continued on next page.

Lotteries (continued)

	Option 1		Option 2		Salient Payoffs (Position)		Position of Probe
Lottery	Position 1	Position 2	Position 3	Position 4	Option 1	Option 2	(s=salient)
35	-150 (0.1)	-2700 (0.9)	-2640 (0.85)	-1200 (0.15)	-150 (1)	-2640 (3)	1 (s)
36	-4500 (0.9)	-250 (0.1)	-4400 (0.85)	-2000 (0.15)	-250 (2)	-4400 (3)	4
37	-250 (0.3)	-15 (0.7)	-65 (0.9)	-240 (0.1)	-15 (2)	-240 (4)	3
38	-7,5 (0.7)	-125 (0.3)	-32,5 (0.9)	-120 (0.1)	-7.5 (1)	-120 (4)	1 (s)
39	-750 (0.3)	-45 (0.7)	-720 (0.1)	-195 (0.9)	-45 (2)	-720 (3)	3 (s)
40	-75 (0.7)	-1250 (0.3)	-1200 (0.1)	-325 (0.9)	-75 (1)	-1200 (3)	4
41	200 (0.5)	0 (0.5)	400 (0.2)	30 (0.8)	0 (2)	400 (3)	3 (s)
42	0 (0.5)	100 (0.5)	200 (0.2)	15 (0.8)	0 (1)	200 (3)	2
43	600 (0.5)	0 (0.5)	90 (0.8)	1200 (0.2)	0 (2)	1200 (4)	1
44	0 (0.5)	1000 (0.5)	150 (0.8)	2000 (0.2)	0 (1)	200 (4)	4 (s)
45	1600 (0.3)	1000 (0.7)	1300 (0.5)	1000 (0.5)	1600 (1)	1000 (4)	3
46	500 (0.7)	800 (0.3)	650 (0.5)	500 (0.5)	800 (2)	500 (4)	2 (s)
47	4800 (0.3)	3000 (0.7)	3000 (0.5)	3900 (0.5)	4800 (1)	3000 (3)	1 (s)
48	5000 (0.7)	8000 (0.3)	5000 (0.5)	6500 (0.5)	8000 (2)	5000 (3)	4
49	180 (0.2)	20 (0.8)	100 (0.4)	20 (0.6)	180 (1)	20 (4)	3
50	10 (0.8)	90 (0.2)	50 (0.4)	10 (0.6)	90 (2)	10 (4)	2 (s)
51	540 (0.2)	60 (0.8)	60 (0.6)	300 (0.4)	540 (1)	60 (3)	1 (s)
52	100 (0.8)	900 (0.2)	100 (0.6)	500 (0.4)	900 (2)	100 (3)	4
53	600 (0.45)	0 (0.55)	300 (0.9)	0 (0.1)	600 (1)	0 (4)	3
54	0 (0.55)	300 (0.45)	150 (0.9)	0 (0.1)	300 (2)	0 (4)	2 (s)
55	1800 (0.45)	0 (0.55)	0 (0.1)	900 (0.9)	1800 (1)	0 (3)	1 (s)
56	0 (0.55)	3000 (0.45)	0 (0.1)	1500 (0.9)	3000 (2)	0 (3)	4
57	-160 (0.3)	-100 (0.7)	-100 (0.5)	-130 (0.5)	-160 (1)	-100 (3)	3
58	-50 (0.7)	-80 (0.3)	-50 (0.5)	-65 (0.5)	-80 (2)	-50 (3)	3 (s)
59	-480 (0.3)	-300 (0.7)	-390 (0.5)	-300 (0.5)	-480 (1)	-300 (4)	4 (s)
60	-500 (0.7)	-800 (0.3)	-650 (0.5)	-500 (0.5)	-800 (2)	-500 (4)	1
61	-600 (0.45)	0 (0.55)	0 (0.1)	-300 (0.9)	-600 (1)	0 (3)	2
62	0 (0.55)	-300 (0.45)	0 (0.1)	-150 (0.9)	-300 (2)	0 (3)	3 (s)
63	-1800 (0.45)	0 (0.55)	-900 (0.9)	0 (0.1)	-1800 (1)	0 (4)	4 (s)
64	0 (0.55)	-3000 (0.45)	-1500 (0.9)	0 (0.1)	-3000 (2)	0 (4)	1
65	600 (0.3)	250 (0.7)	820 (0.25)	200 (0.75)	250 (2)	820 (3)	3 (s)
66	125 (0.7)	300 (0.3)	410 (0.25)	100 (0.75)	125 (1)	410 (3)	2
67	1800 (0.3)	750 (0.7)	600 (0.75)	2460 (0.25)	750 (2)	2460 (4)	1
68	1250 (0.7)	3000 (0.3)	1000 (0.75)	4100 (0.25)	1250 (1)	4100 (4)	4 (s)
69	300 (0.4)	200 (0.6)	360 (0.35)	175 (0.65)	200 (2)	360 (3)	3 (s)
70	100 (0.6)	150 (0.4)	180 (0.35)	875 (0.65)	100 (1)	180 (3)	2
71	900 (0.4)	600 (0.6)	525 (0.65)	1080 (0.35)	600 (2)	1080 (4)	1
72	1000 (0.6)	1500 (0.4)	875 (0.65)	1800 (0.35)	1000 (1)	1800 (4)	4 (s)
73	600 (0.001)	0 (0.999)	300 (0.002)	0 (0.998)	600 (1)	0 (4)	3
74	0 (0.999)	300 (0.001)	150 (0.002)	0 (0.998)	300 (2)	0 (4)	2 (s)
75	1800 (0.001)	(0.999) 0	0 (0.998)	900 (0.002)	1800 (1)	0 (3)	1 (s)
76	0 (0.999)	3000 (0.001)	0 (0.998)	1500 (0.002)	3000 (2)	0 (3)	4
77	400 (0.2)	0 (0.8)	300 (0.25)	0 (0.75)	400 (1)	0 (4)	3
78	0 (0.8)	200 (0.2)	150 (0.25)	0 (0.75)	200 (2)	0 (4)	2 (s)
79	1200 (0.2)	0 (0.8)	0 (0.75)	900 (0.25)	1200 (1)	0 (3)	1 (s)

..continued on next page.

Lotteries (continued)

