Three Essays in Behavioral Economics: On Motives, Beliefs and Motivated Beliefs

Dissertation

zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaften (Dr. rer. pol.)

vorgelegt von

Baiba Renerte

Universität Konstanz

Sektion Politik-Recht-Wirtschaft

Fachbereich Wirtschaftswissenschaften

Konstanz, 2019

Tag der mündlichen Prüfung: 26. September 2019

1. Referent: Prof. Dr. Urs Fischbacher
2. Referent: Prof. Dr. Sebastian Fehrler
3. Referentin: Prof. Dr. Susann Fiedler

Acknowledgments

Throughout the writing of this dissertation, I have received a great deal of support and assistance. I would like to thank everyone who provided continued advice and contributed to turning my doctoral studies into a journey of inspiration, exploration and learning.

First and foremost, I would like to express my deepest appreciation to my dissertation committee. I sincerely thank my first supervisor Urs Fischbacher for his guidance through each stage of writing this dissertation. He has taught me more than I could ever give him credit for here. At all times, he has been approachable, encouraging and supportive of my career goals. Above all, he has shown me, by example, what a great researcher should be.

I would like to extend my sincere appreciation to my second supervisor Sebastian Fehrler who has been a valuable help throughout the writing process, from defining the research objectives up to presenting the research results and implications with great clarity of thought. I would also like to express my deepest gratitude to my third supervisor Susann Fiedler whose research has provided inspiration for my research interests and whose advice has been vital for this dissertation.

My research would have been impossible without the aid and support of the Thurgau Institute of Economics in Switzerland, the Graduate School of Decision Sciences in Germany, the German Research Foundation, and the German-French University. I am thankful for the fellowship and experiment funding that made my dissertation projects possible as well as the travel funding that allowed enriching my dissertation through conference visits and research stays.

I have greatly benefited from the whole team of the Thurgau Institute of Economics, especially (in reverse alphabetical order) Irenaeus Wolff, Katrin Schmelz, Moritz Janas, Jan Hausfeld, David Grammling, Sebastian Fehrler, Fabian Dvorak, David Dohmen, Adrian Chadi, Deepti Bathia, and others, and I would like to thank them for their continued help with discussing and editing my work during our paper discussions, project presentations, and even coffee breaks. And a special thank you goes to Barbara Monstein and Madeleine Hafner for making all of our daily work as smooth and pleasant as possible.

I highly appreciate everyone with whom I have had the pleasure to work with during these three-and-a-half years. I would like to thank the amazing people at the University of Konstanz who gradually turned from colleagues to dear friends. Special thanks goes to Jan Hausfeld, Christian Neumeier, Konstantin Hesler, Thomas Hattenbach, Dominik Bauer, Fabian Dvorak, Franziska Deutschmann, Carl Maier, Torsten Twardawski, and others. I greatly enjoyed this amazing time together!

Finally, I am profoundly grateful to my loving and supportive husband Klavs Renerts whose endless energy, enthusiasm and encouragement made me achieve goals I would have otherwise deemed impossible and who kept a sense of humor when I had lost mine. Many thanks to my dear family and friends in Riga for their never-ending love and support and to my wonderful friends in Konstanz and Zürich for their encouragement and delightful distraction.

Contents

Summary

According to the philosopher Charles Peirce, the central function of human thought is to produce the beliefs that people come to hold—about their world, about others within it and about themselves in comparison to others. These beliefs become rules of action that determine the subsequent decisions and thoughts, which in turn generate subsequent beliefs (Pajares, 2008).

In this dissertation, I examine how people's beliefs and motives affect their economic decision making. On the one hand, informed beliefs can help people make more optimal decisions; on the other hand, holding faulty beliefs can hinder the decision-making process. For example, people's preferences can motivate and affect their beliefs, leading to biased beliefs that feel objective (Epley and Gilovich, 2016). Economically relevant examples of beliefs that fulfill one's psychological and functional needs include confidence in ones' abilities. On an individual level, overconfidence is perhaps the most common manifestation of the motivated-beliefs phenomenon (Bénabou and Tirole, 2016). On an interpersonal level, strategic motives and information neglect are relevant examples of related phenomena.

How do motivated beliefs about oneself affect economic decision making? How do these beliefs cause people to exhibit irrational behavior and make sub-optimal decisions? And how does reasoning about others interact with people's beliefs and preferences? In this dissertation, I examine these and other questions using the methods of behavioral and experimental economics in three distinct chapters.

The first chapter "Motives in economic interactions: An (interactive) eye-tracking study" is joint work with Urs Fischbacher and Jan Hausfeld. In this study, we apply eye-tracking technology interactively, that is, by showing gaze patterns of one person to another. We investigate how untrained observers use decision makers' gaze patterns to uncover motives behind their decisions. Gaze patterns can reveal motives, as decision makers dedicate more attention to items particularly relevant for these motives (Fiedler et al., 2013).

In our laboratory study, the decision makers make a variety of allocation decisions in slider-type Social Value Orientation situations that are commonly used to measure social preferences (Murphy et al., 2011). We then display either non-strategic or strategic gaze

patterns of decision makers to observers and let the latter infer the former's motives and actions. While the non-strategic gaze patterns provide strategically undisturbed information, the strategic (real-time observed) gaze patterns require taking the future consequences of the interaction into account. When gaze is non-strategic, observers can recognize the more prosocial and generous decision makers. In contrast, when gaze is strategic, the eye-tracked decision makers successfully alter their gaze to appear more prosocial. Consequently, the less prosocial decision makers are chosen for future interaction more often in the strategic settings than in the non-strategic settings, leading to increased payoffs for them and decreased payoffs for the observers. We conclude that people can be skillful users of eye-tracking, both as signal receivers and as signal senders.

The title of the second chapter is "Beliefs about others: Information neglect over type projection". It is a collaboration with Sebastian Fehrler and Irenaeus Wolff. In this study, we investigate how people learn about others in a standard laboratory task, in which types (colored balls) are drawn from one out of several states of the world (colored urns). In many games of imperfect information, people make Bayesian inferences about other people's types based on the information that is contained in their own type. Several behavioral theories of belief updating, such as social projection, information projection, or type projection (Breitmoser, 2019, among others), even start from the assumption that people project, or even over-project, their own type onto others.

In our experiment with three consecutively developed treatments, we make the types very salient and thus create favorable conditions for finding rational projection and over-projection. However, we find no evidence for over-inference. Instead, between 50% and 70% of the participants in our experiment neglect this important source of information entirely and base their choices only on the prior probabilities. The consecutive experimental treatments allow us to conclude that the observed phenomenon is very robust. Our findings are thus inconsistent with projection and are instead consistent with the other strand of the literature that documents the neglect of information.

The third chapter "Beliefs about oneself: Overconfidence and group composition in investment decisions" is a first-author study in collaboration with Jan Hausfeld and Torsten Twardawski. In this study, we examine whether groups make less optimal investments in risky prospects when more overconfident individuals are involved in the decision. Overconfidence is a commonly observed motivated belief that can lead to material real-life consequences (Bénabou, 2015). And yet, the focus of overconfidence studies has mostly been on individual decision making, even though important economic decisions are often made by groups rather than individuals.

We design a laboratory experiment with two types of investment situations: either with fixed chances of success or with performance-dependent chances of success. In contrast to other studies, our design thus allows to disentangle the overconfident group members' perceived ability to "beat the odds" of the investment situations, i.e. the over-placement component of overconfidence. We show that group investment levels increase when the group is composed of more overconfident members, but the two types of investment situations are treated differently. The group members decide less optimally when a possibility to "beat the odds" of success is provided, and this mindset, facilitated by communication, spills over to cases where such a possibility is no longer provided. Our results thus allow drawing policy implications for board composition with respect to overconfidence.

Zusammenfassung

Nach Ansicht des Philosophen Charles Peirce besteht die zentrale Funktion des menschlichen Denkens darin, die Überzeugungen hervorzubringen, die die Menschen vertreten—über ihre Welt, über andere in dieser Welt und über sich selbst im Vergleich zu anderen. Diese Überzeugungen werden zu Handlungsregeln, die die nachfolgenden Entscheidungen und Gedanken bestimmen, was wiederum nachfolgende Überzeugungen erzeugt (Pajares, 2008).

In dieser Dissertation untersuche ich, wie sich die Überzeugungen und Motive der Menschen auf ihre wirtschaftliche Entscheidungsfindung auswirken. Einerseits können fundierte Überzeugungen Menschen dabei helfen, optimalere Entscheidungen zu treffen, andererseits kann die Beibehaltung falscher Überzeugungen den Entscheidungsprozess behindern. Zum Beispiel können die Präferenzen der Menschen ihre Überzeugungen motivieren und beeinflussen, was zu voreingenommenen Überzeugungen führt, die sich objektiv anfühlen (Epley and Gilovich, 2016). Wirtschaftlich relevante Beispiele für Überzeugungen, die die eigenen psychologischen und funktionalen Bedürfnisse erfüllen, umfassen z.B. das Vertrauen in die eigenen Fähigkeiten. Auf individueller Ebene ist Selbstüberschätzung vielleicht die häufigste Manifestation des Phänomens der motivierten Überzeugungen (Bénabou and Tirole, 2016). Auf zwischenmenschlicher Ebene sind strategische Motive und Vernachlässigung von Informationen relevante Beispiele für einhergehende Phänomene.

Wie wirken sich motivierte Überzeugungen über sich auf die wirtschaftliche Entscheidungsfindung aus? Wie bewirken die Überzeugungen, dass Menschen irrationales Verhalten zeigen und suboptimale Entscheidungen treffen? Und wie interagiert das Denken über andere mit den Überzeugungen und Präferenzen der Menschen? In dieser Dissertation untersuche ich diese und andere Fragen mit den Methoden der Verhaltens- und experimentellen Ökonomik in drei unterschiedlichen Kapiteln.

Das erste Kapitel "Motive in ökonomischen Interaktionen: Eine (interaktive) Eye-Tracking-Studie" ist eine gemeinsame Arbeit mit Urs Fischbacher und Jan Hausfeld. In dieser Studie wird die Eye-Tracking-Technologie interaktiv eingesetzt, indem die Blickmuster einer Person einer anderen Person gezeigt werden. Wir untersuchen, wie unerfahrene Beobachter die Blickmuster von Entscheidungsträgern verwenden, um die Motive

für deren Entscheidungen aufzudecken. Blickmuster können diese Motive aufdecken, da bestimmte Motive zwingend mit bestimmten Arten von Informationssuche einhergehen. Der Entscheidungsträger sollte dann, den für dieses Motiv besonders relevanten Informationen mehr Aufmerksamkeit widmen (Fiedler et al., 2013).

In unserer Laborstudie treffen die Entscheidungsträger eine Reihe von Zuordnungsentscheidungen in Situationen mit sozialer Wertorientierung, die häufig zur Messung sozialer Präferenzen verwendet werden (Murphy et al., 2011). Wir zeigen den Beobachtern dann entweder nicht-strategische oder strategische Blickmuster von Entscheidungsträgern und lassen sie auf deren Motive und Handlungen schließen. Während die nicht-strategischen Blickmuster strategisch unverfälschte Informationen liefern, müssen die zukünftigen Konsequenzen der Interaktion bei den strategischen Blickmustern berücksichtigt werden. Wenn der Blick nicht-strategisch ist, können Beobachter die prosozialen und großzügigen Entscheidungsträger erkennen. Wenn der Blick dagegen strategisch ist, ändern die Entscheidungsträger erfolgreich ihren Blick, um prosozialer zu wirken. Infolgedessen werden die weniger prosozialen Entscheidungsträger in den strategischen Szenarien häufiger für die künftige Interaktion ausgewählt als in den nicht-strategischen Szenarien, was insgesamt zu höheren Auszahlungen für diese Entscheidungsträger und geringeren Auszahlungen für die Beobachter führt. Daraus schließen wir, dass Menschen geschickte Benutzer von Eye-Tracking sein können—sowohl als Signalempfänger, als auch als Signalsender.

Der Titel des zweiten Kapitels lautet "Überzeugungen über Andere: Vernachlässigung von Informationen anstatt von Typprojektion". Es ist eine Zusammenarbeit mit Sebastian Fehrler und Irenaeus Wolff. In dieser Studie untersuchen wir, wie Menschen in einer Standardlaboraufgabe von anderen Menschen lernen; Typen (farbige Kugeln) werden aus einem von mehreren Zustanden der Welt (farbige Urnen) gezogen werden. In vielen Spielen mit unvollständigen Informationen können Menschen anhand der Informationen, die in ihrem eigenen Typ enthalten sind, bayesianische Schlussfolgerungen über die Typen anderer Menschen ziehen. Einige Verhaltenstheorien zur Aktualisierung der Überzeugungen wie soziale Projektion, Informationsprojektion oder Typprojektion (Breitmoser, 2019, u.a.) gehen sogar davon aus, dass Menschen ihren eigenen Typ auf andere projizieren oder überprojizieren.

In unserem Experiment mit drei nacheinander entwickelten Interventionen machen wir die Typen sehr auffällig und schaffen so günstige Bedingungen für die Suche nach rationaler Projektion und Überprojektion. Wir finden jedoch keine Hinweise auf eine Überinferenz. Stattdessen vernachlässigen zwischen 50% und 70% der Teilnehmer an unserem Experiment diese wichtige Informationsquelle vollständig und stützen ihre Auswahl nur auf die vorherigen Wahrscheinlichkeiten. Die aufeinanderfolgenden experimentellen Interventionen zeigen, dass das Phänomen robust ist. Unsere Ergebnisse stimmen daher nicht mit

These der Projektion überein, sondern mit dem anderen Teil der Literatur, der die Vernachlässigung von Informationen dokumentiert.

Das dritte Kapitel "Überzeugungen von sich selbst: Selbstüberschätzung und Gruppenzusammensetzung in Investitionsentscheidungen" ist eine Studie in Zusammenarbeit mit Jan Hausfeld und Torsten Twardawski. In dieser Studie untersuchen wir, ob Gruppen weniger optimal in riskante Investitionsmöglichkeiten investieren, wenn mehr übermütige Einzelpersonen an der Entscheidung beteiligt sind. Selbstüberschätzung ist eine allgemein beobachtete, motivierte Überzeugung, die zu materiellen Konsequenzen führen kann (Bénabou, 2015). Dennoch lag der Schwerpunkt der bisherigen Studien zu diesem Thema hauptsächlich auf der individuellen Entscheidungsfindung, obwohl wichtige wirtschaftliche Entscheidungen häufig eher von Gruppen als von Einzelpersonen getroffen werden.

Wir konzipieren ein Laborexperiment mit zwei Arten von Investitionssituationen: entweder mit festen Erfolgschancen oder mit leistungsabhängigen Erfolgschancen. Im Gegensatz zu anderen Studien ermöglichen unsere Methoden, die Überplatzierungskomponente der Selbstüberschätzung (d.h. die wahrgenommene Fähigkeit der übermütigen Gruppenmitglieder, bessere Chancen zu haben als es beschrieben ist) zu analysieren. Wir zeigen, dass das Investitionsniveau der Gruppe steigt, wenn sich die Gruppe aus übermütigeren Mitgliedern zusammensetzt, aber die beiden Arten von Investitionssituationen unterschiedlich behandelt werden. Die Gruppenmitglieder entscheiden weniger optimal, wenn die Möglichkeit besteht, die Erfolgschancen zu übertreffen. Weiterhin überträgt sich diese Einstellung, die durch Kommunikation erleichtert wird, auch auf Fälle, in denen eine solche Möglichkeit nicht mehr gegeben ist. Unsere Ergebnisse erlauben es daher, Implikationen für die Zusammensetzung des Boards in Bezug auf Selbstüberschätzung zu ziehen.

Chapter 1

Motives in economic interactions: An (interactive) eye-tracking study

Urs Fischbacher, Jan Hausfeld & Baiba Renerte

Abstract

Gaze patterns can reveal motives, as decision makers dedicate more attention to items particularly relevant for these motives. We investigate how untrained observers use decision makers' gaze patterns to uncover motives behind their decisions. We display either non-strategic or strategic gaze patterns of decision makers to observers and let the latter infer the former's motives and actions. While the *non-strategic* gaze patterns provide strategically undisturbed information, the *strategic* gaze patterns require taking the future consequences of the interaction into account. When the gaze is non-strategic, observers can recognize the more prosocial and generous decision makers. In contrast, when the gaze is strategic, the eye-tracked decision makers successfully alter their gaze to appear more prosocial. Consequently, less prosocial decision makers are chosen for future interaction more often in the strategic settings than in the non-strategic settings, leading to increased payoffs for them and decreased payoffs for the observers.