Lottery	Option 1 Position 1	Option 1 Position 2	Option 2 Position 3	Option 2 Position 4	Salient Payoffs (Position) Option 1	Salient Payoffs (Position) Option 2	Position of Probe (s=salient)
80	0 (0.8)	2000 (0.2)	0 (0.75)	1500 (0.25)	2000 (2)	0 (3)	4
81	-400 (0.2)	0 (0.8)	0 (0.75)	-300 (0.25)	-400 (1)	0 (3)	1 (s)
82	0 (0.8)	-200 (0.2)	0 (0.75)	-150 (0.25)	-200 (2)	0 (3)	4
83	-1200 (0.2)	0 (0.8)	-900 (0.25)	0 (0.75)	-1200 (1)	0 (4)	3
84	0 (0.8)	-2000 (0.2)	-1500 (0.25)	0 (0.75)	-2000 (2)	0 (4)	2 (s)
85	0 (0.999)	-600 (0.001)	0 (0.998)	-300 (0.002)	-600 (2)	0 (3)	3 (s)
86	-300 (0.001)	0 (0.999)	0 (0.998)	-150 (0.002)	-300 (1)	0 (3)	2
87	0 (0.999)	-1800 (0.001)	-900 (0.002)	0 (0.998)	-1800 (2)	0 (4)	1
88	-3000 (0.001)	0 (0.999)	-1500 (0.002)	0 (0.998)	-3000 (1)	0 (4)	4 (s)
89	400 (0.5)	300 (0.5)	0 (0.3)	500 (0.7)	400 (1)	0 (3)	3 (s)
90	150 (0.5)	200 (0.5)	0 (0.3)	250 (0.7)	200 (2)	0 (3)	1
91	1200 (0.5)	900 (0.5)	1500 (0.7)	0 (0.3)	1200 (1)	0 (4)	2
92	1500 (0.5)	2000 (0.5)	2500 (0.7)	0 (0.3)	2000 (2)	0 (4)	4 (s)
93	100 (0.45)	400 (0.55)	440 (0.6)	0 (0.4)	400 (2)	0 (4)	4 (s)
94	200 (0.55)	50 (0.45)	220 (0.6)	0 (0.4)	200 (1)	0 (4)	2
95	300 (0.45)	1200 (0.55)	0 (0.4)	1320 (0.6)	1200 (2)	0 (3)	1
96	2000 (0.55)	500 (0.45)	0 (0.4)	2200 (0.6)	2000 (1)	0 (3)	3 (s)
97	400 (0.3)	1000 (0.7)	1025 (0.8)	0 (0.2)	1000 (2)	0 (4)	2 (s)
98	500 (0.7)	200 (0.3)	512.5 (0.8)	0 (0.2)	500 (1)	0 (4)	3
99	1200 (0.3)	3000 (0.7)	0 (0.2)	3075 (0.8)	3000 (2)	0 (3)	4
100	5000 (0.7)	2000 (0.3)	0 (0.2)	5125 (0.8)	5000 (1)	0 (3)	1 (s)
101	0 (0.3)	1200 (0.7)	1000 (0.6)	600 (0.4)	0 (1)	1000 (3)	3 (s)
102	600 (0.7)	0 (0.3)	500 (0.6)	300 (0.4)	0 (2)	500 (3)	1
103	0 (0.3)	3600 (0.7)	1800 (0.4)	3000 (0.6)	0 (1)	3000 (4)	2
104	6000 (0.7)	0 (0.3)	3000 (0.4)	5000 (0.6)	0 (2)	5000 (4)	4 (s)
105	750 (0.5)	0 (0.5)	200 (0.65)	700 (0.35)	0 (2)	700 (4)	2 (s)
106	0 (0.5)	375 (0.5)	100 (0.65)	350 (0.35)	0 (1)	350 (4)	3
107	2250 (0.5)	0 (0.5)	2100 (0.35)	600 (0.65)	0 (2)	2100 (3)	4
108	0 (0.5)	3750 (0.5)	3500 (0.35)	1000 (0.65)	0 (1)	3500 (3)	1 (s)
109	500 (0.6)	0 (0.4)	40 (0.55)	620 (0.45)	0 (2)	620 (4)	3
110	0 (0.4)	250 (0.6)	20 (0.55)	310 (0.45)	0 (1)	310 (4)	1 (s)
111	1500 (0.6)	0 (0.4)	1860 (0.45)	120 (0.55)	0 (2)	1860 (3)	2 (s)
112	0 (0.4)	2500 (0.6)	3100 (0.45)	200 (0.55)	0 (1)	3100 (3)	4
113	0 (0.5)	400 (0.5)	39 (0.65)	500 (0.65)	0 (1)	500 (4)	2
114	200 (0.5)	0 (0.5)	19.5 (0.65)	250 (0.65)	0 (2)	250 (4)	4 (s)
115	0 (0.5)	1200 (0.5)	1500 (0.35)	117 (0.65)	0 (1)	1500 (3)	3 (s)
116	2000 (0.5)	0 (0.5)	2500 (0.35)	195 (0.65)	0 (2)	2500 (3)	1 (s)
117	340 (0.5)	200 (0.5)	250 (0.6)	300 (0.4)	200 (2)	300 (4)	4 (s)
118	100 (0.5)	170 (0.5)	125 (0.6)	150 (0.4)	100 (1)	150 (4)	2
119	1020 (0.5)	600 (0.5)	900 (0.4)	750 (0.6)	600 (2)	900 (3)	1
120	1000 (0.5)	1700 (0.5)	1500 (0.4)	1250 (0.6)	1000 (1)	1500 (3)	3 (s)

B. Part III. Salience and Nudges: Experimental Design

B.1. Instructions

Welcome and thank you very much for participating in this experiment. This experiment is about decision-making. Please read the following instructions carefully. Everything that you need to know in order to participate in this experiment is explained below. If you have any difficulties in understanding these instructions please raise your hand and I will come to you. Please note that communication between participants is strictly prohibited during the experiment. Communication between participants will lead to the exclusion from the experiment. The experimental procedure will be as follows. You will receive 10 Euro. Please decide how much of the 10 Euro you would like to spend on climate protection. You can choose freely how much, if any, you contribute to climate protection (whole numbers between 0 and 10). Should you decide to contribute, we will realize your contribution to climate pro- tection by buying and retiring carbon emission licenses from the European Union Emissions Trading System (EU ETS) at the end of the experiment (please read the respective paragraph below for a description). By this, you have the possibility to make a real contribution to climate protection. The rest of the money is your private pay-out that you will receive in cash at the end of the experiment.

After making the decision you will be kindly asked to complete a short questionnaire. Please note that your decisions in this experiment are anonymous and will not be revealed at any stage to the other participants. (If relevant) a conrmation of the aggregated real payment to the climate protection fund will be sent to all participants at the end of the whole experiment.

Figure B.1.: Experimental screen for control.

Notes: The figure shows the decision screen shown to participants in the Control group. They could choose any integer between 0 and their endowment of 10 EUR. By clicking on the red OK button, subjects went to the next screen, providing them with information about the consequences of their decision, i.e. their payoff, their contribution, as well as kg of CO_2 offset.

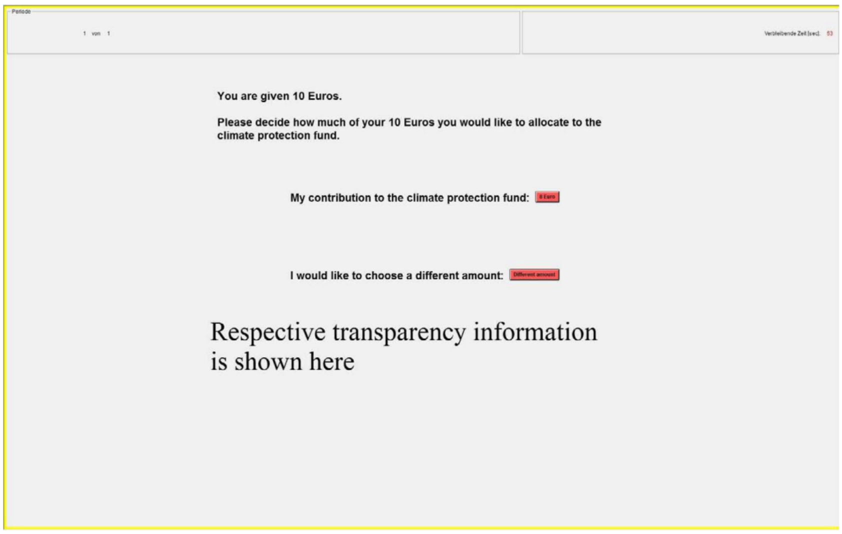

Figure B.2.: Experimental screen for Default and Transparancy.

Notes: The figure shows the decision screen shown to participants in the Default groups. They could choose to contribute the default value of 8 EUR by clocking on the respective red button, or they could click on the button below to choose any other amount. The transparency message was written where indicated in thefigure. The following screen provided subjects with information about the consequences of their decision, i.e. their payoff, their contribution, as well as kg of CO_2 offset.

B.2. The Climate Protection Fund

If a person wants to protect the climate, emitting climate gases such as CO_2 should be avoided. But it is possible to do even more: Individuals can buy and delete emission certificates from the EU Emission Trading System (ETS) through certified organizations and NGOs. By doing so, a private person reduces the amount of CO_2 which can be emitted by European industries, protects the environment and ensures that the development of climate-friendly technologies is accelerated. In this experiment, the participants' contributions to the climate protection fund will be used to buy real carbon dioxide (CO_2) emission licenses on the market of the European Union Emissions Trading Scheme (EU ETS) via the website "TheCompensators.org". It is one example of an NGO that allows ordinary people to directly participate in the EU ETS scheme, and where they can make decisions on CO_2 reductions.

Table B.1 shows how much kilograms of carbon you reduce with your payment, and how much money you receive for yourself. The far right row indicates the respective amount of reduced CO_2 relative to a Dutch citizens' average of 9163 kg of CO_2 emitted per year.

For example, with a payment of 3 Euro to retire carbon licenses, you retire 542 kg CO_2. This corresponds to approximately 6% of the average emissions per capita per year of a Dutch person. As a private pay-out you get 7 Euro. With a payment of 8 Euro to retire carbon licenses, you retire 1447 kg CO_2. This corresponds to approximately 16% of the average emissions per capita per year of a Dutch person. As a private pay-out you get 2 Euro.

Table B.1.: CO_2 reductions.

Payment to retire CO_2-allowances	Private payout in Euro	CO_2 abated in kg	Share of average emissions per year per person in %
0	10	0	0
1	9	181	2
2	8	362	4
3	7	542	6
4	6	723	8
5	5	904	10
6	4	1085	12
7	3	1266	14
8	2	1447	16
9	1	1627	18
10	0	1808	20

C. Part III. Salience and Nudges: Additional Statistical Analyses

C.1.

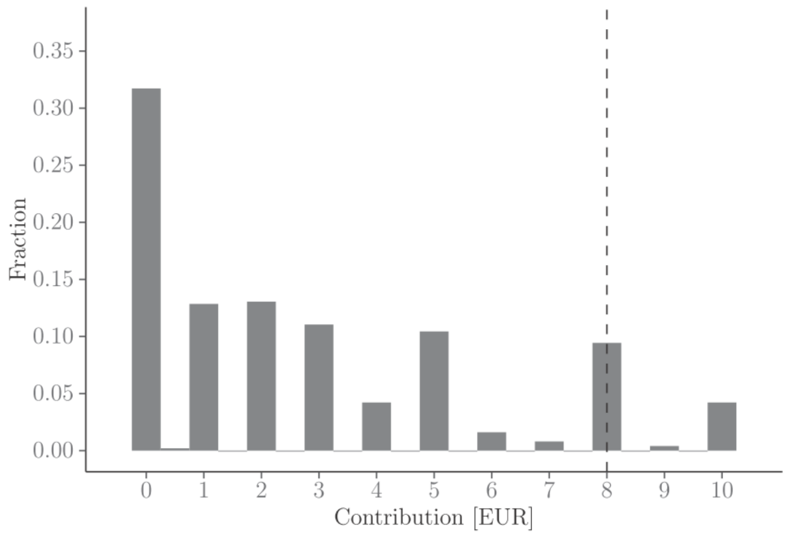

Figure C.1.: Distribution of contributions.

Notes: Shows the distribution of contribution amounts, more precisely the fraction of subjects contributing the respective amount. The dashed line indicates the default value.

Table C.1.: P-values for pairwise MW tests of contributions.

	Control	Default	Default + Info	Default + Purpose
Default	**0.001**			
Default + Info	<**0.001**	0.665		
Default + Purpose	**0.032**	0.843	0.591	
Default + Info + Purpose	**0.046**	0.785	0.606	0.91

Notes: P-values of pairwise Mann-Whitney tests for equality of distributions of contributions to the climate protection fund. Comparisons are indicated by the treatment names provided in the first column and first row, respectively. Significance levels: p < 0.05 in bold, p < 0.1 in cursive.

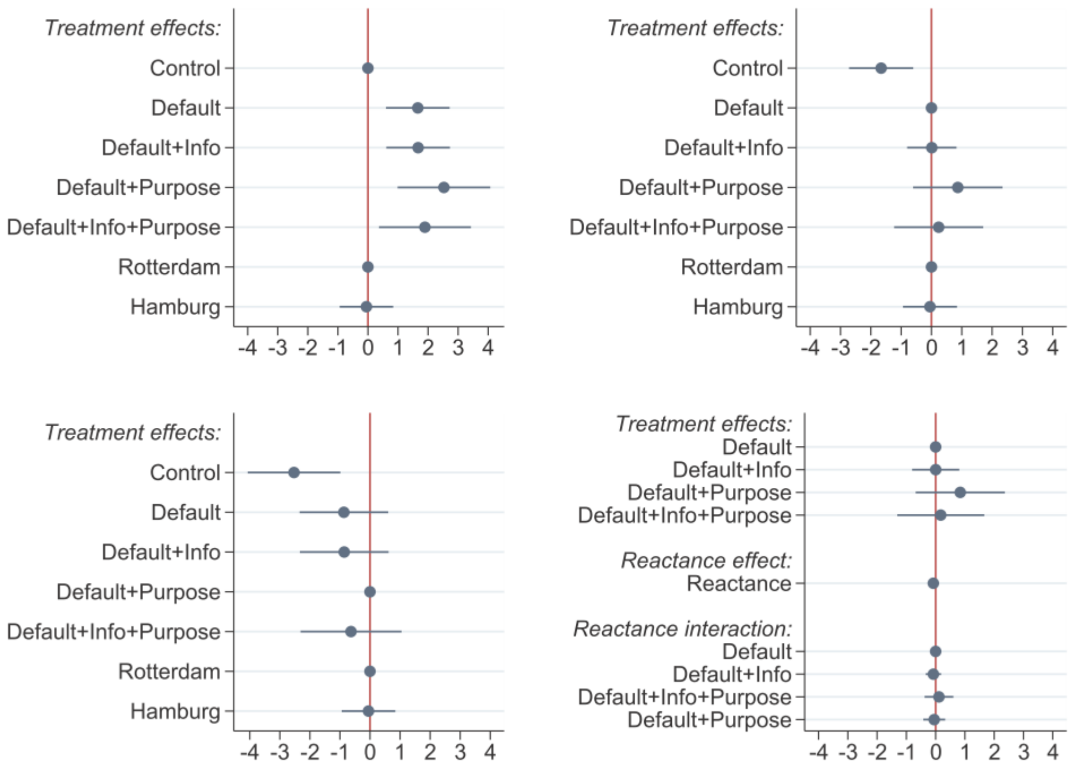

Figure C.2.: Default and transparency effects on contributions for different base-categories.

Notes: The figure graphically depicts results from some of the findings from the Tobit models. Dots with horizontal lines indicate point estimates with 98% conficence intervals. Dots on the zero line without confidence intervals denote the reference category. Models (3) and (8) in table 10.4 display the underlying regression results. The top left panel refers to finding F1, the top right panel to F2 and F3, the bottom left panel to F4, and the panel on the bottom right to F6. Covariates are not shown.

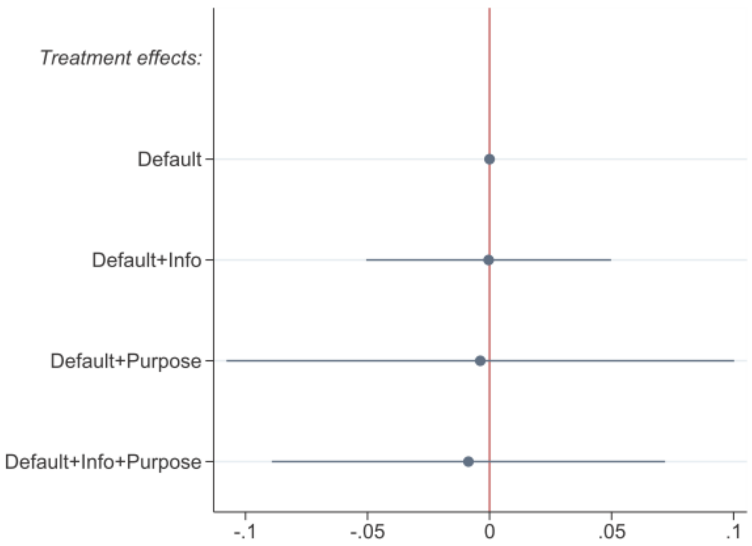

Figure C.3.: Default and transparency effects on perceived threat to freedom.

Notes: Dots with horizontal lines indicate point estimates with 95% confidence intervals from marginal effects of ordered logistic models. Dots on the zero line without confidence intervals denote the reference category. Model (4) in table 10.5 displays the underlying regression results (albeit not showing marginal effects). It refers to finding F5. Covariates are not shown.

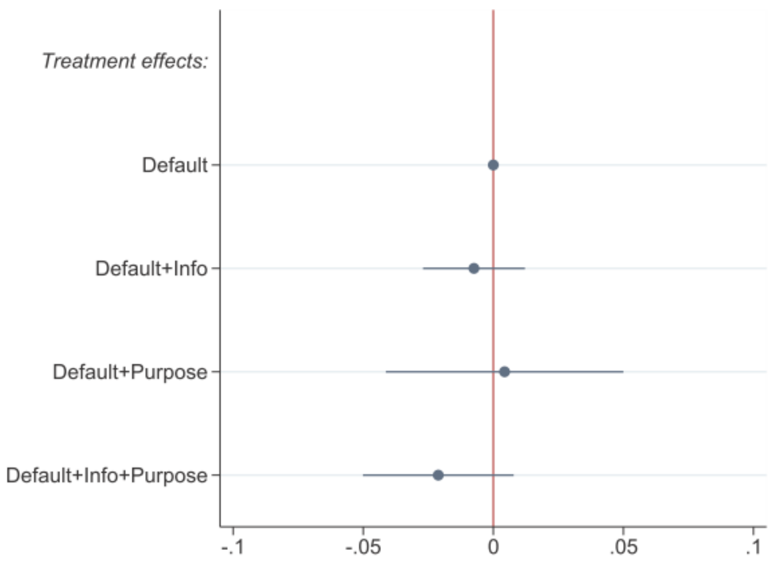

Figure C.4.: Default and transparency effects on anger.

Notes: Dots with horizontal lines indicate point estimates with 95% confidence intervals from marginal effects of ordered logistic models. Dots on the zero line without confidence intervals denote the reference category. Model (5) in table 10.5 displays the underlying regression results (albeit not showing marginal effects). It refers to finding F5. Covariates are not shown.

C.2. Comparing Subjects from Rotterdam and Hamburg

We conducted experimental sessions in two different cities. Findings from the first eleven experimental sessions relied on data solely from Rotterdam, while additional observations where gathered in Hamburg primarily in order to increase the reliability of the null result presented in F2 (and to a minor degree F3-F4 by increasing the n in the control group). The number of additional observations gathered in Hamburg relied on an a priori power analysis. Based on this analysis we conducted additional sessions to gather 284 additional observations for the Control, Default, and Default + Info groups. The experimental protocol in all sessions was identical.

Table C.2 shows summary statistics of the main outcome variables disaggregated by treatment and location of the experiment. Contribution distributions in the Control ($W = 795.5$, $p = 0.329$), Default ($W = 3053.5$, $p = 0.528$), and Default + Info ($W = 2119.5$, $p = 0.092$) groups do not differ by location. The same is true for the remaining outcome variables.