JEL classification: C91, C92, D87, D91.

Keywords: motives, social preferences, social value orientation, eye-tracking.

1.1. Motivation

In this study, we demonstrate that gaze patterns can reveal motives to others. Imagine two decision makers, called allocators, who have to make an allocation choice. Their own payoff is the same in all options and only the payoff for another person differs between the options. The first allocator might only care about her own payoff and randomly chooses an option, which by chance happens to be the most beneficial option for the other person. In contrast, the second allocator might want to maximize the other person's payoff or efficiency and deliberately searches for the most beneficial allocation for the other person. Even though the option set and the chosen action are the same, the underlying motives behind the action are different.

We study these motives using eye-tracking technology. We extend the use of eye-tracking from being a passive tool for decision-process analysis to being an active component of the interaction. Since eyes are natural means of communication, observing recorded *non-strategic* gaze patterns of an allocator should help revealing her true underlying motives to an observer and distinguishing her type. It is, however, unclear whether real-time interactive *strategic* gaze (with incentives for the allocator to disguise her motives) would still help revealing the true underlying motives to an observer.

In our laboratory study, the allocators make a variety of allocation decisions in slider-type Social Value Orientation situations that are commonly used to measure social preferences (Murphy et al., 2011). We show either the eye-tracked gaze patterns during the allocation decision (in GazeVideo or GazePicture format, showing the motive) or the actual chosen allocation (in Choice format, showing the action) of the allocators to the observers. The former thus informs the observers about the search motive and the latter about the action of the allocator. The observers then attempt to identify the more prosocial and generous allocators and choose them for future interaction, as the observers receive the payoff allocated to them by their chosen allocator. We examine the behavior of both the allocators and the observers in strategic compared to non-strategic settings.

We firstly find that non-strategic gaze patterns have significant out-of-sample predictive power for the allocators' prosociality. Secondly, we show that untrained observers in *non-strategic* settings do intuitively understand how to recognize the more prosocial or generous allocators and choose them as partners for further interaction. We thus find evidence that people are able to correctly interpret non-strategic gaze patterns. Thirdly, in *strategic* settings, the eye-tracked allocators understand how to gaze strategically and simulate prosociality or generosity, in order to shift the partner choice in their favor. We thus find evidence that people are able to generate strategic gaze patterns. This results in less prosocial subjects being chosen for future interaction more often. We conclude that people can be skillful users of eye-tracking, both as signal receivers and as signal senders.

1.1.1. *Related literature*

Eye-tracking is an affordable and precise method for measuring information acquisition.[1] It has already been used to study how people interact in games and to infer player types from recorded gaze (Knoepfle et al., 2009; Wang et al., 2010; Fiedler et al., 2013; Polonio et al., 2015; Stewart et al., 2016). We advance this literature by applying eye-tracking interactively instead of passively, that is, by displaying the (strategic) gaze patterns of the allocators to the observers. When eye-tracking is interactive and used as a communication device, gaze becomes a part of the strategic decision of the transmitting allocator. She not only needs to find the relevant information, but also plan what to communicate to the receiving observer.

There is ample evidence that people can learn from observing gaze. Problem-solving performance benefits from seeing the gaze patterns of others who previously worked on the same task (Velichkovsky, 1995; Stein and Brennan, 2004; Litchfield et al., 2010; Litchfield and Ball, 2011). Bilateral real-time gaze transmission positively affects performance in visual search tasks (Brennan et al., 2008; Neider et al., 2010). Meanwhile, the gaze does not simply reflect the search behavior, but also the intended advice (or, on the contrary, hides the intended advice). Thus, the gaze can be difficult to interpret if there are various motives for looking at a piece of information, and benefits of gaze transfer can also be limited (Müller et al., 2013, 2014).

In fact, observing gaze patterns might not be informative at all if the allocators have an incentive to be unpredictable or even to deceive the observers. On the one hand, it could be the case that gaze *always* conveys true information even if this is not in the allocator's interest. After all, gaze patterns of card game players can reflect the numerical value of their hands (Holmes et al., 2016) and gaze patterns and pupil dilation can reveal information about motives in sender-receiver games (Wang et al., 2010).[2] On the other hand, it could be the case that gaze does not convey true information if the allocator chooses to hide her true preference and gazes strategically (Foulsham and Lock, 2015).

However, previous studies do not address the question whether manipulating gaze works also if incentives are commonly known. We are especially interested in a setup, in which it is common knowledge that the allocator might gaze strategically. Hausfeld et al. (2018)

[1]See Duchowski (2002) for a survey of eye-tracking studies across disciplines. See also Lahey and Oxley (2016) for a recent discussion on how eye-tracking can inform economic theory.

[2]Meanwhile, many people, even professional interrogators, mistakenly believe that deception is associated with evasive gaze patterns (Vrij, 2004).

examine how people communicate their intended choice with their eyes in simple coordination and discoordination games.[3] We extend this line of research to economic interactions where social preferences play a crucial role.

1.2. Method

In our experiments, participants make decisions while their eye movements are being tracked. Their gaze patterns are shown to others who attempt to guess the prosociality or generosity of the tracked participants based on this information. We distinguish between *non-strategic* decisions with no further consequences for the tracked participants and *strategic* decisions with consequences in further interactions for the tracked participants.

There are two types of participants in our study: (i) "allocators" whose gaze and choices are either non-strategic because they were recorded in an earlier session or strategic because they were transmitted interactively within the same session and (ii) "observers" who interpret the gaze and choice information of the allocators and attempt to identify the more prosocial or generous allocators for future interaction.

Our study consists of two experiments. In the PILOT experiment, we show allocators' non-strategic gaze patterns and choices to observers. In this experiment, we focus on whether and how gaze patterns reveal motives to others. In the MAIN experiment, we additionally introduce strategic incentives for allocators and transmit allocators' strategic gaze and choices to the observers, as well as compare them to a non-strategic control. This experiment allows us to establish the impact of strategic incentives on the allocators' gaze and choices as well as the observers' comprehension of this information.

In the following sub-section 1.2.1, we introduce how the allocators in all sessions choose their preferred allocations (see 1.2.1.1) and how the observers go through the main stages of the experiments: the Information stage and the Assessment stages (see 1.2.1.2 and 1.2.1.3). We then describe the setup of the two experiments in sub-section 1.2.2 and the hypotheses and experimental procedure in sub-sections 1.2.3 and 1.2.4.

1.2.1. Stages of the experiments

1.2.1.1. Decision situations for allocators

The decision situations are slider-type Social Value Orientation (SVO) allocation decisions with five possible allocation options adjusted from Murphy et al. (2011), as depicted in Fig-

[3] Hausfeld et al. (2018) find that participants naturally understand that gaze conveys information. If participants want to coordinate, people amplify their gaze on the chosen option. However, if the transmitting participant is incentivized to be unpredictable, gaze becomes uninformative.

ure 1.1 below.[4] The SVO situations were originally designed to measure the magnitude of concern that the allocators have for others. The allocation decisions are incentivized via one randomly chosen paid-out allocation (random dictator in a random game procedure). The gaze and choices are recorded and used in the Information stage.

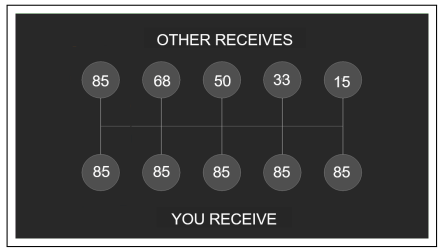


Fig. 1.1. Adjusted decision screen as seen by all allocators (the font sizes were smaller in the experiment to prevent peripheral perception, see the screens in Figure 1.2 below).

1.2.1.2. Information stage for observers

We initially use three information formats for showing the allocators' decisions to the observers: (i) "Choice" (action information), (ii) dynamic gaze pattern "GazeVideo" (motive information) and (iii) static gaze pattern "GazePicture" (motive information). For comparability, the information in all formats is shown for the length of time the allocator took to make the decision.

In Choice information, a rectangle highlights the allocator's chosen action, as depicted in Figure 1.2 (left sub-figure) below. In GazeVideo information, whenever an option is being looked at, the respective payoff lights up for as long at is being inspected. This results in a sequence of areas of interest (AOIs) lighting up, as illustratively depicted in Figure 1.2 (middle sub-figure) below.[5] In GazePicture[6] information, a scanpath connects the inspected AOI in the sequence they are looked at, starting from the first AOI (green dot) and ending with the last AOI (red dot), as depicted in Figure 1.2 (right sub-figure) below.[7]

[4]An overview of all 12 situations is depicted in Appendix A. These situations were used twice—once jittered and once non-jittered—to construct a total of 24 situations. Note that for the purposes of eye-tracking processing, we reduce the number of available displayed options within each SVO allocation decision to five and standardize the displayed values from 0 to 100.

[5]We define the payoff AOIs to include also the rectangle area in the close vicinity of the circles, to allow for slight imprecisions in the eye-tracking data. Accordingly, we define 10 AOIs in the vicinity of the payoff numbers, 2 AOIs in the vicinity of the legend texts, and 2 AOIs covering the rest of the screen. In result, the whole screen is divided into (more or less informative) AOIs.

[6]We explore this information type only in the INFOPILOT sessions, as additionally explained in sub-section 1.4.1 below.

[7]The hollow circles increase in size with the time spent looking at the respective AOIs in accordance to the rule $100 \times (1 - e^{-0.25 \times DurationSeconds})$ and move toward the center of the picture with decision time.

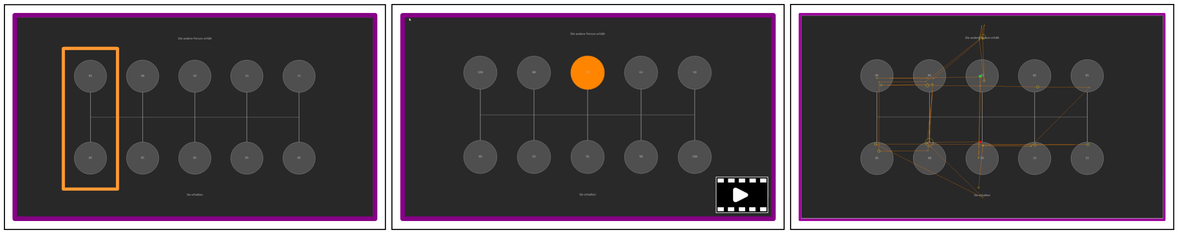


Fig. 1.2. Decision screen as seen by all observers in the Information stage in the Choice (left), GazeVideo (middle) and GazePicture (right) treatment.

1.2.1.3. Assessment stages for observers

The allocators and observers are rematched in new groups in each period. We examine how the observers assess the preferences of the matched allocators using two methods: predictions of next decisions and partner choice.

(i) Prediction stage for observers

After the Information stage, (three) next randomly drawn situations are shown to the observers. They do not receive any information about the matched allocators' Choice, GazeVideo or GazePicture in these situations but need to state their predictions (and non-incentivized prediction certainty) about the allocators' actions in these situations. The observers also see an "average choice" line, positioned closest to having equal numbers of choices on both sides.[8] The correctness of the predictions is incentivized using a two-step procedure. Firstly, the observers receive a fixed payment if they correctly guess on which side of the average choices of others (1 of 2) was the allocator's choice. Secondly, they additionally receive a fixed payment if they correctly guess the exact chosen option (1 of 5).

(ii) Partner choice stage for observers

The observers first see the information about two different allocators and then have to decide with which one they want to interact in the next situations.[9] The implicit task for the observers is thus to identify the more prosocial or generous of the two matched allocators in the group. We always use purple and blue colored frames around the decision screens to color-code the first and second allocator for the observers.

[8]For the "average choice" line, we draw an orange line right between two successive options in the given situation, such that its position is the closest possible to the actual average allocators' choices in previous sessions, see decision screens in Appendix B. We chose the two-step procedure because the observers might be able to predict the correct direction from the average but not necessarily the exact correct option.

[9]One exception is the INFOPILOT sessions, as explained in sub-section 1.4.1 below.

1.2.2. *Design of the experiments*

In the following sub-sections, we introduce the design of the non-strategic PILOT experiment (see 1.2.2.1) and the MAIN experiment that includes strategic incentives (see 1.2.2.2), as summarized in Table 1.1 below.

Experiment	Sessions	Focus
PILOT	INFOPILOT	Comparison of the presentation formats for gaze patterns
	PARTNERPILOT	Test for the recognition of prosocial and generous types
MAIN	INTERACTIVE	Test for the introduction of strategic incentives
	RECORDED	Control with the same participants without strategic incentives

Table 1.1: Summary of the experiment design.

1.2.2.1. *Design of the PILOT experiment*

In the PILOT experiment, we use choice and gaze data from earlier sessions, in which we simply recorded gaze and choices. We carry out two types of sessions with recorded data, in which the allocators had no incentives to gaze strategically. The two session types differ in the following way. In the "INFOPILOT" sessions, each observer sees information about *one* single allocator in each round and then states only her predictions of this allocator's actions in the next three decisions, as summarized in Figure 1.3 below. In the "PARTNERPILOT" sessions, the observer sees the information about *two* allocators one after another and then not only states her predictions of the two allocators' actions in the next three decisions, but also makes a partner choice, i.e. she picks with which of the two allocators to interact in these decisions, as summarized in Figure 1.4 below. The observer receives the payoff allocated to her by the chosen allocator in these decisions.

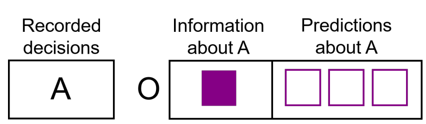


Fig. 1.3. Overview of the INFOPILOT session design in the non-strategic PILOT experiment with an allocator (A) and an observer (O) per group.

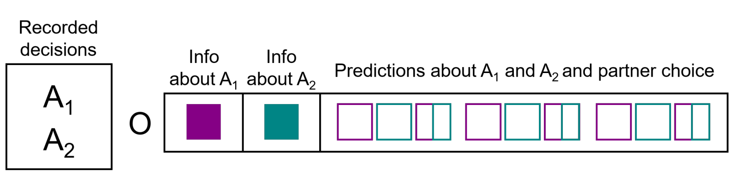


Fig. 1.4. Overview of the PARTNERPILOT session design in the non-strategic PILOT experiment with two allocators (A) and an observer (O) per group.

The INFOPILOT sessions comprise 36 periods in total: 12 periods for each of the three information formats (Choice, GazePicture and GazeVideo) respectively in counter-balanced blocks. The PARTNERPILOT sessions comprise 24 periods in total: 12 periods for Choice and GazeVideo information formats respectively in counter-balanced blocks. The incentives for the observers in the INFOPILOT and PARTNERPILOT sessions remain the same: (i) prediction correctness for each information treatment about a randomly chosen allocator's decision in one of the three next decisions and (ii) payoff (for the other) that was chosen by the picked allocator (PARTNERPILOT) in one of the next three decisions.

1.2.2.2. Design of the MAIN experiment

In the MAIN experiment, we carry out two different types of sessions. In the "INTERACTIVE" sessions, the observers see strategic interactive information from allocators within the same session. In control "RECORDED" sessions, the observers see non-strategic recorded information from allocators from an already conducted session. Importantly, the *same* eye-tracked allocators are the transmitters in RECORDED and INTERACTIVE sessions.[10]

In the INTERACTIVE sessions, it is known to all participants that the allocators' gaze or choice (from the first decision but not the second decision) will be shown to the observers, as summarized in Figure 1.5 below. The allocators have an incentive to gaze strategically, as it is also common knowledge that their payoff in the second decision will be *multiplied* by the number of observers (two, one or zero) who chose this allocator in the partner choice stage after the first decision. In the control RECORDED sessions, however, the recorded allocators had no incentives to gaze strategically, as summarized in Figure 1.6 below.

Both INTERACTIVE and RECORDED sessions comprise 9 periods in total: 4 periods with GazeVideo information and 4 periods with Choice information (counter-balanced blocks), and 1 final period in which the information format is determined by the observers (either GazeVideo or Choice). The incentives for the *observers* in the INTERACTIVE and RECORDED sessions remain the same: (i) payoff (for the other) that was chosen by a randomly determined allocator in the first decision, (ii) payoff (for the other) that was chosen by the picked allocator in the second decision and (iii) prediction correctness about a randomly determined allocators' second decision. Importantly, the incentives for the *allocators* in the INTERACTIVE sessions are: (i) payoff that they themselves chose in the first decision and (ii) payoff that they themselves chose in the second decision, multiplied by the number of the observers (two, one or zero) who chose this allocator in the partner choice stage.

[10]This feature of our design was organized as follows. Before the INTERACTIVE (and all other) sessions, we added a part that allows all participants to play through all decision situations once. Accordingly, in the RECORDED sessions, we show the recorded gaze and choices from this part of the INTERACTIVE sessions.

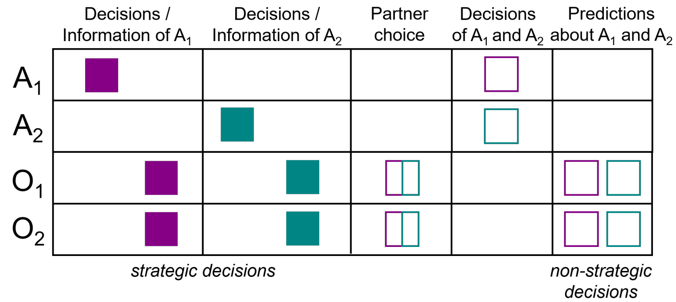


Fig. 1.5. Overview of the *strategic* INTERACTIVE session design in the MAIN experiment with strategic incentives for the allocators A_1 and A_2 in the first (transmitted) decision. Their payoff from the second (non-transmitted) decision is multiplied by the number of the observers (two, one or zero) who chose them in the partner choice stage.

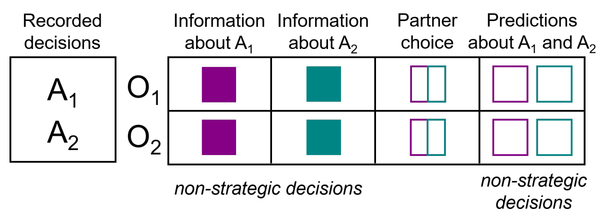


Fig. 1.6. Overview of the *control non-strategic* RECORDED session design in the MAIN experiment (where A_1 and A_2 are recorded allocators from an earlier session and O_1 and O_2 know this) for direct comparison with the INTERACTIVE sessions.

1.2.3. *Hypotheses*

This eye-tracking study allows to investigate whether and how gaze patterns reveal underlying motives behind an action to observers. We examine the PILOT and MAIN experiments with respect to (i) the allocators' allocated attention and payoffs and (ii) the observers' predictions and partner choices. In general, we expect that the observers are able to recognize the allocators' motives based on their gaze and, conversely, that the allocators are able to conceal their motives in their gaze if aware that they will be observed and their choices or gaze could bear consequences.

In the INFOPILOT sessions, we expect that it will be easier for the observers to interpret GazeVideo information than GazePicture information.[11] In the PARTNERPILOT sessions, we generally expect that the observers are able to correctly interpret non-strategic gaze patterns and (i) choose the more prosocial or generous allocator significantly more often and (ii) state predictions with correctness significantly above chance levels. These hypotheses mostly stem from the findings of Fiedler et al. (2013) who show that the SVO angles correlate

[11]We use the more preferred information format from this comparison in further sessions.

with measures of attention and Hausfeld et al. (2018) who show that participants understand how gaze is informative of the action.

In the INTERACTIVE sessions of the MAIN experiment, we hypothesize that the allocators strategically allocate more equal payoffs in the first (strategic because transmitted) decision than in the second (non-strategic because non-transmitted) decision. We also expect that they gaze more strategically during the first decision than in the second decision and than in the RECORDED session decisions.[12] Consequently, for the observers in the INTERACTIVE sessions, we expect less correct partner choices and less correct predictions than in the RECORDED sessions (in line with Hausfeld et al. (2018) who find that incentives shift gaze behavior and that participants are good at both pointing to an action and disguising an action).[13] In general, we expect to find evidence that observers are able to correctly interpret non-strategic gaze patterns, but not strategic gaze patterns.

1.2.4. Procedure

We use Tobii EyeX eye-trackers (60Hz frequency, with 1920×1080 pixel resolution monitors and chin-rests at a 58 cm distance[14]) connected to z-Tree (Fischbacher, 2007) such that (real-time) gaze data can be displayed and integrated in the interaction. We use ORSEE (Greiner, 2015) for recruiting student participants at the Lakelab in Konstanz, Germany.

In the PILOT experiment, the participants in the recording sessions ($n = 46$ allocators[15]), the INFOPILOT ($n = 56$ observers) and the PARTNERPILOT ($n = 54$ observers) sessions are fairly similar—average age of 21.3, 21.3 and 21.9 years, at least half (56.3%, 50.0%, 59.3%) female. In terms of the SVO orientation, there were no altruistic or competitive types in either of the sessions, and at least half of both the allocators (50.0%) and the observers (62.5% and 66.7%) in all sessions were prosocial.

In the MAIN experiment, the participants in the RECORDED ($n = 94$ observers) and INTERACTIVE ($n = 48$ allocators and $n = 48$ observers) sessions, respectively, also are fairly similar—average age of 22.4 and 21.5 years, majority (56.4% and 54.2%) female. Also in terms of the SVO orientation, there were again no altruistic or competitive types in either of the sessions, and the majority of both the allocators (62.5%) and the observers (58.5% and 58.3%) were prosocial in all sessions.

[12]Namely, using the measures proposed by Fiedler et al. (2013), we hypothesize that they allocate more attention to the other's payoff and make more comparisons that include the observers' payoffs, as well as spend more time and inspect more information during the first (strategic) decision.

[13]We define correctness of the partner choice here as the likelihood of identifying the allocator who allocates higher payoffs to the observer. We define correctness of the predictions as the likelihood of correctly guessing the chosen option and the chosen side from the "average choice" line of the allocator.

[14]In accordance to Gibaldi et al. (2017), Tobii EyeX is thus an appropriate apparatus for our parameters.

[15]Two of the initial 48 participants needed to be excluded due to technical and comprehension problems.

1.3. Allocator results

In this section, we firstly discuss the gathered allocator gaze data used in the INFOPILOT and PARTNERPILOT sessions of the PILOT experiment (see 1.3.1). Afterwards we turn to the effects of strategic incentives on them in the INTERACTIVE compared to RECORDED sessions in the MAIN experiment (see 1.3.2). Throughout this section, we focus on four measures[16] of allocator gaze patterns as proposed by Fiedler et al. (2013): (i) total decision time (seconds) per decision, (ii) share of inspected information (share of the total 10 AOIs seen at least once) per decision, (iii) share of the decision time spent gazing on own payoffs (percent) per decision, (iv) transition index[17] for self and other (index from -1 to +1) per decision.

1.3.1. Allocator results in the PILOT experiment

1.3.1.1. Non-strategic gaze and SVO

In this sub-section, we describe the non-strategic gaze patterns of the allocators, as shown in the INFOPILOT and PARTNERPILOT sessions. Note that we analyze the allocators' eye-tracking data in a format that exactly reflects how the observers saw the gaze patterns in the Information stage. We demonstrate that the allocators' gaze patterns can indeed be telling about their motives to the observers.

We examine the four measures in relation to the SVO angles (where higher SVO angle values indicate more pronounced prosociality), as depicted in Figure 1.7 and Table 1.2 below.[18] We can conclude that the more prosocial allocators inspect the AOIs significantly

[16] Fiedler et al. (2013) in addition focus on a fifth measure that is not suitable for our analysis due to the crude differences in how we define and transmit the allocators' fixations in our study. Namely, in our study, due to computational reasons, an AOI lights up immediately after an allocator's gaze enters the wider area, not after it fixates on an AOI.

[17] The transition index (TI) is defined as search direction

$$TI = \frac{\frac{trans_{other}}{k_{other}} - \frac{trans_{self}}{k_{self}}}{\frac{trans_{other}}{k_{other}} + \frac{trans_{self}}{k_{self}}}$$

where values above zero indicate that the allocator prefers comparisons involving the payoffs of the observer, whereas values below zero indicate a preference of transitions only involving own payoffs. The variables are defined as follows: $trans_{other}$ indicate the number of transitions between AOIs, for which at least one AOI contains payoff information of the observer and $trans_{self}$ indicates the number of transitions between solely own payoff information. We correct for the fact that there are more possible transitions including $trans_{other}$ than $trans_{self}$ by using $k_{self} = 20$ and $k_{other} = 20 + 10 = 30$ to indicate the number of *possible* (non-diagonal) transactions involving solely own or also other payoffs, respectively (Fiedler et al., 2013). Note that our results are robust to omitting the transitions that correspond to the shortest fixations of less than 50 ms.

[18] The significance levels in Figure 1.7 correspond to those of the coefficients in Table 1.2. The regressions include 46 recorded allocators in 24 allocation situations, but some observations lack either transitions including the observer or attention on the observer.

more carefully and spend significantly more time making the decisions. They also spend proportionally more time inspecting the observers' payoffs and have a higher transition index. We thus largely replicate the results of Fiedler et al. (2013): The differences in the SVO angles are reflected in the patterns of information search in social decision making.

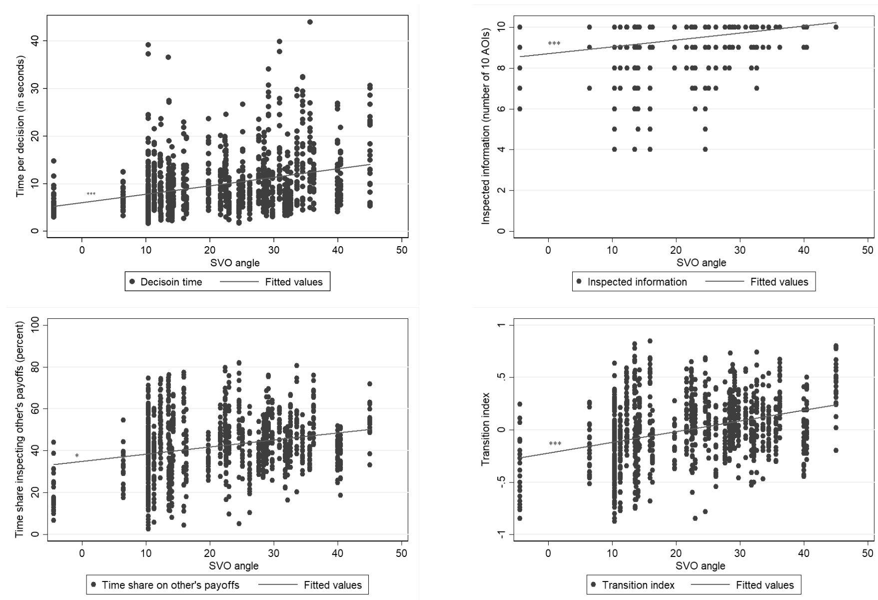


Fig. 1.7. Relation between the allocators' SVO and their shares of inspected information (top left), decision times (top right), attention to other's payoffs (bottom left) and transition index (bottom right) in the recording sessions.

	Log decision time	Log inspected AOIs	Time share percent	Transition index
SVO angle	0.018***	0.004***	0.338*	0.010***
	(-0.004)	(-0.001)	(-0.129)	(-0.002)
Period	-0.021***	-0.002**	-0.039	-0.002
	(-0.002)	(-0.001)	(-0.058)	(-0.001)
Constant	2.022***	2.175***	35.522***	-0.198**
	(-0.092)	(-0.030)	(-3.545)	(-0.064)
R^2	0.183	0.108	0.063	0.129
N	1104	1104	1090	1087

Table 1.2: Results of regression models for the allocators' gaze patterns. Unstandardized regression weights, robust standard errors clustered at participant level in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Furthermore, we estimate out-of-sample predictions of each participant's SVO angles using the four gaze-pattern measures in each situation to demonstrate that these measures have significant predictive power.[19] Our best model that uses all four measures together indicates a medium-to-strong positive correlation between the actual and predicted SVO angles (Pearson's $\rho = 0.60, p < 0.001$).[20] The four other models that use each of the four gaze pattern measures separately indicate a weak-to-moderate Pearson's ρ of $0.32, 0.49, 0.21$ and 0.43 (all $p < 0.001$), respectively, for the share of inspected information, the decision time, the attention to other's payoffs and the transition index.

1.3.2. Allocator results in the MAIN experiment

1.3.2.1. Strategic gaze and SVO

We further compare the results of the strategic INTERACTIVE sessions with the results from the non-strategic RECORDED sessions. Overall, the allocators had three different types of choices: strategic (transmitted) and non-strategic (non-transmitted) from the INTERACTIVE sessions of the experiment and non-strategic from the RECORDED sessions of the experiment. We expect strategic gazing in the strategic decisions but not in the two different non-strategic decisions.

In Figure 1.8 below, we again consider the same four gaze pattern measures as above, but this time we separate whether GazeVideo or Choice was transmitted and divide the allocators into prosocial and individualistic using the SVO angle of 22.45° as the threshold, in line with Murphy et al. (2011). Note that we can consider within-participant comparisons, as we use the same allocators in the RECORDED and INTERACTIVE sessions.[21]

In line with our hypotheses, we observe mimicry or "peacock" effects in strategic compared to non-strategic decisions. Namely, both the prosocial and the individualistic allocators attempt to appear more prosocial (in terms of the four gaze measures) in the strategic decisions, although the individualistic allocators exhibit this behavior more pronouncedly. In particular, the prosocial allocators have more "prosocial-like" gaze patterns (elaborated in the previous sub-section) in strategic compared to non-strategic decisions in 4 of the 8 sub-figures of Figure 1.8 below, while the individualistic participants have this tendency in all 8 of the sub-figures of Figure 1.8 below.[22]

[19]We estimate the parameters using all subjects except one (two for the partner choices) and use the estimates to predict the angle for the remaining subject(s).

[20]This corresponds to 61% of allocators being classified in the *same* categories—namely prosocial or individualistic—according to their actual SVO angles and their estimated SVO angles (significantly different from chance level, Wilcoxon signed-rank test, $p < 0.001$). See further allocator results in Appendix C.

[21]As mentioned, we record the allocators in non-strategic RECORDED situations in an introduction part preceding the strategic INTERACTIVE situations. Hence, we can consider the *same* allocators in RECORDED and INTERACTIVE settings.