Table C.2.: Descriptive statistics of all outcome variables by experimental group and location.

Group	Location	Contribution Mean	SD	Contributed Mean	Picked default Mean	n
Control	R	1.67	2.68	46.67	0	45
Control	H	2	2.66	57.5	0	40
Default	R	3.24	3.21	73.91	19.57	46
Default	H	2.84	2.9	69.6	9.6	125
Default + Info	R	2.49	2.95	67.44	6.98	43
Default + Info	H	3.24	2.98	76.47	8.4	119
Default + Purpose	R	2.92	3.19	71.79	15.38	39
Default + Info + Purpose	R	2.85	2.95	65.85	17.07	41

Notes: The table reports summary statistics (means and standard deviations) of different outcome variables, as well as the number of subjects per experimental group. Outcome variables are: contributions to the climate protection fund, the percentage of subjects contributing a positive amount, as well as the percentage of subjects contributing the default value. Statistics are disaggregated by experimental group and location of the experiment.

Figure C.5 shows the mean contributions disaggregated by location and treatments, including bars indicating 95% confidence intervals. Mann-Whitney tests indicate that, while the default effect is significant in the Rotterdam sample ($W = 707.5$, $p = 0.007$), but not the Hamburg sample ($W = 2040.5$, $p = 0.074$), this is reversed with respect to the Default + Info effect, which is significant in Hamburg ($W = 1732.5$, 0.009), but not in Rotterdam ($W = 769.5$, $p = 0.084$). Differences between Default and Default + Info are insignificant in both samples (R: $W = 1113$, 0.302; H: $W = 6799$, $p = 0.24$).

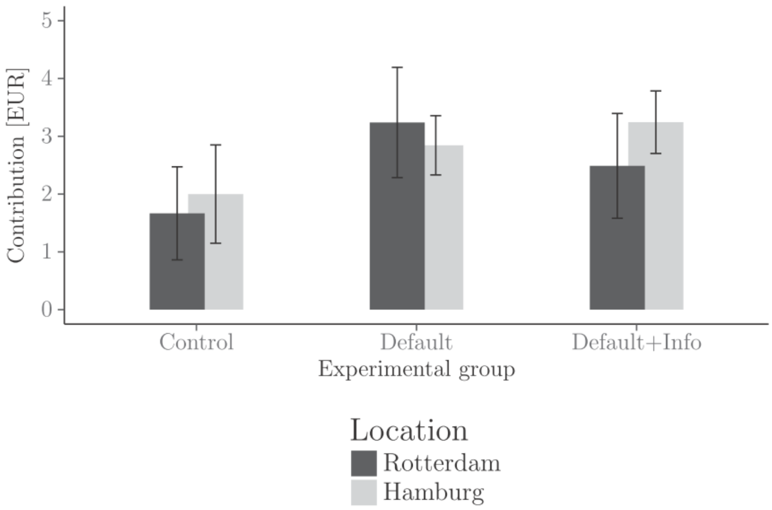

Figure C.5.: Mean contributions per experimental group and location.

Notes: Shows mean contributions by experimental group and location, including 95% confidence intervals.

Table C.3 shows summary statistics of the covariates included in the regression models disaggregated by treatment and location of the experiment. Aggregated over treatments, participants in Hamburg are on average older than participants in Rotterdam ($M = 24.94$ ($SD = 4.81$) vs. $M = 22.16$ ($SD = 3.45$), $t(494.84) = 7.517$, $p = 0.001$), less likely to be male ($M = 39.08$ vs. $M = 57.01$, $\chi^2(1) = 15.038, p < 0.001$), and also have a different distribution of study areas ($\chi^2(6) = 156.65, p < 0.001$). Additionally, participants in Hamburg are more likely than their Rotterdam colleagues to rate climate protection as (very) important ($\chi^2(1) = 37.06, p < 0.001$). They do not differ with respect to prior experience in experiments ($\chi^2(1) = 0.16, p < 0.69$) or their views regarding the effectiveness of the EU ETS ($\chi^2(1) = 0.002, p < 0.961$). Aggregated over location, subjects are not balanced among treatments according to some variables. Subjects' ratings of the importance of climate protection correlate with the treatment ($\chi^2(4) = 34.37, p < 0.001$). So does age ($H(4) = 16.294$, $p = 0.003$), and the distribution of study areas ($\chi^2(6) = 156.65, p < 0.001$).

Table C.3.: Descriptive statistics of covariates by experimental group and location.

		Age		Gender (Male)	Importance of CP	No exp. Experience	EU ETS not effective
Group	Location	Mean	SD	Mean	Mean	Mean	Mean
Control	R	21.8	3.08	60	57.78	31.11	57.78
Control	H	25.95	5.7	35	97.5	15	62.5
Default	R	22.02	2.79	60.87	78.26	30.43	60.87
Default	H	24.95	4.48	36.8	84	29.6	60
Default + Info	R	22.07	2.96	51.16	79.07	20.93	53.49
Default + Info	H	24.59	4.81	42.86	91.6	27.73	57.98
Default + Purpose	R	22.28	4.65	53.85	51.28	20.51	64.1
Default + Info + Purpose	R	22.68	3.72	58.54	63.41	19.51	58.54

Notes: The table reports summary statistics (means and standard deviations) of different covariates per experimental group. Covariates are: age of participants, percentage of males, percentage of subjects perceiving climate protection as (very) important, percentage of subjects without prior experience with experiments, as well as the percentage of subjects judging license retirement as an ineffective mean for climate protection. Statistics are disaggregated by experimental group and location of the experiment.

Figure C.6 shows standardized effect sizes and 95% confidence intervals of the relevant pairwise comparisons for which we gathered additional data. While the effect size of the default effect (Con vs. Def) included zero in the Hamburg sample, it does not include zero in the Rotterdam- and the aggregate sample. The default + info effect size (Con vs. Def + Inf) is different from zero in the Hamburg and aggregated sample, but not in the Rotterdam sample. Although the standardized effect sizes for the Def vs. Def + Inf comparison is opposite between Hamburg and Rotterdam, neither those nor the aggregated sample exclude an effect size of zero.

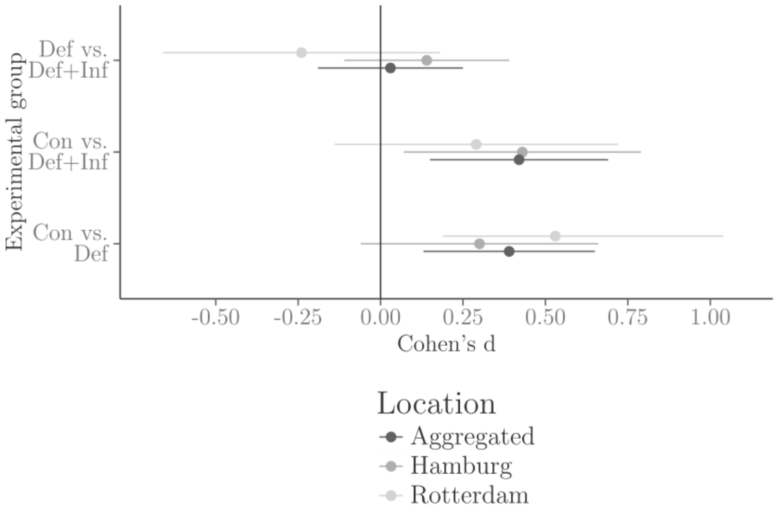

Figure C.6.: Effect sizes by location and for aggregated data.

Notes: Shows Cohen's d for each pairwise comparison for which additional data in Hamburg was gathered, including 95% confidence intervals.

Figure C.7 shows the regression coefficients and 95% confidence intervals from Tobit model (3). These are qualitatively similar to the respective effect sizes, with the exception that the standardized effect size for the Con vs. Def comparison in Hamburg includes zero, whereas this is not the case for the respective regression coefficient.

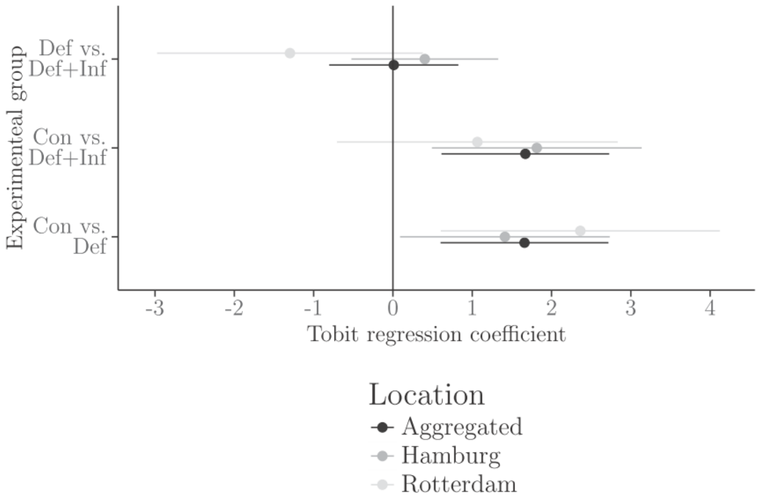

Figure C.7.: Coefficients from tobit model by location and for aggregated data.

Notes: Shows estimated coefficients from Tobit model (3) for effect for which additional data in Hamburg was gathered, including 95% confidence intervals.

C.3. Questionnaires

C.3.1. Questionnaire on covariates

What is you gender? O Male O Female
What is your age?
Have you participated in other experiments before today? O Yes O No
How important is climate protection for you? Please circle the most suitable answer.
O Not important at all O Not important O Indifferent O Important O Very important
Do you think that buying real carbon dioxide (CO_2) emissions licenses on the market of the European Union Emissions Trading Scheme (EU ETS) is an effective method to contribute to climate protection? O Yes O No

C.3.2. Questionnaire on state reactance

Please indicate to what extent do you agree with the following statements on a 5-point response scale that ranges from the statement ”strongly disagree” to the statement ”strongly agree”. (Perceived threat to freedom)

- The default value threatened my freedom to choose.
- The default value tried to make a decision for me.
- The default value tried to manipulate me.
- The default value tried to pressure me.
  Please indicate to what extent do you agree with the following statements on a 5-point response scale that ranges from the statement ”Not at all” to the statement ”Very”. (anger)

- Please indicate how irritated you were with regard to the given default value.
- Please indicate how angry you were with regard to the given default value.
- Please indicate how annoyed you were with regard to the given default value.
- Please indicate how aggravated you were with regard to the given default value.

C.3.3. Questionnaire on trait reactance

Please indicate to what extent do you agree with the following statements on a p-point response scale that ranges from the statement ”strongly disagree” to the statement ”strongly agree”.

- Regulations trigger a sense of resistance in me.
- I find contradicting others stimulating.
- When something is prohibited, I usually think, ”that’s exactly what I am going to do”.
- The thought of being dependent on others aggravates me.
- I consider advice from others to be an intrusion.
- I become frustrated when I am unable to make free and independent decisions.
- It irritates me when someone points out things, which are obvious to me.
- I become angry when my freedom of choice is restricted.
- Advice and recommendations usually induce me to do just the opposite.
- I am content only when I am acting on my own free will.
- I resist the attempts of others to influence me.
- It makes me angry when another person is held up as a role model for me to follow.
- When someone forces me to do something, I feel like doing the opposite.
- It disappoints me to see others submitting to standards and rules.

Bibliography

Albert, D., & Steinberg, L. (2011). Peer influences on adolescent risk behavior. In M. T. Bardo, D. H. Fishbein, & R. Milich (Eds.), *Inhibitory control and drug abuse prevention* (pp. 211–226).

Alemanno, A., & Sibony, A.-L. (2015). *Nudge and the law: A european perspective*. Bloomsbury Publishing.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole americaine. *Econometrica*, *21*(4), 503–546.

Allcott, H., & Mullainathan, S. (2010). Behavior and energy policy. *Science*, *327*(5970), 1204–1205.

Allport, F. H. (1924). *Social psychology*. Boston, Mass.: Mifflin [u.a.]

Altmann, S., Grunewald, A., & Radbruch, J. (2018). Passive choices and cognitive spillovers.

Arad, A., & Rubinstein, A. (2017). The people's perspective on libertarian-paternalistic policies. *Tel Aviv*.

Araña, J. E., & León, C. J. (2013). Can defaults save the climate? evidence from a field experiment on carbon offsetting programs. *Environmental and Resource Economics*, *54*(4), 613–626.

Armel, K. C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making*, *3*(5), 396–403.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Baron, R. S. (1986). Distraction-conflict theory: Progress and problems. In *Advances in experimental social psychology* (Vol. 19, pp. 1–40).

Bartling, B., Fehr, E., & Herz, H. (2014). The intrinsic value of decision rights. *Econometrica*, *82*(6), 2005–2039.

Baxter, M. G., & Murray, E. A. (2002). The amygdala and reward. *Nature reviews. Neuroscience*, *3*(7), 563–573.

Bechara, A. (2004). The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and cognition*, *55*(1), 30–40.

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, *52*(2), 336–372.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex*, *10*(3), 295–307.

Ben Zur, H., & Breznitz, S. J. (1981). The effect of time pressure on risky choice behavior. *Acta psychologica*, *47*(2), 89–104.

Bergman, O., Ellingsen, T., Johannesson, M., & Svensson, C. (2010). Anchoring and cognitive ability. *Economics Letters*, *107*(1), 66–68.

Bhargava, S., Loewenstein, G., & Sydnor, J. (2017). Choose to lose: Health plan choices from a menu with dominated option*. *The Quarterly Journal of Economics*, *132*(3), 1319–1372.

Bigdely-Shamlo, N., Kreutz-Delgado, K., Kothe, C., & Makeig, S. (2013). Eyecatch: Datamining over half a million eeg independent components to construct a fully-automated eye-component detector. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, *2013*, 5845–5848.

Bock, O., Baetge, I., & Nicklisch, A. (2014). Hroot: Hamburg registration and organization online tool. *European Economic Review*, *71*, 117–120.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012a). Salience theory of choice under risk. *The Quarterly Journal of Economics*, *127*(3), 1243–1285.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012b). Online appendix of salience theory of choice under risk.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012c). Salience in experimental tests of the endowment effect. *American Economic Review*, *102*(3), 47–52.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2015). Salience theory of judicial decisions. *The Journal of Legal Studies*, *44*(S1), S7–S33.

Borghans, L., & Golsteyn, B. H. H. (2015). Susceptibility to default training options across the population. *Journal of Economic Behavior & Organization*, *117*, 369–379.

Bovens, L. (2009). The ethics of nudge. In *Preference change* (pp. 207–219). Springer.

Bradley, B. P., Mogg, K., & Millar, N. H. (2000). Covert and overt orienting of attention to emotional faces in anxiety. *Cognition and Emotion*, *14*(6), 789–808.

Brehm, J. W. (1966). A theory of psychological reactance.

Brehm, S. S., & Brehm, J. W. (2013). *Psychological reactance: A theory of freedom and control.* Academic Press.

Breiter, H. C., & Rosen, B. R. (1999). Functional magnetic resonance imaging of brain reward circuitry in the human. *Annals of the New York Academy of Sciences*, *877*(1 Advancing Fro), 523–547.

Brown, C. L., & Krishna, A. (2004). The skeptical shopper: A metacognitive account for the effects of default options on choice. *Journal of consumer research*, *31*(3), 529–539.

Brown, J. R., Farrell, A. M., & Weisbenner, S. J. (2012). The downside of defaults. National Bureau of Economic Research.

Buckert, M., Schwieren, C., Kudielka, B. M., & Fiebach, C. J. (2014). Acute stress affects risk taking but not ambiguity aversion. *Frontiers in neuroscience*, *8*, 82.

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(19), 7745–7750.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432.

Caplin, A., & Martin, D. (2012). *Defaults and attention: The drop out effect.*

Cappelletti, D., Mittone, L., & Ploner, M. (2014). Are default contributions sticky? an experimental analysis of defaults in public goods provision. *Journal of Economic Behavior & Organization*, *108*, 331–342.

Cardinal, R. N., Parkinson, J. A., Hall, J., & Everitt, B. J. (2002). Emotion and motivation: The role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience & Biobehavioral Reviews*, *26*(3), 321–352.

Carlin, B. I., Gervais, S., & Manso, G. (2013). Libertarian paternalism, information production, and financial decision making: Table 1. *Review of Financial Studies*, *26*(9), 2204–2228.

Carroll, G. D., Choi, J. J., Laibson, D., Madrian, B. C., & Metrick, A. (2009). Optimal defaults and active decisions *. *The Quarterly Journal of Economics*, *124*(4), 1639–1674.

Carter, C. S., Botvinick, M. M., & Cohen, J. D. (1999). The contribution of the anterior cingulate cortex to executive processes in cognition. *Reviews in the neurosciences*, *10*(1), 56.

Carter, R. M., Bowling, D. L., Reeck, C., & Huettel, S. A. (2012). A distinct role of the temporal-parietal junction in predicting socially guided decisions. *Science (New York, N.Y.)* *337*(6090), 109–111.

Chand, G. B., & Dhamala, M. (2016). Interactions among the brain default-mode, salience, and central-executive networks during perceptual decision-making of moving dots. *Brain connectivity*, *6*(3), 249–254.

Charles, S. T., Reynolds, C. A., & Gatz, M. (2001). Age-related differences and change in positive and negative affect over 23 years. *Journal of personality and social psychology*, *80*(1), 136.