[22]Note that the significance levels correspond to Wilcoxon signed rank tests (differences between strategic and non-strategic decisions) and Mann-Whitney-Wilcoxon rank sum tests (differences between individualistic

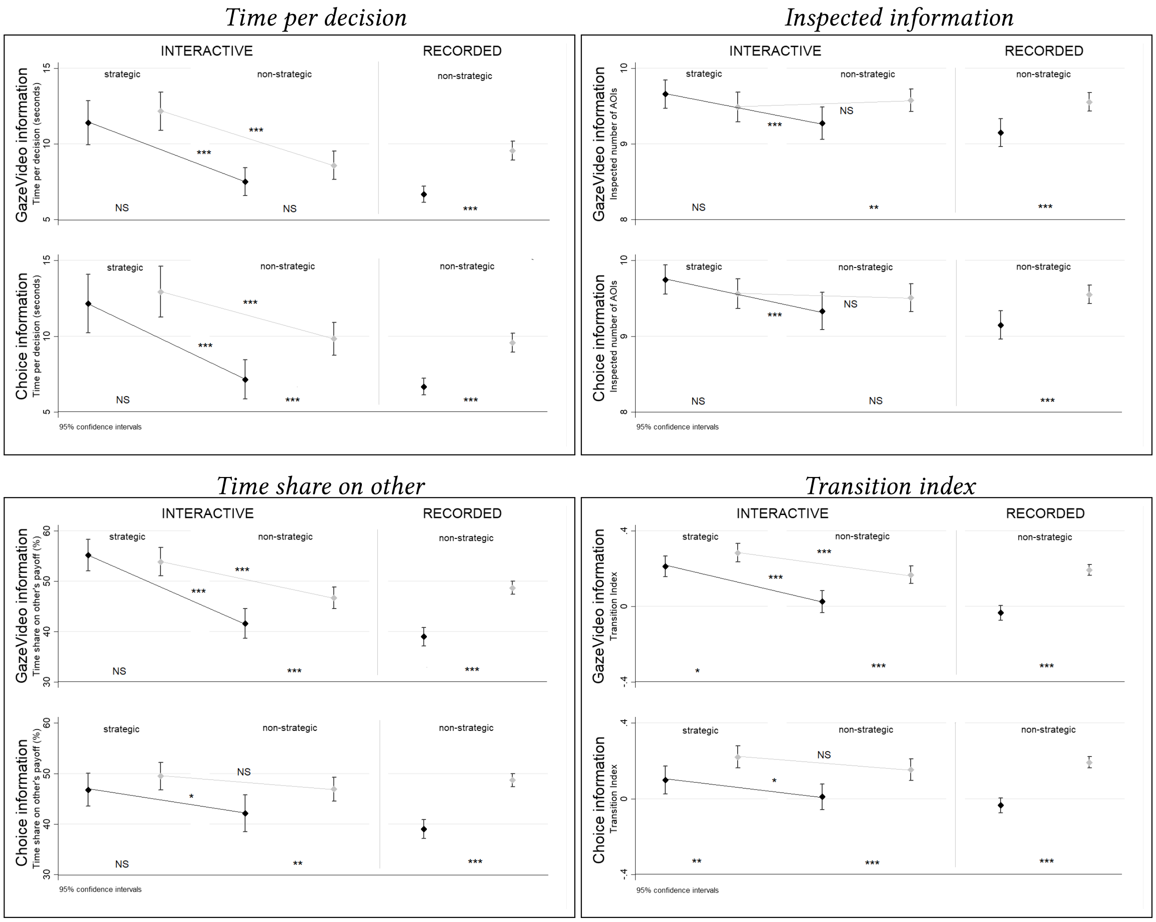


Fig. 1.8. The four measures separated into prosocial (gray) and individualistic (black) allocators at the 22.45° SVO angle; decisions in the RECORDED sessions (right side of the sub-figures and INTERACTIVE sessions (mid- and left side of the sub-figures), separated by whether the decisions were strategic and whether GazeVideo (top) or Choice (bottom) was transmitted to the observers.

In the *strategic* INTERACTIVE decisions, both prosocial and individualistic allocators thus put effort into appearing to care more about the situations and the observers' payoffs. But both prosocial and individualistic allocators return close to the baseline (non-strategic RECORDED on the right side) in the *non-strategic* INTERACTIVE decisions.

Moreover, the allocators' statements in the post-experiment questionnaire of the INTERACTIVE sessions further support the strategic gazing result. We ask the allocators the following question: Did you have a particular strategy for gazing when your gaze was transmitted? The provided answers often (for 48% or 23 out of 48 allocators) include some strategic gazing.[23] For example, "Yes, I focused my gaze on the best option for the observer",

and prosocial participants).

[23]Furthermore, the share is even higher when considering a subset of the individualistic allocators alone: 56% of the individualistic allocators specifically mention strategic gazing. Note, however, that the strategy

"I sometimes tried to mislead using my gaze", or "Yes, I looked at a positive option for the partner and took the best option for myself afterwards", among others.

1.3.2.2. Strategic choices and SVO

Up to now, we described support for strategic gaze when the eye movements were transmitted in the GazeVideo format. Note, however, that we also find strategic choices when the action information was transmitted in the Choice format. The actual choices underlying the discussed strategic and non-strategic gaze patterns are summarized in Figure 1.9 below. We find that the allocators act strategically and choose particularly equal allocations between themselves and the observers when only Choice information is transmitted, while the allocations are comparatively more individualistic when the GazeVideo information is transmitted. The allocations are most individualistic when neither Choice, nor GazeVideo information is transmitted. However, the prosocial allocators do not change their behavior in response to the strategic incentives or treatments as much as the individualistic allocators.

Allocated payoffs to self and other

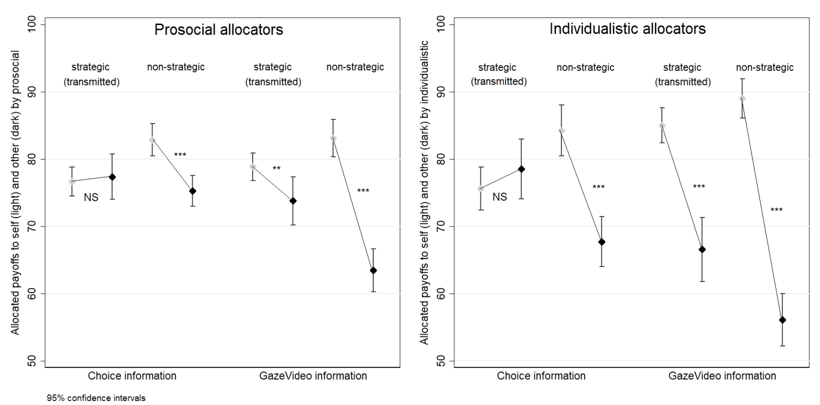


Fig. 1.9. Allocated payoffs (self gray, other black) in the INTERACTIVE decisions by prosocial allocators (left) and individualistic allocators (right), separately by whether the decision was strategic (in Choice or GazeVideo format, same situations) or non-strategic.

1.4. Observer results: Predictions

In this section, we firstly discuss the observers' prediction results in the INFOPILOT and PARTNERPILOT sessions of the PILOT experiment (see 1.4.1). Afterwards we turn to the effects of the strategic incentives in the INTERACTIVE compared to RECORDED sessions of the MAIN experiment (see 1.4.2).

statements were not incentivised and the question was phrased very generally.

1.4.1. Observer prediction results in the PILOT experiment

Throughout this section, we consider two measures for prediction correctness, in line with the applied two-step incentive procedure. As depicted in Table 1.3 below, we examine (i) side predictions (correctly guessing on which side of the average choice was the chosen option) and (ii) option predictions (correctly guessing the allocator's exact chosen option). This two-step procedure was applied because we expected that the observers might be able to predict the correct direction from the average choice but not necessarily the exact correct option.

			Side predictions			Option predictions		
Sessions	N	Periods	Choice	Video	Picture	Choice	Video	Picture
INFOPILOT	56	12×3	70.9%	65.3%	66.0%	59.5%	53.0%	53.5%
Empirical chance			*57.7%*	*57.7%*	*57.7%*	*55.4%*	*55.4%*	*55.4%*
PARTNERPILOT	54	12×3	65.4%	61.0%	-	53.0%	48.8%	-
Empirical chance			*56.0%*	*56.0%*	-	*47.0%*	*47.0%*	-
RECORDED	94	4	62.5%	59.4%	-	48.0%	45.1%	-
INTERACTIVE	48	4	57.3%	57.3%	-	44.5%	43.8%	-

Table 1.3: Summary of prediction correctness in the PILOT and MAIN experiments. "N" indicates the total number of observers, "Periods" indicate the number of periods per information type (three information types in the INFOPILOT sessions, two information types in all other sessions).

In general, considering all observers and all situations together, it appears that Choice information outperforms GazePicture and GazeVideo information in the INFOPILOT sessions. All three information formats consistently lead to significantly better prediction correctness than chance level in Wilcoxon signed-rank tests.[24] This holds both for the side predictions (one of two sides from the average allocator: median empirical chance level at 57.7%, $p < 0.001$; naïve chance level at 50.0%, $p < 0.001$) and partly for the option predictions (one of five allocations: median empirical chance level at 55.4%, $p = 0.962$; naïve chance level at 20.0%, $p < 0.001$).

Note that in all other sessions, we use only the GazeVideo format for motive information, as it appears to outperform the GazePicture format (see elaborated in Appendix D). In the PARTNERPILOT sessions, we observe similar trends in the prediction correctness to those in the INFOPILOT sessions, as depicted in Table 1.3 above—again for the side predictions (empirical chance level median at 56.0%, $p < 0.001$; naïve chance level at 50.0%, $p < 0.001$)

[24] The naïve chance levels are at 0.20 or 0.50. The empirical chance levels are higher in accordance to $\sum p_i^2$, where p denotes the empirical option or side probabilities for each situation, respectively.

and the option predictions (empirical chance level median at 47.0%, $p < 0.001$; naïve chance level at 20.0%, $p < 0.001$). We thus also conclude that the possibility to compare two allocators in the same situation does not improve the observers' prediction correctness.

1.4.2. Observer prediction results in the MAIN experiment

In the MAIN experiment, however, the allocators in the INTERACTIVE sessions (but not in the RECORDED sessions, and not in the PILOT experiment) tend to gaze and choose strategically. In Table 1.3 above, we also present the prediction correctness in the RECORDED compared to INTERACTIVE sessions. In line with our hypotheses, the predictions are more correct in the RECORDED sessions than in the INTERACTIVE sessions.

1.5. Observer results: Partner choice

In this section, we firstly discuss the observer partner choice results (in terms of both prosociality and generosity) in the INFOPILOT and PARTNERPILOT sessions of the PILOT experiment (see 1.5.1). Afterwards we turn to the effects of strategic incentives in the INTERACTIVE compared to RECORDED sessions of the MAIN experiment (see 1.5.2).

1.5.1. Observer partner choice results in the PILOT experiment

Throughout this sub-section, we consider two measures for partner choice correctness: (i) choosing the more prosocial partner in general and (ii) choosing the more generous partner in terms of the allocated payoffs in a next decision. To evaluate the optimality of the observers' decisions, we define partner choice as the average picked partner number (between 1 and 2) and compare it to whether the 1^{st} or the 2^{nd} partner is the "correct choice".

The first measure for partner choice correctness is *prosociality* of the allocators. We find that the observers are able to correctly interpret non-strategic gaze patterns and choose the correct partners in both Choice and GazeVideo treatments, as firstly depicted in Figure 1.10 below (left sub-figure with Choice information, right sub-figure with GazeVideo information). Indeed, we can compare this partner choice result to an optimal partner choice of our model that uses out-of-sample SVO angle predictions (as described in the sub-section 1.3.2 above) and conclude that the observers perform only moderately worse than the benchmark set by our best-performing model.

Importantly, the GazeVideo information allows choosing the more prosocial partner *even if* the two partners made the exact same choice in the Information stage, as depicted in Figure 1.11 below. Here, the partner decisions are separated by whether the same choice (top sub-figures) or different choices (bottom sub-figures) were made by the two matched allocators. The result in the top right sub-figure where the two partners made the exact same choice in the Information stage indicates that gaze patterns can convey motive information beyond that of action information.

Partner choice and prosociality

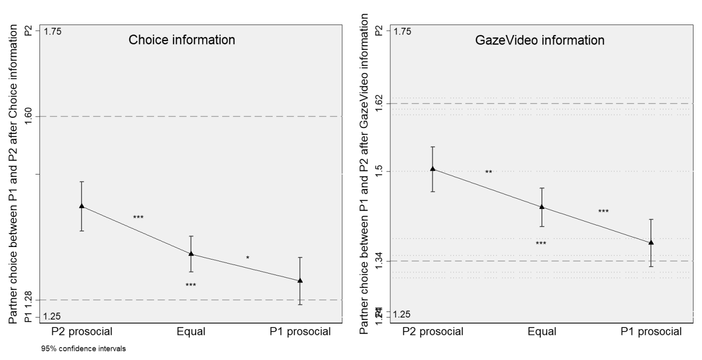


Fig. 1.10. Partner choice correctness (in terms of prosociality) in the non-strategic PARTNERPILOT sessions; dashed lines show the benchmark model with the optimal choices from out-of-sample SVO angle predictions.

Partner choice and prosociality

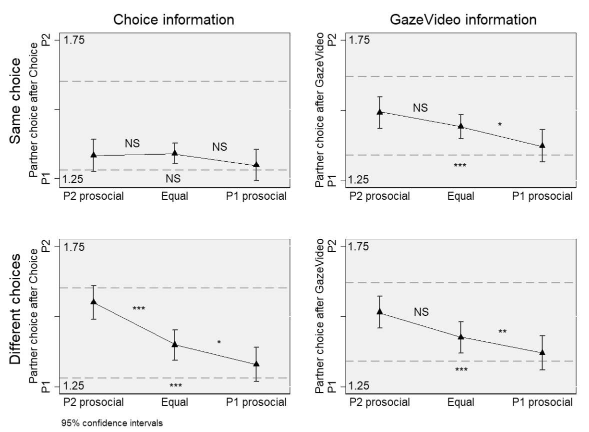


Fig. 1.11. Partner choice correctness (in terms of prosociality) in the non-strategic PARTNERPILOT sessions, separated by whether the two shown allocators made the same choice.

Partner choice and prosociality

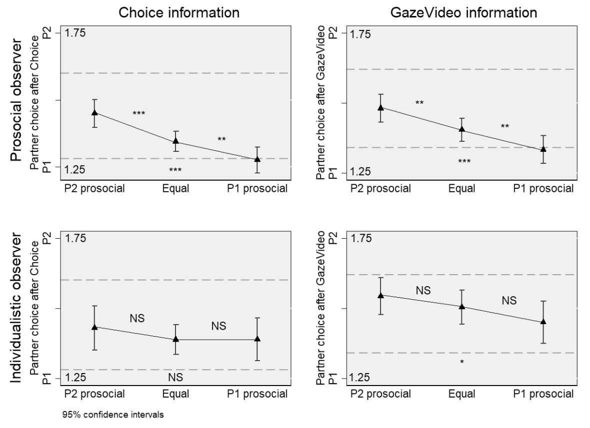


Fig. 1.12. Partner choice correctness (in terms of prosociality) in the non-strategic PARTNERPILOT sessions, separated by the observers' SVO.

Furthermore, if we additionally separate partner choice decisions of Figure 1.10 by the observers' SVO angles, we can conclude that the above results are mainly driven by the more prosocial observers rather than the individualistic observers, as depicted in Figure 1.12 above (top sub-figures compared to bottom sub-figures, respectively).[25]

The second measure for partner choice correctness is *generosity* of the allocated payoffs. In particular, we consider whether the observers choose the allocators who were more generous in the subsequent non-strategic decision, which is directly relevant for the observers' payoffs, as depicted in Figure 1.13 below. We conclude that the generosity result is similar to the prosociality result in Figure 1.10 above: The observers are able to correctly interpret non-strategic gaze patterns and choose the correct partners more often in both Choice and GazeVideo treatments.

If we again separate the partner choice decisions of Figure 1.13 by the observers' SVO angles, we can conclude that also these generosity results are mainly driven by the more prosocial observers rather than the individualistic observers, as depicted in Figure 1.14 below (top sub-figures compared to bottom sub-figures, respectively). Note that this result is similar to the prosociality result in Figure 1.12 above.

[25]Please refer to Appendix E for further partner choice results.

Partner choice and generosity in the next decision

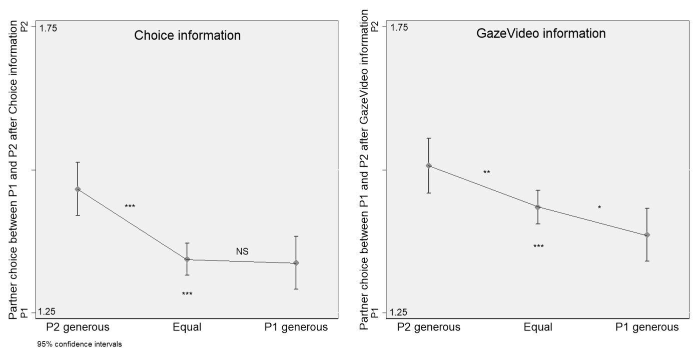


Fig. 1.13. Partner choice correctness (in terms of payoffs) in the non-strategic PARTNERPILOT sessions.

Partner choice and generosity in the next decision

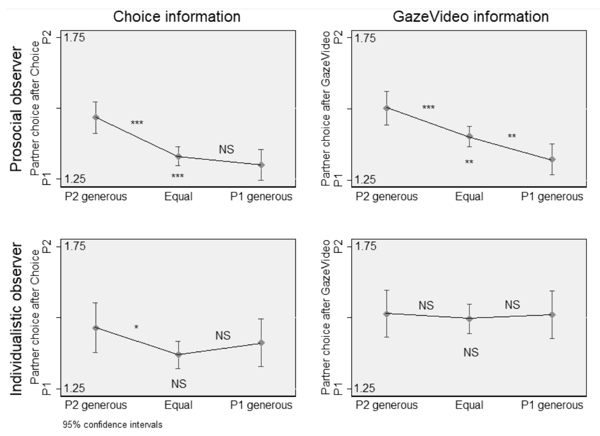


Fig. 1.14. Partner choice correctness (in terms of payoffs) in the non-strategic PARTNERPILOT sessions, separated by the observers' SVO.

1.5.2. *Observer partner choice results in the MAIN experiment*

Similarly, in the MAIN experiment, we again ask the observers to pick one of the two transmitted (strategic) allocators with which they would like to interact in a next (non-strategic) situation. As already demonstrated, the allocators in the INTERACTIVE sessions (but not the RECORDED sessions) tend to gaze and choose strategically in the transmitted decisions. To examine the effects of the introduced strategic incentives, we again define the optimality of the observers' partner choice in terms of both the prosociality of the chosen allocator and the generosity of the allocated payoffs, considering the payoffs from the non-strategic decision, which is relevant for the observers' payoffs.

We conclude that partner choices are significantly more correct in the non-strategic RECORDED situations than in the strategic INTERACTIVE situations and result in significant material consequences for the observers. As depicted in Figures 1.15 and 1.16 below, the observers in the non-strategic RECORDED sessions (left sub-figures) do indeed pick the more prosocial and generous allocators significantly more often.[26] However, in the strategic INTERACTIVE sessions (right sub-figures), the observers are unable to pick a partner optimally.

Partner choice and prosociality
in non-strategic (left) and strategic (right) situations

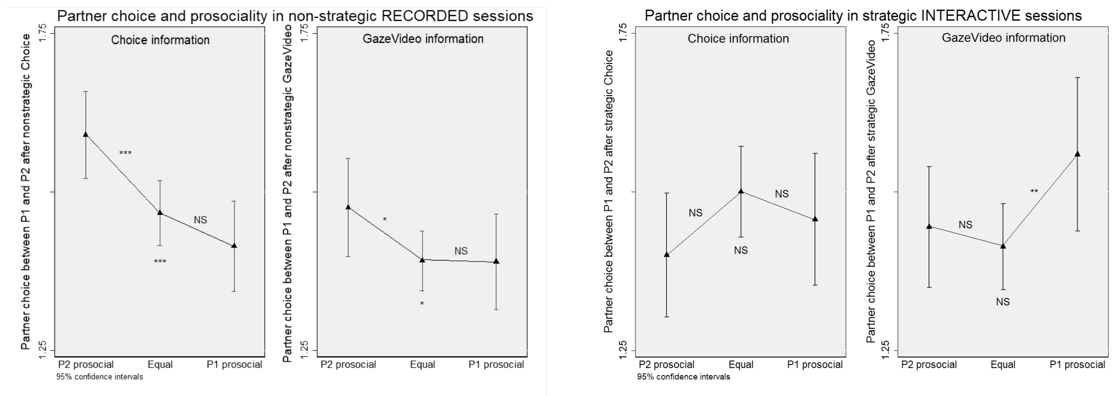


Fig. 1.15. Partner choice correctness (in terms of prosociality) in the RECORDED sessions (left sub-figures) and the INTERACTIVE sessions (right sub-figures).

Partner choice and generosity in the next decision
in non-strategic (left) and strategic (right) situations

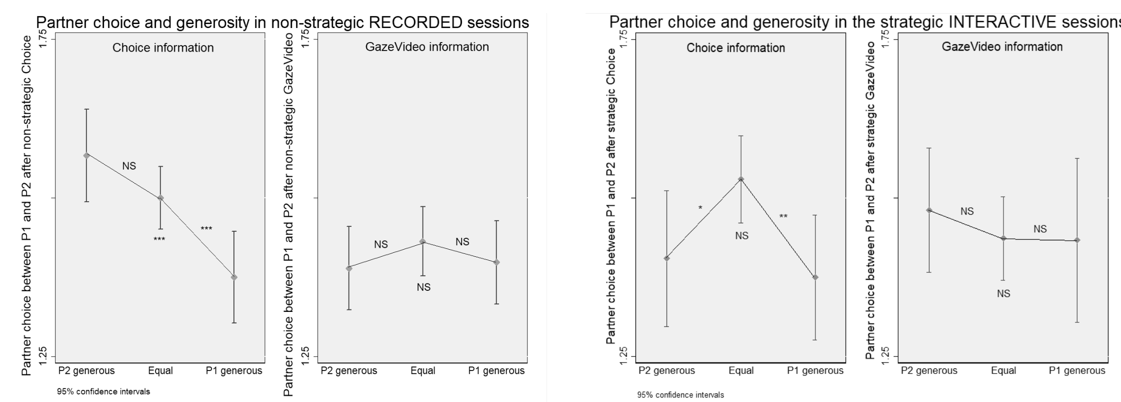


Fig. 1.16. Partner choice correctness (in terms of payoffs) in the RECORDED sessions (left sub-figures) and the INTERACTIVE sessions (right sub-figures).

[26]Compare Figure 1.16 to equivalent analysis of the payoffs from the decision transmitted in the Information stage in the INTERACTIVE compared to RECORDED decisions (as well as PARTNERPILOT decisions) in Appendix E. For the correct partner choice, the Information stage payoffs constitute a less noisy indicator of the observers' comprehension of the allocators' motives than the Prediction stage payoffs, as the slopes appear to be steeper in the Information stage in Appendix E compared to the Prediction stage depicted here.

In particular, the observers make the optimal partner choice significantly less often on average in the strategic INTERACTIVE sessions than the non-strategic RECORDED sessions: 58.6% compared to 53.8% correct choices in terms of generosity (excluding all cases where both allocators in a given decision situation allocate the same payoffs). This difference translates into an average payoff decrease of 5.8 points or 8.0% per decision for the observers (Mann-Whitney-Wilcoxon rank sum test, $p = 0.015$) and an average payoff increase of 7.4 points or 9.6% per decision for the allocators ($p < 0.001$, excluding the payoff multipliers). That is, if the allocators have incentives to gaze or choose strategically, the observers are less likely to choose the more generous partner for future interaction. In line with our hypotheses, we observe consistently less optimal partner choices in the INTERACTIVE sessions than in the RECORDED sessions.

1.6. Discussion and conclusions

In this study, we use interactive eye-tracking to examine social decisions in non-strategic compared to strategic settings. Using benchmark models for out-of-sample SVO predictions, we establish that gaze patterns can help to predict levels of prosociality. We then find that observers are capable of such type recognition in non-strategic settings, leading to similar performance as our best-performing benchmark model. That is, the observers can use the allocators' gaze patterns to identify the more prosocial and generous allocators, even if the allocators' actual choices are unknown, and this ability has a material impact on future interactions.

In the non-strategic settings, the allocators convey their type truthfully, albeit noisily. In contrast, in the strategic settings, the allocators are able to strategically manipulate their gaze and choices, as to appear more generous and affect the observer's preferences in favor of these allocators. We thus observe consistently less optimal actions by the observers in the strategic settings than in the non-strategic settings, leading to significantly worse payoffs for them.

Interestingly, we find that non-strategic gaze patterns are revealing even in situations, in which the underlying actual choices of the allocators are the same. Although it might seem trivial at first, note that the Choice treatment included the choice *and* the decision times. It is thus purely the gaze that helps the observers to identify the prosocial allocators.

In general, our findings reveal that it is an easier task to identify which allocator is more generous than to predict which option the allocator chose. Note, however, that especially the prosocial observers tend to choose the more generous allocators. It could be due to the prosocial participants identifying their own types better, although this reasoning should then also hold for individualistic participants, who should avoid choosing their own type. This result presents a potential direction for future research, namely examining the heterogeneity in how different types learn from others' motives and actions.

Interactive eye-tracking is quite a new research field in economics, and this study in combination with Hausfeld et al. (2018) uncovers numerous further questions that still need to be answered, for example, about the types of information that can be conveyed using interactive eye-tracking and about recognition of the conveyed information's quality or truthfulness.

Appendix A. Investment situations

Note that Figure A1.1 below illustrates the situations in Table A1.1 below graphically.[27]

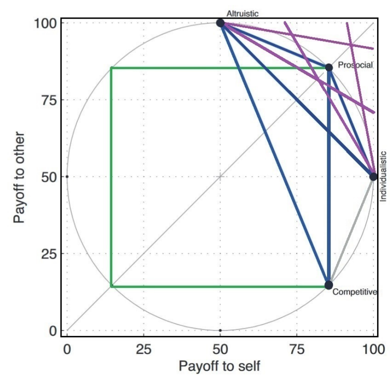


Fig. A1.1. Graphic illustration of the twelve different choice situations (selected SVO primary items in blue, selected SVO secondary items in purple, additional custom-created items in green).

Primary items

You receive	85	85	85	85	85
Other receives	85	68	50	33	15
You receive	100	96	93	89	85
Other receives	50	59	68	76	85
You receive	50	59	68	76	85
Other receives	100	96	93	89	85
You receive	50	59	68	76	85
Other receives	100	79	58	36	15
You receive	100	88	75	69	50
Other receives	50	63	75	88	100

Secondary items

You receive	100	93	85	78	70
Other receives	50	63	75	88	100
You receive	100	88	75	63	50
Other receives	70	78	85	93	100
You receive	50	63	75	88	100
Other receives	100	98	95	93	90
You receive	90	93	95	98	100
Other receives	100	88	75	63	50

Additional items

You receive	15	15	15	15	15
Other receives	85	68	50	33	15
You receive	85	68	50	33	15
Other receives	85	85	85	85	85
You receive	85	68	50	33	15
Other receives	15	15	15	15	15

Table A1.1: The twelve different choice situations (in monetary units).

[27] The full dataset comprised the twelve situations two times, once as depicted and once with jittered values (only in the PILOT experiment). In addition, a thirteenth situation between "Competitive" and "Individualistic" points (the final SVO primary item, by Murphy et al., 2011) was added in the recording stages for SVO angle calculation.

Appendix B. Decision screens

The main components of the observers' decision screens, as depicted in Figures A1.2 to A1.4 below, include (i) predictions of the allocators' choices and prediction certainty elicitation (all treatments) and (ii) partner choices for further interaction (all treatments except INFOPILOT).

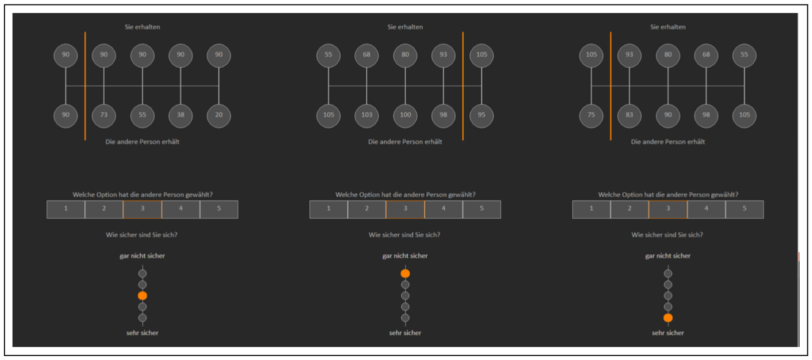


Fig. A1.2. Decision screen as seen by all observers in the Assessment stages of the non-strategic INFOPILOT sessions.

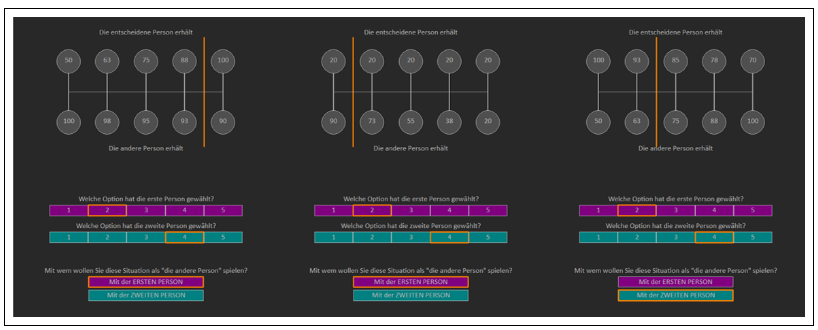


Fig. A1.3. Decision screen as seen by all observers in the Assessment stages of the non-strategic PARTNER-PILOT sessions.

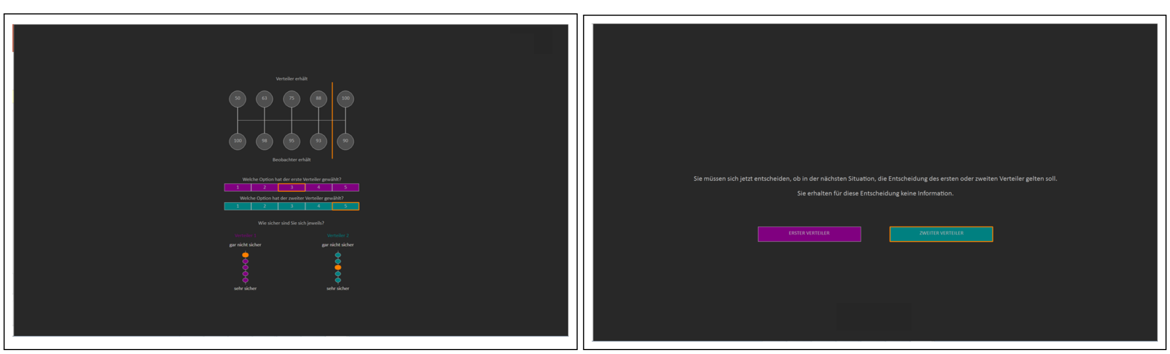


Fig. A1.4. Decision screens as seen by all observers in the Assessment stages of both the strategic INTERACTIVE and non-strategic RECORDED sessions.

Appendix C. Further allocator results

As depicted in Figure A1.5 below, almost all decisions were made after looking at all ten AOIs (top left sub-figure). The decisions were not too quick either (median decision time at 8.3 seconds; top right sub-figure). The allocators considered both their own payoffs and the observer's payoffs, although usually own payoffs more carefully (majority below 50%; bottom left sub-figure). Finally, the allocators' search direction was quite balanced (bottom right sub-figure). We conclude that the allocators in our sample appear to be quite attentive.

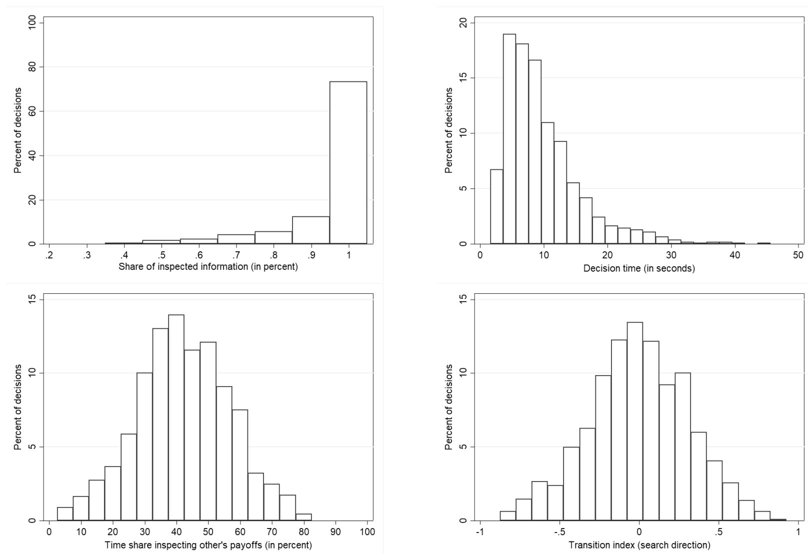


Fig. A1.5. Descriptive results including the allocators' shares of inspected information, decision times, attention to other's payoffs and transition index in the recording sessions.