Chartrand, T. L., Dalton, A. N., & Fitzsimons, G. J. (2007). Nonconscious relationship reactance: When significant others prime opposing goals. *Journal of Experimental Social Psychology*, *43*(5), 719–726.

Chein, J., Albert, D., O'Brien, L., Uckert, K., & Steinberg, L. (2011). Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry. *Developmental Science*, *14*(2), F1–F10.

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *29*(39), 12315–12320.

Chung, D., Christopoulos, G. I., King-Casas, B., Ball, S. B., & Chiu, P. H. (2015). Social signals of safety and risk confer utility and have asymmetric effects on observers' choices. *Nature neuroscience*, *18*(6), 912–916.

Cisler, J. M., & Koster, E. H. W. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clinical psychology review*, *30*(2), 203–216.

Cooper, J. C., & Knutson, B. (2008). Valence and salience contribute to nucleus accumbens activation. *NeuroImage*, *39*(1), 538–547.

Cooper, R. M., & Langton, S. R. H. (2006). Attentional bias to angry faces using the dot-probe task? it depends when you look for it. *Behaviour research and therapy*, *44*(9), 1321–1329.

Critchley, H. D., & Harrison, N. A. (2013). Visceral influences on brain and behavior. *Neuron*, *77*(4), 624–638.

Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron*, *29*(2), 537–545.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*(2), 448–474.

Cummings, J. L. (1995). Anatomic and behavioral aspects of frontal-subcortical circuits. *Annals of the New York Academy of Sciences*, *769*(1 Structure and), 1–14.

Damasio, A. R. (2003). *Looking for spinoza: Joy, sorrow, and the feeling brain*. Orlando, Fla.: Harcourt.

Dambacher, M., & Hübner, R. (2015). Time pressure affects the efficiency of perceptual processing in decisions under conflict. *Psychological Research*, *79*(1), 83–94.

Danchin, E., Giraldeau, L.-A., Valone, T. J., & Wagner, R. H. (2004). Public information: From nosy neighbors to cultural evolution. *Science (New York, N.Y.) 305*(5683), 487–491.

de Haan, T., & Linde, J. (2018). 'good nudge lullaby': Choice architecture and default bias reinforcement. *The Economic Journal*, *128*(610), 1180–1206.

de Kloet, E. R. (2004). Hormones and the stressed brain. *Annals of the New York Academy of Sciences*, *1018*, 1–15.

de Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science (New York, N.Y.) 313*(5787), 684–687.

de Paola, M., & Gioia, F. (2016). Who performs better under time pressure? results from a field experiment. *Journal of Economic Psychology*, *53*, 37–53.

Deck, C., & Jahedi, S. (2015). The effect of cognitive load on economic decision making: A survey and new experiments. *European Economic Review*, *78*, 97–119.

Defoe, I. N., Dubas, J. S., Figner, B., & van Aken, M. A. G. (2015). A meta-analysis on age differences in risky decision making: Adolescents versus children and adults. *Psychological Bulletin*, *141*(1), 48–84.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, *284*(5416), 970–974.

Delorme, A., & Makeig, S. (2004). Eeglab: An open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9–21.

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in eeg data using higher-order statistics and independent component analysis. *NeuroImage*, *34*(4), 1443–1449.

Dhingra, N., Gorn, Z., Kener, A., & Dana, J. (2012). The default pull: An experimental demonstration of subtle default effects on preferences. *Judgment and Decision Making*, *7*(1), 69.

Dias, R., Robbins, T. W., & Roberts, A. C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the wisconsin card sort test: Restriction to novel situations and independence from "on-line" processing. *Journal of Neuroscience*, *17*(23), 9285–9297.

Diederich, J., & Goeschl, T. (2014). Willingness to pay for voluntary climate action and its determinants: Field-experimental evidence. *Environmental and Resource Economics*, *57*(3), 405–429.

Dillard, J. P., & Shen, L. (2005). On the nature of reactance and its role in persuasive health communication. *Communication Monographs*, *72*(2), 144–168.

Dinner, I., Johnson, E. J., Goldstein, D. G., & Liu, K. (2011). Partitioning default effects: Why people choose not to choose. *Journal of Experimental Psychology: Applied*, *17*(4), 332.

Elliott, R., Rees, G., & Dolan, R. (1999). Ventromedial prefrontal cortex mediates guessing. *Neuropsychologia*, *37*(4), 403–411.

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, *14*(4), 583–610.

Ernst, M., & Paulus, M. P. (2005). Neurobiology of decision making: A selective review from a neurocognitive and clinical perspective. *Biological Psychiatry*, *58*(8), 597–604.

Ernst, M., Nelson, E. E., McClure, E. B., Monk, C. S., Munson, S., Eshel, N., ... Pine, D. S. (2004). Choice selection and reward anticipation: An fmri study. *Neuropsychologia*, *42*(12), 1585–1597.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science*, *8*(3), 223–241.

Fehr, E., Herz, H., & Wilkening, T. (2013). The lure of authority: Motivation and incentive effects of power. *American Economic Review*, *103*(4), 1325–1359.

Felsen, G., Castelo, N., & Reiner, P. B. (2013). Decisional enhancement and autonomy: Public attitudes towards overt and covert nudges. *Judgment and Decision Making*, *8*(3).

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140.

Fiedler, S., & Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in psychology*, *3*, 335.

Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.

Franco-Watkins, A. M., Rickard, T. C., & Pashler, H. (2010). Taxing executive processes does not necessarily increase impulsive decision making. *Experimental psychology*, *57*(3), 193–201.

Frewen, P. A., Dozois, D. J. A., Joanisse, M. F., & Neufeld, R. W. J. (2008). Selective attention to threat versus reward: Meta-analysis and neural-network modeling of the dot-probe task. *Clinical psychology review*, *28*(2), 307–337.

Frith, C. D., & Singer, T. (2008). The role of social cognition in decision making. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *363*(1511), 3875–3886.

Gabriel, M., Burhans, L., & Kashef, A. (2003). Consideration of a unified model of amygdalar associative functions. *Annals of the New York Academy of Sciences*, *985*(1), 206–217.

Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: An experimental study. *Developmental Psychology*, *41*(4), 625–635.

Gehring, W. J., & Knight, R. T. (2000). Prefrontal-cingulate interactions in action monitoring. *Nature neuroscience*, *3*(5), 516–520.

Goel, V., Gold, B., Kapur, S., & Houle, S. (1998). Neuroanatomical correlates of human reasoning. *Journal of Cognitive Neuroscience*, *10*(3), 293–302.

Goswami, I., & Urminsky, O. (2016). When should the ask be a nudge? the effect of default amounts on charitable donations. *Journal of Marketing Research*, *53*(5), 829–846.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with orsee. *Journal of the Economic Science Association*, *1*(1), 114–125.

Grimshaw, G. M., Foster, J. J., & Corballis, P. M. (2014). Frontal and parietal eeg asymmetries interact to predict attentional bias to threat. *Brain and cognition*, *90*, 76–86.

Haggag, K., & Paci, G. (2014). Default tips. *American Economic Journal: Applied Economics*, *6*(3), 1–19.

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

Hansen, P. G., & Jespersen, A. M. (2013). Nudge and the manipulation of choice. *European Journal of Risk Regulation*, *4*(01), 3–28.

Harkins, S. G. (2006). Mere effort as the mediator of the evaluation-performance relationship. *Journal of personality and social psychology*, *91*(3), 436–455.

Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic Literature*, *42*(4), 1009–1055.

Hausman, D. M., & Welch, B. (2010). Debate: To nudge or not to nudge*. *Journal of Political Philosophy*, *18*(1), 123–136.

Hedlin, S., & Sunstein, C. R. (2016). Does active choosing promote green energy use: Experimental evidence. *Ecology Law Quarterly*, *43*, 107.

Heinze, H. J., Luck, S. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. i. evidence for early selection. *Electroencephalography and Clinical Neurophysiology*, *75*(6), 511–527.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *353*(1373), 1257–1270.

Hong, S.-M., & Faedda, S. (1996). Refinement of the hong psychological reactance scale. *Educational and Psychological Measurement*, *56*(1), 173–182.

House of Lords Report. (2011). Behaviour change.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision research*, *40*(10), 1489–1506.

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

Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., & Koeppe, R. A. (1997). Verbal working memory load affects regional brain activation as measured by pet. *Journal of cognitive neuroscience*, *9*(4), 462–475.

Jung, T.-P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, *111*(10), 1745–1758.

Kahneman, D. (2011). *Thinking, fast and slow.* Farrar, Straus and Giroux New York.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *Volume 47, Issue 2*, 263–292.

Kahnt, T., & Tobler, P. N. (2013). Salience signals in the right temporoparietal junction facilitate value-based decisions. *Journal of Neuroscience*, *33*(3), 863–869.

Kahnt, T., Park, S. Q., Haynes, J.-D., & Tobler, P. N. (2014). Disentangling neural representations of value and salience in the human brain. *Proceedings of the National Academy of Sciences*, *111*(13), 5000–5005.