Appendix D. Information treatments

As depicted in Figure A1.6 below, we compare the GazePicture and GazeVideo formats by examining self-reported prediction certainty and post-experiment evaluations. We conclude that the observers slightly prefer GazeVideo over GazePicture information (Mann-Whitney-Wilcoxon rank sum tests, $p < 0.1$). We thus focus on the GazeVideo information format in the PARTNERPILOT, INTERACTIVE and RECORDED sessions.

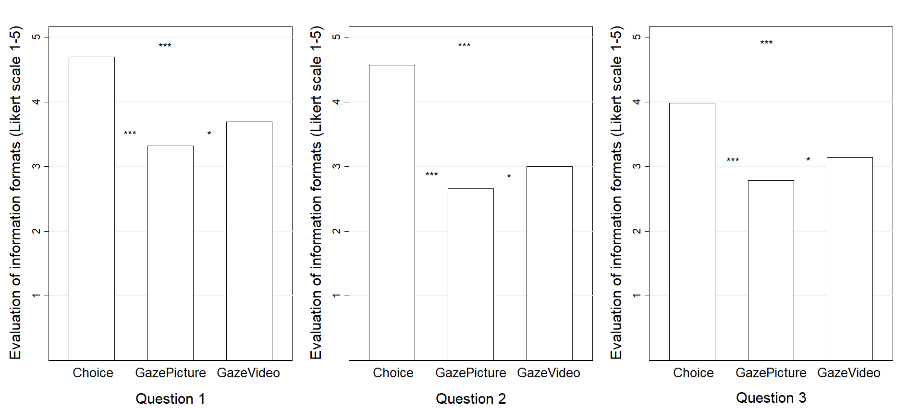


Fig. A1.6. Self-reported evaluations of the information treatments in the INFOPILOT sessions, with the three evaluation questions as follows. Question 1: How well did you understand the information that you saw? Question 2: How well could you tell from the seen information which option the other person chose? Question 3: How well did the seen information help you predict which option the other person would choose in another situation?

Appendix E. Further observer results: Partner choice

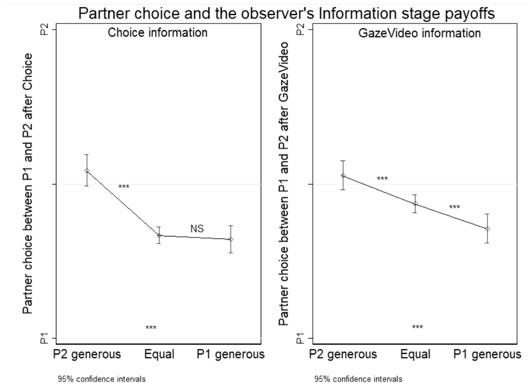


Fig. A1.7. Partner choice correctness (in terms of Information stage payoffs) in the non-strategic PARTNER-PILOT sessions.

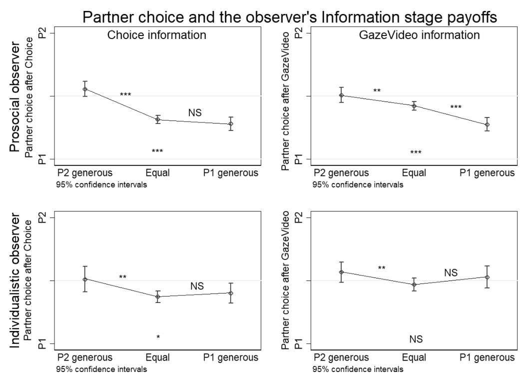


Fig. A1.8. Partner choice correctness (in terms of Information stage payoffs) in the non-strategic PARTNER-PILOT sessions, separated by the observers' SVO.

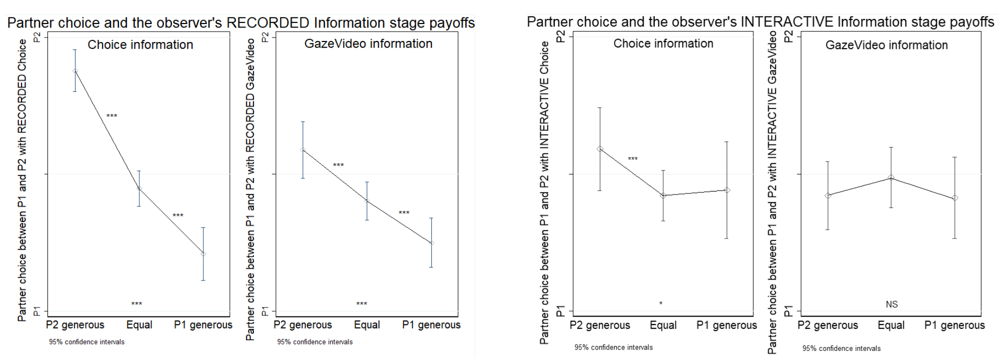


Fig. A1.9. Partner choice correctness (in terms of Information stage payoffs) in the non-strategic RECORDED sessions (left sub-figures) and the strategic INTERACTIVE sessions (right sub-figures).

Appendix F. Instructions

The following instructions (translated from German) correspond to the INTERACTIVE treatment of the MAIN experiment that includes both allocators and observers in the same session. It is, in a sense, the fullest version of instructions from all four session designs, and the other three sessions can be seen as partial versions of these instructions. Namely, (i) the RECORDED sessions in the MAIN experiment exclude the instructions for allocators and only include equivalent instructions for the observers—except that the allocators' decisions are not transmitted live but from a recording at an earlier session (thus effectively excluding the strategic component in the transmitted decisions). On top of this, (ii) the PARTNERPILOT sessions in the PILOT experiment additionally exclude the second observer in the same group, and (iii) the INFOPILOT sessions in the PILOT experiment additionally exclude the second allocator in the same group (and the partner choice stage, accordingly).

Instructions in the INTERACTIVE sessions of the MAIN experiment

---

Printed instructions 1 start here

General instructions
We warmly welcome you to this economic experiment. Your decisions and, if applicable, the decisions of other participants influence your payout in this experiment. It is therefore very important that you read these instructions carefully. **You are not allowed to communicate with other participants for the entire duration of the experiment.** We therefore ask you not to speak with each other. Please switch off your mobile phones and tablets as well. If you do not understand something, first take a look at the experiment instructions again. If you still have questions, please give us a sign by raising your hand. We will then come to you and answer your questions personally.

The experiment consists of 2 parts. Part 1 consists of 26 decisions, of which 1 decision is paid out. You can find the instructions for Part 1 on the screen. Part 2 consists of 9 rounds, of which the decisions of 1 round are paid out. The detailed instructions for this part will be provided later. After Part 2, a questionnaire will follow.

During the experiment, we are not talking about euro, but about points. Your total income is therefore calculated first in points. The total score achieved by you during the experiment is converted into euros at the end, such that 10 points = 1 euro. In addition to your income from the experiment, you will receive a 6 euro show-up fee.

Please note that this experiment is an eye-tracking experiment. The eye trackers must first be calibrated for this. The calibration takes place after reading the instructions. On the next page, you will see an explanation of how the calibration works. Please use the chin rest no later than when you can read that calibration now follows on your screen. After completing Part 1, you can remove your head from the chin rest. You will receive a corresponding message for this. After you will have read the instructions for Part 2 (you will receive these later), a calibration will follow again. Please click "Next" on your screen once you have read and understood these instructions.

Printed instructions 1 end here

Please click "Next" once you have read and understood the printed instructions.

Welcome to the Lakelab. Please note the following rules: 1. There is a ban on communication while you are in the lab. 2. If you have questions, please direct them to the experimenters. 3. You may only use the computer functions provided for the experiment. 4. Mobile phones must be switched off. Thank you for supporting our research. Please press "Next" (with the left mouse button) to continue.

Part 1
In this part, groups of two participants are formed. So, you are in a group with another participant. You do not know who that person is, and vice versa, that person does not know that they are in a group with you. This part consists of 26 decision situations. In each decision situation, there are 5 ways you can split points between yourself and the other participant in the group. You can choose one of the 5 options. At the end of the experiment, a random draw will determine which situation will be paid out and whether you or the other group member will decide on the allocation. The points in this experiment are converted as follows: 1 point = 0.10 euro. Furthermore, you will receive additional 6 euro for your participation. After the instructions, control questions will follow. Until then, you can always click "Back". Please click "Next" to continue.

Decision on the computer 1
The decision will be shown at the bottom of the page. Each option consists of a circle at the top and a circle below, which are connected with a grey line. The upper number indicates how many points you receive. The bottom number shows the points the other person in your group receives. For example, in the fourth option from the left, you get 90 points and the other group member gets 50 points. You select an option by typing 1, 2, 3, 4 or 5. When you have done so, an orange frame appears around the selected option. As a suggestion, please use the keys 1, 2, 3, 4 and 5, which are located above Q W E R T. You can try this out now. If you press "1", a frame around the leftmost option appears, etc. Click "Next" to continue.

Decision on the computer 2
During the first part, you have to confirm your choice with the spacebar. This means that you first have to select the respective option with 1, 2, 3, 4 or 5, and then an orange-colored frame appears. Then you have to confirm your choice with the spacebar. During the first part, you no longer need the mouse. You can place your fingers on the 5 keys and the spacebar. You must also press the spacebar once before each decision. Then a cross appears on the screen for about 1 second. Only then the decision situation follows. If you click on "control questions", you will get to the control questions and cannot read the instructions anymore.

Control questions
1. Please select the option that earns you 60 points. [...] 2. Please select the option that earns 70 points to the other person. [...] 3. Please choose the option that gives you 80 points and 60 points for the other person. [...]

That was correct. It will continue soon. [or] That was wrong, click the spacebar to get back to the question.

Now part 1 follows. As a reminder: There are 26 decision situations. You select an option with 1, 2, 3, 4 or 5 and then confirm with the spacebar. Please place your fingers on the 5 keys and the spacebar. You do not need the mouse now. Please click the spacebar to start the first part.

Please click the spacebar to start. Part 1 then begins. [...]

You are done with Part 1. The instructions for Part 2 will be distributed right away. Please wait. Please read the printed instructions. If you have any questions, let us know by raising your hand. Please click on the "w" button after reading and understanding the printed instructions.

Announcement*: We will now distribute the instructions for Part 2. Part 2 consists of 9 rounds. In general, there are always 2 observers and 2 allocators in a group, with the groups being shuffled before each round. If you have an eye tracker in front of you, you are an allocator. First, the allocators decide successively. Here, the observers either see the allocator's decision itself or a live video that shows where an allocator is looking. Then the observers decide with which allocator they choose to play in the next decision. The allocators then make that second decision and get the points as often as they have been chosen: twice, once or not at all. The observers get the points from the option picked by the chosen allocator. The observers also have to make predictions for both allocators. Everything else can be found in the instructions.*

Printed instructions 2 start here

Part 2

The second part of the experiment consists of a total of 9 rounds with several decisions per round. In each round, a new group of **two allocators and two observers** is formed. Each round runs as follows: **decision 1, information, choice, decision 2, prediction**.

Allocators: The allocators (people with eye trackers) make distribution decisions as in the first part, and they make **two different decisions in each round**. Decision 1 is always relevant for payment. Decision 2 counts for the allocators twice, once or not at all, depending on how many observers have chosen them after decision 1.

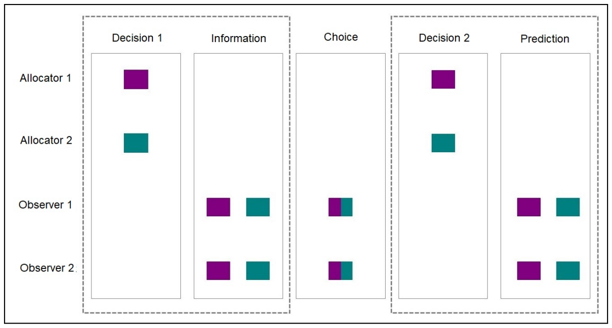
Observers: The observers (people without eye trackers) receive information (either the allocator's decision or the eye movement) about **decision 1** from two allocators and then select one of the allocators for decision 2. The **decision 2** of this allocator is then relevant for this observer. After this selection, the observers see a situation and must predict what both allocators will choose in decision 2.

Allocation and payment: In decision 1, the allocators and observers are randomly assigned to each other; each one allocator is assigned to one observer. In decision 2, the assignment is based on the choice of observers. Each observer is assigned to the selected allocator, but the allocators can be assigned to two, one or no observer. Accordingly, the decision for the allocators counts twice, once or not at all.

In Part 2, the following colors are used. The observer sees the decision for allocator 1 framed purple and for allocator 2 framed turquoise. The observer briefly sees a green frame at the time when the indicated allocator has decided. The allocator sees the decision situation framed in color when information (either green for the choice or orange for a live gaze video) is transmitted to the observers. The figure explains the sequence of rounds 1-8.

There are 9 rounds in total. After the first 4 rounds, the information type is changed, and after the 8^{th} round, the observers decide on the information type for the 9^{th} round. For each group in rounds 1-8, the following applies:

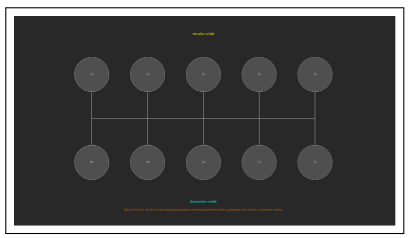
1) Each allocator is at most once in a group with the same allocator.
2) Each observer sees any allocator only once per type of information.



Decision 1

Like in Part 1, there are again 5 ways in which the allocators can divide points between themselves and an observer in each decision-making situation in Part 2. The decision 1 for the allocators is framed in color: green (CHOICE) or orange (GAZE VIDEO). A colored frame around the screen indicates to the allocators that information is transmitted to the observers.

In this case, both allocators and observers see the decision situation exactly as the allocator sees it. This means that the points for the observer are "Observer obtains" and the points for the allocator are "Allocator obtains".



The observers see the decision situation before the allocators decide and can familiarize themselves with the decision that the allocators will then make. The allocators do not see the situation until they decide.

Information

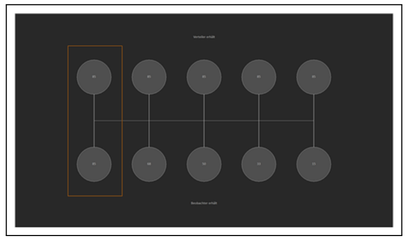
The observers first see decision 1 with the information from the first allocator. This information is framed purple for the observers. After information about the first allocator, the decision 1 follows with information about the second allocator. This is framed turquoise. Which allocator starts is randomly determined before each round. In each round, there are 2 observers and 2 allocators in a group. The following applies: Each allocator is at most once with the same other allocator in a group. Within an information type, the observers encounter the same allocator only once.

There are 2 ways in which the information is displayed: **choice** and **gaze video**. You will be informed before each round which type of information you will receive or share about the decision. The information type is the same in rounds 1-4, then changes in rounds 5-8. At the beginning of the 1^{st} and 5^{th} round, all participants see **3 examples**, which are the same between the 2 different types of information.

CHOICE

When the **choice** is displayed, the exact decision of the allocator is displayed. An orange frame appears around the option that the allocator has chosen. The observers will see the decision for as many seconds as the respective allocator required for the decision. The observers will see a green frame after the respective allocator has decided. After the observers have seen the information from the first allocator, the second allocator decides. Then also the information from the second allocator follows in the same way.

The observers see this information in 4 rounds. First, they see the information of the first allocator and then the information of the second allocator. The allocators transmit the information in these rounds. A green frame on the screen signals this to the allocators.

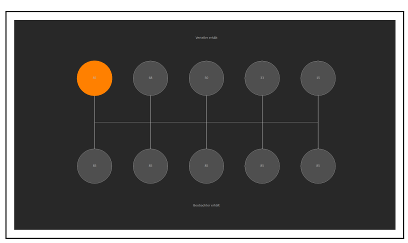


GAZE VIDEO

When the **gaze video** is displayed, a video of the eye movements of the allocator is displayed during the respective decision. Here, the viewing order during the decision is displayed. It depicts whether the gaze is on one of the ten payoffs (see picture). More specifically, an orange-filled circle indicates that a person is looking at this information or in the immediate vicinity. As soon as they look at one of these areas, it is marked. For example, if the person looks at the information "85" on "the other person receives", the circle fills orange for the respective time.

The observers see the decision and gaze behavior for as many seconds as the respective allocator needs for the decision. The observers will see a green-colored frame after the respective allocator has decided. After the observers have seen the information from the first allocator, the second allocator decides. Then the information from the second allocator follows in the same way.

The observers see this information in 4 rounds. First, they see the information of the first allocator and then the information of the second allocator. The allocators transmit the gaze video live in these rounds. An orange frame around the screen signals this to the allocator.



Partner choice

Having seen the information (choice or gaze video) of the two allocators in decision 1, the observers must decide whether decision 2 from the first or the second allocator should apply to them in the upcoming decision 2.

Decision 2

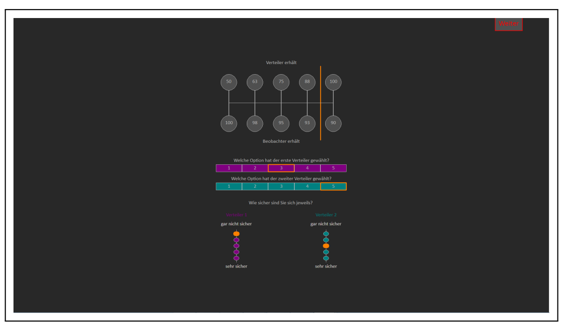
After decision 1, **decision 2** follows with a new randomly drawn decision situation. Here, the allocators decide without any information being passed on. Accordingly, there is no orange frame for the allocators, and the observers do not know about this decision.

For observers, this decision includes payoff from the allocator that was previously selected. For allocators, this decision is payoff-relevant as often as they were chosen by the observers after decision 1: twice, once or not at all.

Prediction

Regardless of their partner choice, the observers in the prediction phase have to predict the **decisions of the two allocators** for the new decision situation (options 1, 2, 3, 4 or 5). In addition, the observers receive additional information about how the whole group of participants decided in a very similar experiment in June. In this earlier experiment, there were only the decision situations and no transmission of information. An orange line divides the decisions in such a way that roughly as many June participants chose the decisions on the left side of the orange line as on the right side. For example, if the options were selected by 10%, 20%, 30%, 20% and 20% of the participants, then the orange line is after the third option; the options to the left of the orange line were chosen by 60% of the people and the options to the right of the orange line by 40% of the participants.

For these situations, the observers have to predict which of the 5 possibilities the two displayed allocators selected. In addition, the observers must also state how confident they feel about their predictions. In the made-up example below, we have indicated that an observer predicted the 3^{rd} option for the first allocator and the 5^{th} option for the second allocator. The observer was not at all sure about the first allocator, while the prediction for the second allocator was made with half-certainty. The information about certainty is not payout relevant.



After the prediction has been made, the observers continue using the "Next" button.

Decisions in round 9

Before round 9, the observers choose what information (CHOICE or GAZE VIDEO) is displayed in this round. After the observers have chosen, all participants learn how the observers in the group have chosen. There are 2 scenarios: either both observers choose the same information type, or the observers choose different information types. If both observers choose the same type, round 9 is equivalent to the rounds in the respective type of information. If the two observers select different types of information, the respective type is displayed to the respective observer. For the allocators, this means that one observer is shown CHOICE and the other observer is shown GAZE VIDEO.

Payout

At the end of part 2, one round will be randomly drawn and paid out. It will also be randomly determined whether the prediction for the first or the second allocator is relevant for payment. The payout consists of several parts.

Allocator: You will receive the points from the first decision, as well as the points from the second decision as often as you were elected: twice, once or not at all.

Observer: You will be assigned to one of both allocators in the first decision and will receive the respective points. In decision 2, you get the points from the option chosen by the allocator you have chosen.

Regardless of their choice of allocator, the observers receive points from their predictions. First, they get **30 points** if they have predicted the other participant's decision on the **correct side** (to the left or right of the displayed average decision). Furthermore, they get another **30 points** if they have also predicted the **correct option** (options 1, 2, 3, 4 and 5).

The points in this part of the experiment are converted again as follows: 1 point = 0.10 euro.

During Part 2, there will be a short summary of the respective section before the 1^{st} round and before the 5^{th} round.

Summary (observer)

In Part 2, you will first see the information of 2 different allocators in the format CHOICE or GAZE VIDEO. In each of the 4 rounds, you are in a group with different allocators. First, you will learn exactly what the decision situation will look like and you will familiarize yourself with the situation. When you have done this, you can continue using the spacebar.

In decision 1, you see the decision situation (first from the first allocator, then from the second allocator) with the additional information CHOICE or GAZE VIDEO for as long as the respective allocator needed to decide. You will receive the points from the option chosen by one of the allocators. Which allocator this is will be decided randomly.

After seeing the information from both allocators, you have to decide whose next decision should apply to you. This decision 2 is a different decision situation and you will receive the points from the option chosen by the allocator you have chosen.

Either way, you need to make predictions for both allocators about what they have chosen in decision 2 and indicate how confident you are of your prediction. The prediction can give you a total of 60 points. In the decisions you are to predict, you will see an orange line that marks the average decision of all participants from an earlier experiment (June 2017). You get 30 points if you indicate the other participant's option on the correct side (left or right of the average). You get another 30 points if you predict the exact option.

Summary (allocator)

In Part 2, you will make 2 decisions in each round. Furthermore, your information is viewed by two different observers in the format CHOICE (green screen frame) or GAZE VIDEO (orange screen frame). In each of the 4 rounds, you will transfer this information and you will play each round with a different allocator and different observers. Chance will determine whether you or the other allocator will decide first.

In decision 1, the observers see the decision situation (first from the first allocator, then from the second allocator) with the additional information CHOICE or GAZE VIDEO for as long as you or the other allocator needed to decide. Your decision is relevant for your payout, as well as for one of the assigned observers.

After the first decision, the observers can choose whether your decision or the decision of the other allocator should apply in decision 2.

After the observers have decided, you make decision 2. This decision is a new situation. No information is transmitted here. You will receive the selected points as follows: If both observers have chosen you, you will receive the selected points twice. If one observer has chosen you, you will receive the selected points once. If no observer has chosen you, you will not receive any points from this decision.

So, you get the points from decision 1, and you get the points from decision 2 (without transmission) either double, once or not at all.

Printed instructions 2 end here

Please wait until all participants have read the instructions. The first type of information transfer follows. [...] The second type of information transfer follows. [...]

You are an allocator and transmit this information. [or] You are an observer and see this information transmitted from the allocators.

In the next 4 rounds, the information type CHOICE [or] GAZE VIDEO applies. You will now see 3 examples of CHOICE [or] GAZE VIDEO from a previous experiment. Press the "w" button to continue to the examples.

[...] Please press the spacebar to continue. Then the second example follows. [...] Please press the spacebar to continue. Then the third example follows. [...] Please press the spacebar to continue.

Please click the "w" button to get to the summary for part 2.

Summary (observer) [or] Summary (allocator) [with the same text as in the printed instructions above] [...] Now rounds [x] to [x + 3] follow. In these 4 rounds, you will see the information as CHOICE [or] GAZE VIDEO.

Please wait until the experiment continues. An experimenter will come to you. [or] Please wait until the people are done with the eye-tracker calibration.

[Round 9]

You now have to decide if you want to see the information as CHOICE or GAZE VIDEO in the next round. [...]

[or] The observers have both chosen CHOICE. Accordingly, only CHOICE is transmitted. [or] The observers have both chosen GAZE VIDEO. Accordingly, only GAZE VIDEO is transmitted. [or] One observer opted for GAZE VIDEO and one observer opted for CHOICE. Accordingly, one CHOICE and one GAZE VIDEO is transmitted. Please press the spacebar to continue.

As a reminder, you are an allocator. The observer sees the screen exactly as you see it. Accordingly, their points will be displayed on Observer obtains. Round [x] now follows. The information is transmitted as CHOICE and has a green screen frame. [or] The information is transmitted as a GAZE VIDEO and has an orange screen frame.

[or] As a reminder, you are an observer. The allocator saw the screen exactly as you see it now. Round [x] now follows. You will now see the information CHOICE [or] GAZE VIDEO (re-live) from both allocators. Please press the spacebar to continue. [...]

Now decision 2 from round [x] follows. NO information will be disclosed. As a reminder: You will receive the selected points as often as you have just been chosen. [or] You will receive the selected points from the person just chosen and for your prediction from one of the allocators. Whether your prediction for the first or second allocator counts will be randomly drawn at the end. Please press the spacebar to continue.

You have completed round [x]. Round [x + 1] now follows. The information is transmitted in the same format (CHOICE [or] GAZE VIDEO). [or] The information changes format. [or] In the next round, the observers choose the type of information.

You are done with the experiment now. There are a few questions left. Please press the spacebar to continue.

[Questionnaire]
This experiment is also about the types of visualizations of gaze directions. Since this is the first experiment with gaze video, there are a few more questions left. Please click OK to continue.

Gaze video
Do you have comments on the visualization of the GAZE VIDEO? (For example: Was it understandable, was the time sufficient, were the colors too bright? ...) You can answer the question in keywords. [...] Please click OK to continue.

Comparison
How well did you understand the information you saw? CHOICE: [...] GAZE VIDEO: [...]
How well could you tell from the information shown what the other person chose in the shown situation? CHOICE: [...] GAZE VIDEO: [...]
How well did the information help you predict what the other person would do in a different situation? CHOICE: [...] GAZE VIDEO: [...]
Which of the two (CHOICE or GAZE VIDEO) information types did you find more informative? [...]
Do you have any further comments or suggestions for improvement? (You can answer in keywords) [...] Please click OK to continue.

Announcement*: We will now roll dice to determine all decisions or decisive rounds. Throw 1 determines which decision from Part 1 is paid out, throw 2 determines whether an odd or an even number corresponds to your decision (you can see this on the screen). Throw 3 determines which round from Part 2 is paid out; throw 4 determines, for the observers, which allocator's prediction counts in this round.*

[Random draws for payout]
Drawn situation from Part 1 (1-26): [...] Number determining who makes the decision in Part 1 (1-6): [...] Drawn information number from Part 2 (1-9): [...] Drawn allocator from Part 2 (1-2): [...]

Your income
Part 1: Your decision is relevant. [or] The decision of the other group member is relevant. The following situation was drawn: [...] You receive: [...] The other group member receives: [...]

Part 2: In part 2, the following round was drawn: [...] From the first decision, you receive: [...] Therefore, from the second decision in this round, you receive: [...] In the whole experiment, 1 point equals: [...] In addition, you will receive: [...] In total, you will receive: [...]

Chapter 2

Beliefs about others: Information neglect over type projection

Sebastian Fehrler, Baiba Renerte & Irenaeus Wolff

Abstract

In many games of imperfect information, standard theories assume that players can make Bayesian inferences about other players' types based on the information that is contained in their own type. Several behavioral theories of belief updating even start from the assumption that players project (or even over-project) their own type onto others. We investigate such inferences in a standard laboratory task, in which types (colored balls) are drawn from one out of several states of the world (colored urns). However, we find no evidence for over-inference. Instead, between 50% and 70% of the participants in our experiment neglect this important source of information entirely and base their choices only on the prior probabilities. Using several experimental interventions, we show that the phenomenon is very robust. Our findings are thus inconsistent with projection but are instead consistent with the other strand of the literature that documents the neglect of information.

JEL classification: C91, D83.

Keywords: projection, Bayesian inference, biased belief updating.

2.1. Introduction

We examine how people make inferences from their own type about other people's types. There are many situations in which such inferences play a role. Just to name two real-world examples that have been extensively discussed in the literature, inferences about others' types play a critical role in auctions with correlated valuations (Cremer and McLean, 1985, 1988; Brusco, 1998; Breitmoser, 2019) and elections with correlated political preferences (Goeree and Großer, 2007; Taylor and Yildirim, 2010; Agranov et al., 2017; Tolvanen, 2017).

In settings like the above, standard theory predicts that people make rational Bayesian inferences about the type distribution based on the information that is contained in their own type. Such inferences are standard in economic theories with fully rational agents. Several behavioral theories, such as social projection, information projection, or type projection (Dawes, 1989; Madarász, 2012, 2015; Breitmoser, 2019) even assume that people over-infer from their own type and thus over-project it onto others. This assumption is supported by a number of empirical studies (e.g. Messé and Sivacek, 1979; Offerman et al., 1996; Aksoy and Weesie, 2012; Blanco et al., 2014; Danz et al., 2018).

However, it is also in stark contrast with findings from other studies, in which inferences are not about another person's type. These studies show that people often neglect available sources of information (e.g., Friedman, 1998; Page, 1998; Hanna et al., 2014; Esponda and Vespa, 2014; Enke and Zimmermann, 2017). In light of this evidence, one might ask whether the projection idea is generally valid, as our interpretation of the projection literature suggests, or whether people might, at least in some situations, neglect the information contained in their own type when updating their beliefs.

In a laboratory experiment with three consecutively developed treatments, we make the players' types very salient and thus create favorable conditions for finding rational projection and over-projection. However, we find no evidence for over-projection of own types. Instead, we find behavior in the other direction, namely under-projection and under-inference, even when projection would be rational. We test the robustness of our results with a battery of interventions that nudge participants toward making use of the information available to them via their type. However, after conducting our intervention treatments, we conclude that the under-projection results are very robust. Our findings are not consistent with projection but are instead consistent with the other strand of the literature, which documents the neglect of information.

The basic set-up that we use to study projection is a central ingredient in many studies looking at imperfect-information setups (e.g., Guarnaschelli et al., 2000; Holt and Smith, 2009, 2016; Bouton et al., 2016; Fehrler and Hughes, 2018). There are two possible states of the world, a dark urn and a light urn. One of the states/urns is randomly drawn. Then, player A's type (a dark ball or a light ball) is drawn from the true state of the world. After replacement of A's ball, player B's type (ball) is drawn from the same urn. Player A does not get to see B's type. Instead, player A's task is then to guess B's type or to state the probability that she is of the same type. Between rounds, we vary the prior probabilities of the states of the world.

Many subjects react to this change in the prior but not at all to the information that is contained in their own type. When asked about the type probability in the easiest setting of a 50-50 prior for either urn, a majority of around 70%(!) of subjects state the prior probability, for example. The interventions that we employ include (i) belief-elicitation tasks (with respect to both the other's type and the state of the world) between the projection guesses to highlight the importance of belief updating, (ii) tasks without replacement to highlight the importance of replacement, (iii) tasks with several additional draws from the selected urn to stimulate learning, (iv) the use of physical urns and balls to avoid potential confusion about computerized processes, and (v) questions about the subject's reasoning behind previous choices to increase deliberation and facilitate the identification of reasoning errors. None of these interventions is very effective in reducing information neglect—only asking for beliefs slightly reduces it.

The self-described reasoning of participants suggests that many believe that nothing has changed when the ball is put back into the urn as compared to the situation before receiving the ball, and hence that the probability that the other person has the same type must be the prior probability. This resembles the "no change principle" that many people appear to be following in the more complicated Monty Hall problem—a classic setting, in which information neglect is the rule rather than the exception (Burns, 2017).[1]

2.2. Related literature and hypotheses

2.2.1. Projection

"Projection" can be defined as follows. A projecting person infers from their own type that another person shares their type with a probability that is higher than the *prior* probability.[2] This can be either a rational result of Bayesian updating—when one's type is indeed

[1] We describe and discuss the Monty Hall problem in sub-section 2.2.2.

[2] Note the similarities to the definition of the (truly-)false-consensus effect by Engelmann and Strobel (2012, p.680): "A (truly) false consensus effect is considered to be present if people, when forming expectations concerning other people's decisions, weight their own decision more heavily than that of a randomly selected person from the same population."

informative about the other player's type—or irrational. Our definition includes all sorts of projection, from social projection (Dawes, 1989) to information projection (Madarász, 2012), action projection (Al-Nowaihi and Dhami, 2015) and type projection (Breitmoser, 2019).