Kameda, T., Ohtsubo, Y., & Takezawa, M. (1997). Centrality in sociocognitive networks and social influence: An illustration in a group decision-making context. *Journal of personality and social psychology*, *73*(2), 296–309.

Kappenman, E. S., Farrens, J. L., Luck, S. J., & Proudfit, G. H. (2014). Behavioral and erp measures of attentional bias to threat in the dot-probe task: Poor reliability and lack of correlation with anxiety. *Frontiers in psychology*, *5*, 1368.

Karlan, D., McConnell, M., Mullainathan, S., & Zinman, J. (2016). Getting to the top of mind: How reminders increase saving. *Management Science*, *62*(12), 3393–3411.

Kling, J. R., Mullainathan, S., Shafir, E., Vermeulen, L. C., & Wrobel, M. V. (2012). Comparison friction: Experimental evidence from medicare drug plans. *The Quarterly Journal of Economics*, *127*(1), 199–235.

Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fmri. *Neuroreport*, *12*(17), 3683–3687.

Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M., & Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: Characterization with rapid event-related fmri. *NeuroImage*, *18*(2), 263–272.

Kocher, M. G., & Sutter, M. (2006). Time is money—time pressure, incentives, and the quality of decision-making. *Journal of Economic Behavior & Organization*, *61*(3), 375–392.

Kocher, M. G., Pahlke, J., & Trautmann, S. T. (2013). Tempus fugit: Time pressure in risky decisions. *Management Science*, *59*(10), 2380–2391.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience*, *13*(10), 1292–1298.

Kroese, F. M., Marchiori, D. R., & de Ridder, D. T. D. (2016). Nudging healthy food choices: A field experiment at the train station. *Journal of public health (Oxford, England)*, *38*(2), e133–7.

Kuang, X., Weber, R. A., & Dana, J. (2007). How effective is advice from interested parties? *Journal of Economic Behavior & Organization*, *62*(4), 591–604.

Labudda, K., Woermann, F. G., Mertens, M., Pohlmann-Eden, B., Markowitsch, H. J., & Brand, M. (2008). Neural correlates of decision making with explicit information about probabilities and incentives in elderly healthy subjects. *Experimental Brain Research*, *187*(4), 641–650.

Laury, S., & Holt, C. A. (2005). Further reflections on prospect theory. *SSRN Electronic Journal.*

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual review of neuroscience*, *23*, 155–184.

Lee, D. (2008). Game theory and neural basis of social decision making. *Nature neuroscience*, *11*(4), 404–409.

Litvak, V., Mattout, J., Kiebel, S., Phillips, C., Henson, R., Kilner, J., … Friston, K. (2011). Eeg and meg data analysis in spm8. *Computational intelligence and neuroscience*, *2011*, 852961.

Liu, Y., Zhang, D., & Luo, Y. (2015). How disgust facilitates avoidance: An erp study on attention modulation by threats. *Social cognitive and affective neuroscience*, *10*(4), 598–604.

Loewenstein, G. F., & O'Donoghue, T. (2004). Animal spirits: Affective and deliberative processes in economic behavior. *SSRN Electronic Journal*.

Loewenstein, G., Rick, S., & Cohen, J. D. (2008). Neuroeconomics. *Annual review of psychology*, *59*, 647–672.

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

Lourenço, J. S., Ciriolo, E., Rafael Almeida, S., & Troussard, X. (2016). *Behavioural insights applied to policy: European report 2016*. EUR. Luxembourg: Publications Office.

Luck, S. J., Heinze, H. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. ii. functional dissociation of p1 and n1 components. *Electroencephalography and Clinical Neurophysiology*, *75*(6), 528–542.

Luck, S. J., Fan, S., & Hillyard, S. A. (1993). Attention-related modulation of sensory-evoked brain activity in a visual search task. *Journal of cognitive neuroscience*, *5*(2), 188–195.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cognitive neuroscience series. Cambridge Mass.: MIT Press.

Luck, S. J., & Kappenman, E. S. (2012a).

Luck, S. J., & Kappenman, E. S. (2012b). Erp components and selective attention. *The Oxford handbook of event-related potential components*, 295–327.

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of abnormal psychology*, *95*(1), 15.

Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *The Quarterly Journal of Economics*, *116*(4), 1149–1187.

Mangun, G. R., Hopfinger, J. B., Kussmaul, C. L., Fletcher, E. M., & Heinze, H.-J. (1997). Covariations in erp and pet measures of spatial selective attention in human extrastriate visual cortex. *Human Brain Mapping*, *5*(4), 273–279.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of eeg- and meg-data. *Journal of neuroscience methods*, *164*(1), 177–190.

Mathalon, D. H., Whitfield, S. L., & Ford, J. M. (2003). Anatomy of an error: Erp and fmri. *Biological psychology*, *64*(1-2), 119–141.

McClure, S. M., Berns, G. S., & Montague, P. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, *38*(2), 339–346.

McKenzie, C. R. M., Liersch, M. J., & Finkelstein, S. R. (2006). Recommendations implicit in policy defaults. *Psychological Science*, *17*(5), 414–420.

Mitterschiffthaler, M. T., Kumari, V., Malhi, G. S., Brown, R. G., Giampietro, V. P., Brammer, M. J., ... Andrew, C. (2003). Neural response to pleasant stimuli in anhedonia: An fmri study. *Neuroreport*, *14*(2), 177–182.

Mogg, K., Philippot, P., & Bradley, B. P. (2004). Selective attention to angry faces in clinical social phobia. *Journal of abnormal psychology*, *113*(1), 160–165.

Monfardini, E., Redouté, J., Hadj-Bouziane, F., Hynaux, C., Fradin, J., Huguet, P., ... Meunier, M. (2015). Others' sheer presence boosts brain activity in the attention (but not the motivation) network. *Cerebral cortex (New York, N.Y. : 1991)*.

Montague, P., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, *36*(2), 265–284.

Muller, D., Atzeni, T., & Butera, F. (2004). Coaction and upward social comparison reduce the illusory conjunction effect: Support for distraction–conflict theory. *Journal of Experimental Social Psychology*, *40*(5), 659–665.

Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *Journal of experimental psychology: Human perception and performance*, *15*(2), 315.

Murooka, T., & Schwarz, M. A. (2018). The timing of choice-enhancing policies. *Journal of Public Economics*, *157*, 27–40.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329–337.

Olsson, A., & Ochsner, K. N. (2008). The role of social cognition in emotion. *Trends in Cognitive Sciences*, *12*(2), 65–71.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). Fieldtrip: Open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Computational intelligence and neuroscience*, *2011*, 156869.

Ossewaarde, L., Qin, S., van Marle, H. J. F., van Wingen, G. A., Fernández, G., & Hermans, E. J. (2011). Stress-induced reduction in reward-related prefrontal cortex function. *NeuroImage*, *55*(1), 345–352.

Owens, D., Grossman, Z., & Fackler, R. (2014). The control premium: A preference for payoff autonomy. *American Economic Journal: Microeconomics*, *6*(4), 138–161.

Pachur, T., Hertwig, R., Gigerenzer, G., & Brandstätter, E. (2013). Testing process predictions of models of risky choice: A quantitative model comparison approach. *Frontiers in psychology*, *4*, 646.

Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature neuroscience*, *5*(2), 97–98.

Perino, G. (2015). Climate campaigns, cap and trade, and carbon leakage: Why trying to reduce your carbon footprint can harm the climate. *Journal of the Association of Environmental and Resource Economists*, *2*(3), 469–495.

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception i: The neural basis of normal emotion perception. *Biological Psychiatry*, *54*(5), 504–514.

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, *400*(6741), 233–238.

Polich, J. (2007). Updating p300: An integrative theory of p3a and p3b. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–2148.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*(2), 160–174.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature reviews. Neuroscience*, *9*(7), 545–556.

Rebonato, R. (2014). A critical assessment of libertarian paternalism. *Journal of Consumer Policy*, *37*(3), 357–396.

Reisch, L. A., & Sunstein, C. R. (2016). Do europeans like nudges? *SSRN Electronic Journal*.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., … Laland, K. N. (2010). Why copy others? insights from the social learning strategies tournament. *Science (New York, N.Y.) 328*(5975), 208–213.

Rieskamp, J., Busemeyer, J. R., & Mellers, B. A. (2006). Extending the bounds of rationality: Evidence and theories of preferential choice. *Journal of Economic Literature*, *44*(3), 631–661.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionst model of decision making. *Psychological Review*, *108*(2), 370.

Salamone, J. D., & Correa, M. (2002). Motivational views of reinforcement: Implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behavioural Brain Research*, *137*(1-2), 3–25.

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, *1*(1), 7–59.

Sanders, G. S., & Baron, R. S. (1975). The motivating effects of distraction on task performance. *Journal of personality and social psychology*, *32*(6), 956–963.

Sanfey, A. G., Hastie, R., Colvin, M. K., & Grafman, J. (2003). Phineas gauged: Decision-making and the human prefrontal cortex. *Neuropsychologia*, *41*(9), 1218–1229.

Schilbach, F., Schofield, H., & Mullainathan, S. (2016). The psychological lives of the poor. *The American Economic Review*, *106*(5), 435–440.

Schmukle, S. C. (2005). Unreliability of the dot probe task. *European Journal of Personality*, *19*(7), 595–605.

Schoenbaum, G., & Setlow, B. (2003). Lesions of nucleus accumbens disrupt learning about aversive outcomes. *Journal of Neuroscience*, *23*(30), 9833–9841.

Schwabe, L., Tegenthoff, M., Höffken, O., & Wolf, O. T. (2012). Simultaneous glucocorticoid and noradrenergic activity disrupts the neural basis of goal-directed action in the human brain. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *32*(30), 10146–10155.

Schwager, S., & Rothermund, K. (2013). Motivation and affective processing biases in risky decision making: A counter-regulation account. *Journal of Economic Psychology*, *38*, 111–126.

Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area lip) of the rhesus monkey. *Journal of neurophysiology*, *86*(4), 1916–1936.

Shafir, E. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory & Cognition*, *21*(4), 546–556.

Sharma, D., Booth, R., Brown, R., & Huguet, P. (2010). Exploring the temporal dynamics of social facilitation in the stroop task. *Psychonomic Bulletin & Review*, *17*(1), 52–58.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature neuroscience*, *6*(12), 1317–1322.

Singer, T., & Tusche, A. (2014). Understanding others. In *Neuroeconomics* (pp. 513–532).

Sjöberg, A., Sjöberg, S., & Forssén, K. (2006). Predicting job performance. *Assessio International, Stockholm.*

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*(1), 3.

Somerville, L. H., Jones, R. M., & Casey, B. J. (2010). A time of change: Behavioral and neural correlates of adolescent sensitivity to appetitive and aversive environmental cues. *Brain and cognition*, *72*(1), 124–133.

Sprenger, A. M., Dougherty, M., Atkins, S. M., Franco-Watkins, A. M., Thomas, R., Lange, N., & Abbs, B. (2011). Implications of cognitive load for hypothesis generation and probability judgment. *Frontiers in psychology*, *2*, 129.

Starcke, K., & Brand, M. (2012). Decision making under stress: A selective review. *Neuroscience and biobehavioral reviews*, *36*(4), 1228–1248.

Stasser, G., Stewart, D. D., & Wittenbaum, G. M. (1995). Expert roles and information exchange during discussion: The importance of knowing who knows what. *Journal of Experimental Social Psychology*, *31*(3), 244–265.

Staugaard, S. R. (2009). Reliability of two versions of the dot-probe task using photographic faces. *Psychology Science Quarterly*, *51*(3), 339–350.

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

Stevens, S., Rist, F., & Gerlach, A. L. (2009). Influence of alcohol on the processing of emotional facial expressions in individuals with social phobia. *The British journal of clinical psychology*, *48*(Pt 2), 125–140.

Studer, B., Cen, D., & Walsh, V. (2014). The angular gyrus and visuospatial attention in decision-making under risk. *NeuroImage*, *103*, 75–80.

Stutzer, A., Goette, L., & Zehnder, M. (2011). Active decisions and prosocial behaviour: A field experiment on blood donation*. *The Economic Journal*, *121*(556), F476–F493.

Sunstein, C. R. (2014). Nudging: A very short guide. *Journal of Consumer Policy*, *37*(4), 583–588.

Sunstein, C. R. (2015). Fifty shades of manipulation. *SSRN Electronic Journal*.

Sunstein, C. R. (2016). Do people like nudges. *Administrative Law Review*, *68*(2), 177–232.

Sunstein, C. R., & Reisch, L. A. (2016). Behaviorally green: Why, which and when defaults can help. In *New perspectives for environmental policies through behavioral economics* (pp. 161–194). Springer.

Suzuki, M., & Gottlieb, J. (2013). Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. *Nature neuroscience*, *16*(1), 98–104.

Taylor, S. E., & Thompson, S. C. (1982). Stalking the elusive "vividness" effect. *Psychological Review*, *89*(2), 155–181.

Taylor, S. F., Welsh, R. C., Wager, T. D., Phan, K. L., Fitzgerald, K. D., & Gehring, W. J. (2004). A functional neuroimaging study of motivation and executive function. *NeuroImage*, *21*(3), 1045–1054.

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

Thaler, R. H., & Sunstein, C. R. (2003). Libertarian paternalism. *American Economic Review*, *93*(2), 175–179.

Torrence, R. D., & Troup, L. J. (2017). Event-related potentials of attentional bias toward faces in the dot-probe task: A systematic review. *Psychophysiology*.

Treue, S. (2003). Visual attention: The where, what, how and why of saliency. *Current Opinion in Neurobiology*, *13*(4), 428–432.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, Mass.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323.

Uddin, L. Q. (2015). Salience processing and insular cortical function and dysfunction. *Nature reviews. Neuroscience*, *16*(1), 55–61.

Vogel, E. K., & Luck, S. J. (2000). The visual n1 component as an index of a discrimination process. *Psychophysiology*, *37*(2), 190–203.

von Dawans, B., Kirschbaum, C., & Heinrichs, M. (2011). The trier social stress test for groups (tsst-g): A new research tool for controlled simultaneous social stress exposure in a group format. *Psychoneuroendocrinology*, *36*(4), 514–522.

Weber, E., & Kirsner, B. (1997). Reasons for rank-dependent utility evaluation. *Journal of Risk and Uncertainty*, *14*(1), 41–61.

Williams, L. M., Liddell, B. J., Kemp, A. H., Bryant, R. A., Meares, R. A., Peduto, A. S., & Gordon, E. (2006). Amygdala-prefrontal dissociation of subliminal and supraliminal fear. *Human Brain Mapping*, *27*(8), 652–661.

Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, *125*(4), 387.

World Bank. (2015). *World development report 2015: Mind, society, and behavior*. World Bank Group.

Yechiam, E., & Hochman, G. (2013). Loss-aversion or loss-attention: The impact of losses on cognitive performance. *Cognitive psychology*, *66*(2), 212–231.

Yoshida, W., Seymour, B., Friston, K. J., & Dolan, R. J. (2010). Neural mechanisms of belief inference during cooperative games. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *30*(32), 10744–10751.

Young, D. L., Goodie, A. S., Hall, D. B., & Wu, E. (2012). Decision making under time pressure, modeled in a prospect theory framework. *Organizational behavior and human decision processes*, *118*(2), 179–188.

Zajonc, R. B. (1965). Social facilitation. *Science*, *149*(3681), 269–274.

Zaki, J., Schirmer, J., & Mitchell, J. P. (2011). Social influence modulates the neural computation of value. *Psychological Science*, *22*(7), 894–900.

Zhang, D., Liu, Y., Wang, L., Ai, H., & Luo, Y. (2017). Mechanisms for attentional modulation by threatening emotions of fear, anger, and disgust. *Cognitive, affective & behavioral neuroscience*, *17*(1), 198–210.

Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., & Berns, G. S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron*, *42*(3), 509–517.

Declaration of Co-authorship

The dissertation is based on three studies, which are depicted in part II and III.

Part II embraces two studies investigating the salience in lotteries as defined by Salience Theory. They are based on an experiment where reaction time data and EEG data was recorded. The first study describes results from behavioral data (chapter 8). I am the single author of this study. The second study looks at results from EEG data (chapter 9). This study was conducted together with Professor Wolfgang H.R. Miltner and Dr. Marcel Franz. Within the division of tasks, I came up with the research question and worked on conceptualisation, theory and data gathering overproportionally. Data analysis was done in a proportional way.

Part III (chapter 10) is based on the publication *Bruns, H., Kantorowicz-Reznichenko, E., Klement, K., Jonsson, M. L., Rahali, B. (2018). Can nudges be transparent and yet effective? Journal of Economic Psychology, 65, 41-59.* Within the division of tasks, I worked overproportionally on conceptualisation, theory and literature research. Data gathering in pre-tests was done proportionally.

Statutory Declaration

Erklärung nach §4 Abs. 1 Promotionsordnung

Hiermit erkläre ich,

1. dass mir die geltende Promotionsordnung bekannt ist;

2. dass ich die Dissertation selbst angefertigt, keine Textabschnitte eines Dritten oder eigener Prüfungsarbeiten ohne Kennzeichnung übernommen und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben habe;

3. dass ich bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskriptes keine unzulässige Hilfe in Anspruch genommen habe;

4. dass ich nicht die Hilfe eines Promotionsberaters in Anspruch genommen habe und dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir f?r Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen;

5. dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe;

6. dass ich nicht die gleiche, eine in wesentlichen Teilen ähnliche oder eine andere Abhandlung bei einer anderen Hochschule bzw. anderen Fakultät als Dissertation eingereicht habe.

Katharina Klement