Social projection is an established concept in psychology, incorporating the assumption that people tend to project their own preferences, behavior and intentions onto others (Ross et al., 1977). A mathematical foundation for social projection was first suggested by Dawes in his rationalization of the false-consensus effect (Dawes, 1989). Today, false-consensus bias and social-projection bias often are used synonymously. Closely related to social projection, Madarász (2012, 2015) studies projection of information from an informed participant to an uninformed one. Another closely related concept is type projection (Breitmoser, 2019). It incorporates a social-projection bias similar to the one defined above and also relates to information projection.

Game-theoretically—though perhaps not psychologically—all bits of private information, preferences and informative signals can be subsumed under the term "type" and treated equivalently with the set of tools that are available for the analysis of games of imperfect information. Breitmoser (2019) sets up an elegant model of type projection in Bayesian games, where players can overestimate the probability that other people in their group share their type. He argues that people over-infer from their own type and project it onto others. He defines "type" in the context of games with two-sided imperfect information as a person's signal about the object value. A type-projecting agent believes that with some probability p, their opponents have the same type as herself, and with probability $1 - p$, their opponents have the types that Bayesian updating would indicate. The projecting agents then best-respond to the mixture of Bayesian and projected types.

There is evidence for projection both in non-strategic and strategic environments and both between individuals (Messé and Sivacek, 1979; Offerman et al., 1996; Aksoy and Weesie, 2012; Blanco et al., 2014; Danz et al., 2018) and within individuals but between different points in time (Gilbert et al., 1998; Read and Van Leeuwen, 1998; Conlin et al., 2007; Danz et al., 2015), which we interpret as suggesting that it might be a general phenomenon. In line with the literature on projection biases, we thus state the hypothesis that participants may over-infer and thus over-project their own types in our simple belief-updating tasks.

Projection (Null) Hypothesis: *Participants over-infer from their own types and thus over-project their own types onto others.*

However, a second strand of literature has revealed evidence that subjects often neglect available sources of information.

2.2.2. Information neglect

The other side of the story consists of theories that we subsume under the label "information-neglect theories". It includes different theories which all predict that people will fail to take into account one or another element of the information structure.

Most prominently, there is the work by Eyster and Rabin (2005, 2010), who argue that there are two ways in which people in strategic settings may err when inferring information—first, an extreme form of inferential neglect (cursed equilibrium) and, second, a form of inferential naïvety. Type projection is conceptually related to cursed equilibrium (Eyster and Rabin, 2005). Similar to the type-projection equilibrium, cursed equilibrium is also a solution concept for Bayesian games with subjective type distributions that differ from the objective one. The difference between cursed equilibrium and type-projection equilibrium is that a type-projecting agent projects his own type onto his opponents, while a cursed agent believes that she faces a random type that is drawn from the prior type distribution even at a later stage in the game when beliefs could have been updated. A cursed player neglects the implicit information about the type she is facing, which is being revealed through the other player's behavior. Here, the implicit information stems from the different incentive structures of different types to choose certain actions. Both models thus distort the correlation between the opponents' actual types and their strategies. In another contribution, the same authors contrast cursed "rational herding" and "naïve herding" (or inferential naïvety; Eyster and Rabin, 2010). Under inferential naïvety, players fail to fully attend to the strategic logic of the setting they are in. In consequence, they naïvely believe that each previous player's action reflects solely that player's private information.

There are further information-neglect concepts. First, correlation neglect in the updating process can make beliefs excessively sensitive to well-connected information sources (Levy and Razin, 2015; Enke and Zimmermann, 2017). Second, pivotality neglect and difficulties to infer from hypothetical events cause biases that can explain naïve voting and bidding behavior (Esponda and Vespa, 2014). Third, base-rate neglect (Bar-Hillel, 1980) and other fallacies are known to cause ignorance of some relevant (for example, statistical) information in favor of using other irrelevant information. Fourth, Hanna et al. (2014) develop a theory of learning through noticing and present supporting evidence from a field experiment, in which farmers often fail to notice readily available information to optimize their production technology.

Finally, the most studied situation of information neglect is the Monty Hall problem (Friedman, 1998; Page, 1998). In this puzzle, a game-show participant faces three doors. Behind two of the doors, there is a worthless prize (a "goat"), while behind the third, there is a highly valuable prize (a "car"). The most popular protocol now proceeds as follows.

The participant chooses one of the three doors, which remains closed. The game-show host (famously, Monty Hall in "Let's make a deal") opens one of the remaining two doors but never opens the door hiding the car (uniformly randomizing which door to open when the participant's initially chosen door hides the car). Finally, the participant is asked whether she wants to stick to her original choice of doors or to switch to the other unopened door that she had not chosen at the outset. Under this set of rules (in particular, that the host always opens a door, that the opened door never hides a car, and that the host randomizes between goat-hiding doors), it should be clear that switching yields a probability of two thirds of getting the car.[3] However, most people who are confronted with the puzzle prefer to stick to their initial choice, even after many repetitions (e.g., Friedman, 1998).

In the Monty Hall problem, most people seem to be "cursed", as they neglect the information contained in the host's choice of which alternative door to open. If this choice was random (and not restricted to doors hiding goats), then switching and non-switching would lead to the same probabilities. Note though that cursedness alone would not be enough to explain choices in the Monty Hall problem. If people saw the two remaining options as equivalent, they should not *favor* sticking to their initial choice.[4]

After Friedman (1998) introduced the problem into economics, a number of researchers have studied variants of the Monty Hall problem, looking at whether individual bias in the task would be reflected in the market prices of assets on the possible doors (Kluger and Wyatt, 2004; Siddiqi, 2009), at whether communication and competition eliminate the bias[5] or at how intelligence and communication relate to learning the optimal strategy (Palacios-Huerta, 2003). Tor and Bazerman (2003, p.353) attribute non-switching (as well as two other empirical phenomena) to missing attention and a resulting failure "to consider all of the information needed".

This review of information-neglect studies and theories is certainly not exhaustive, but it is sufficient to show that there are many reasons to believe that people may *not* be able to make use of the information contained in their own type. Hence, people may not project their type onto others, even when it would be rational. We thus formulate the following alternative hypothesis to projection.

Information Neglect (Alternative) Hypothesis: *Participants do not over-infer from their own types but instead neglect the information that their type contains.*

[3] See, e.g., Selvin (1975a,b), Nalebuff (1987), or Gillman (1992).

[4] See, e.g., Friedman (1998) for a discussion of additional forces that may combine with cursedness to yield the observed non-switching behavior.

[5] "Competition" here includes detailed empirical feedback on the competitors' strategies and performance; see Slembeck and Tyran (2004).

In several of the studies discussed in this sub-section, nudges are used to make subjects use the available information. We discuss these nudges and how they relate to our own in the following.

2.2.3. Nudging people to use the available information

Enke and Zimmermann (2017) show that pointing the subjects to the correlation structure by juxtaposing a setup with correlation and a setup without correlation, effectively reduces correlation neglect and makes subjects take this element of the information structure into account. In our setup, we contrast the basic setup with replacement with a setup without replacement to point the subjects to the fact that the second ball is always drawn from the same urn.

Esponda and Vespa (2014) try several interventions to make subjects condition their behavior on the hypothetical information contained in the event of being pivotal. These include feedback, sequential rather than simultaneous play and several verbal hints. However, when they switch back to the main task in the experiment, error rates jump back to the initial, high level. Hanna et al. (2014) summarize the relevant information for the farmers who are then better able to notice them and take them into account. In our experiment, we go to great length to make the setup as clear and easy to understand as possible. We try two different decision screens and even use physical urns and balls to demonstrate the information structure at the beginning of the session. Not noticing or understanding what happens at which point in time can thus be ruled out as an explanation for information neglect in our experiment.

For the Monty Hall setup, Page (1998) shows that increasing the number of doors from three to at least ten increases the percentage of people who switch, but that switching in the many-doors problems does not mean participants would switch in a subsequent or simultaneously administered three-doors task. This inspired another of our interventions, in which we show people several draws from the randomly selected urn before they have to guess the color of the other person's ball.

2.3. Setup

The main, focal task is the Guessing Task. In this task, participants have to guess the color of another participant's ball after being informed about their type (receiving a colored ball). To this task, we added different parts in the different treatments. In Treatment 1, the Guessing Task was followed by a Belief-Elicitation Task. In the Belief-Elicitation Task, we asked participants for the probabilities of the other participant's ball being of specific colors and about the probabilities of the different states of the world. We added this Belief-Elicitation Task as a diagnostic tool to understand the participants' reasoning behind their Guessing-Task choices.

After finding evidence for information neglect in Treatment 1, we conducted two subsequent treatments as robustness checks. Also in Treatments 2 and 3, the focal task remained the Guessing Task. In both treatments, the session started with the Guessing Task. Then, in alternating fashion, participants faced three different intervention tasks and the Guessing Task. We chose the different intervention tasks to nudge participants into realizing that they could learn relevant information from the color of their own ball. Before Treatment 3, we physically enacted the whole process behind the Guessing Task in front of participants: how an urn would be randomly selected and how the balls would be drawn with replacement. We enacted the process to ensure participants' failure to update was not due to the task being too abstract or poorly understood.

2.3.1. *Design and theoretic background*

2.3.1.1. *Guessing Task*

Nature draws one of the two states of the world, or one of eight urns, as depicted in Figure 2.1 below. The states of the world can be either Dark (D) as on the left side or Light (L) as on the right side of Figure 2.1, with three balls in each urn. Of the balls in urn D, a fraction $q_d = 2/3$ is dark (d) and $q_l = 1/3$ is light (l). Of the balls in urn L, a fraction $q_l = 2/3$ is light (l) and $q_d = 1/3$ is dark (d). The possible priors are: $\omega \in \Delta \equiv \{D, L\}$ with $Pr(D) = 4/8$ or $Pr(D) = 5/8$, or $Pr(D) = 7/8$.

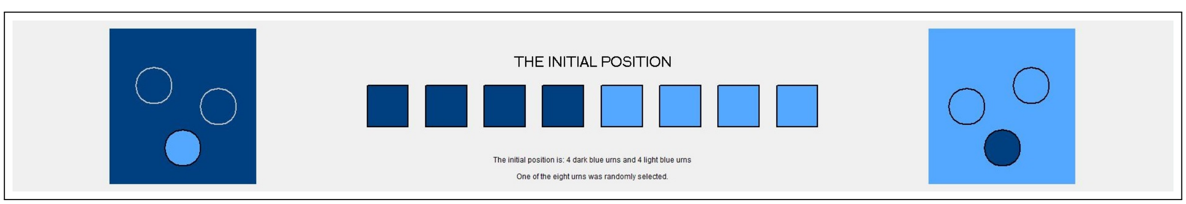


Fig. 2.1. Representation of the prior distribution in the experiment in the 4/8-situation (four dark and four light urns).

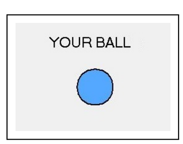


Fig. 2.2. Representation of the drawn signal in the experiment (one light ball).

Let us for now focus on the $4/8$-situation (prior) as in Figure 2.1 above, which is the most straight-forward for the participants and the most interesting one for our purposes. First, the participants are divided into pairs, and for each pair an urn is drawn. Nature's draw of the urn is not revealed, thus it is unknown in the remainder of the particular situation whether the balls are drawn from a D or an L urn. Then each of the two players receives a signal from this randomly selected urn, as depicted in Figure 2.2 above. The signals are also either d or l. The two signals were drawn independently with replacement for each player, which participants knew.

The projection is introduced as follows. As mentioned, in a group of two participants, the first participant i receives a private signal s_i in the form of a ball drawn from the urn. After replacement, the second participant j in the group receives a private signal s_j from the same urn ω. The participant i then has to guess whether participant j saw a light or a dark signal s_j.

In all situations of our laboratory experiment, a participant's own type is informative about the state of the world she and her partner both are in. Therefore, projecting one's own type onto the other participant (in other words, "going after one's signal") can be rational. In the example in Figures 2.1 and 2.2, the prior probability $Pr(D)$ is the uniform $4/8$, but then the participant i receives a light s_i signal. This should make i think that it is more likely that she and her partner j are in a light state of the world ω than that they are in a dark one. It is thus rational for i to project her own signal onto j by guessing that also j's signal s_j is light.

In the Guessing-Task setting, we thus define projection as follows. Rational projection, allowing for some errors in Bayesian inference, occurs if the participants guess the same color as their signal (in other words, "go after their signal") in: (i) close to 100% of the cases when the correct Bayesian posterior is larger than 50% and (ii) close to 0% of the cases when the correct Bayesian posterior is smaller than 50%. Over-projection occurs if the participants guess the same color as their signal in close to 100% of the cases when the correct Bayesian posterior is smaller than 50%. Conversely, under-projection—or information neglect—occurs if the participants *do not* guess the same color as their signal in close to 100% of the cases when the correct Bayesian posterior is larger than 50%. Note that in most cases it is rational for the subjects to go after their own signal when making the guess. We let the participants perform the Guessing Task 18 times (six times per prior, including counter-balancing of the colors).[6]

[6]The participants performed the Guessing Task in six sequences of the following order: 4/8-situation and 5/8-situation (and 7/8-situation in Treatment 1); counter-balanced 4/8-situation (light and dark instead of dark and light balls and urns) and counter-balanced 5/8-situation (and counter-balanced 7/8-situation in Treatment 1). The guesses were incentivized as follows. One iteration in one out of all Guessing Tasks was randomly

2.3.1.2. Belief-Elicitation Task

After the Guessing Task, we elicited the participants' beliefs. We asked them for the probability that the other participant's ball is dark or light if they saw a dark or a light signal, and for the probability that the urn is dark or light if they saw a dark or a light signal.[7] This way we can see if participants update their beliefs in the right or the wrong direction, or if they update at all, and at which step they fail to update correctly.

In the Belief-Elicitation Task, similarly as in the Guessing Task, we define projection as follows. We classify participants as rationally projecting players in the 4/8-situation if they indicate ball-probability beliefs between 50% and 59% (56% being the correct Bayesian posterior ball probability; we thus allow for some error). We define over-projection as participants indicating ball-probability beliefs at 59% or above. Conversely, under-projection—or information neglect—occurs if the participants indicate their ball probability beliefs at 50%. We avoid classifying the very few participants who indicate their ball probability beliefs below 50%.

For the 5/8- and 7/8-situations' theoretic predictions and definitions, please see Appendix A. Note that in most cases, participants' rational choice would have been to state their probability beliefs strictly above 50%.

2.3.2. Interventions

After finding evidence for information neglect instead of projection in Treatment 1, we conduct two subsequent treatments as robustness checks. In Treatments 2 and 3, we include four iterations of the Guessing Task each. We then try to facilitate projection (rational updating) by three interventions each, which are all tailored to nudge participants into realizing that they could learn something about the other participant's type from their own type. Please refer to Table 2.1 below for an overview of the three treatments.[8]

selected, and participants received 12 points if the corresponding guess was correct.

[7]The participants performed the Belief-Elicitation Tasks in the following order: ball-probability questions for the 4/8-situation and for the 5/8-situation (and for the 7/8-situation in Treatment 1), followed by urn-probability questions for the 4/8-situation and for the 5/8-situation (and 7/8-situation in Treatment 1). The beliefs in Treatment 1 were incentivized as follows. Participants received one point for each belief that did not differ by more than 2.5 percentage points from the correct probability. In Treatment 2, participants also received one additional point for each belief that did not differ by more than 0.5 percentage points from the correct probability. In Treatment 3, we used the loss-aversion-adjusted scoring rule introduced by Offerman and Palley (2016).

[8]In each of the four Guessing Task iterations of Treatments 2 and 3, the participants perform the Guessing Task 12 times (six times per prior, including counter-balancing). In contrast to Treatment 1 with three different priors, we only provide scenarios with two different priors in Treatments 2 and 3: $Pr(D) = 4/8$ and $Pr(D) = 5/8$.

Treatment 1	Treatment 2: Robustness checks	Treatment 3: Robustness checks
		Intervention 0: Physical urns
Guessing Task	Guessing Task	Guessing Task
Belief elicitation I	*Intervention 1: w/o replacement draws*	*Intervention 1: Belief elicitation II*
	Guessing Task	Guessing Task
	Intervention 2: Belief elicitation I	*Intervention 2: Strategy elicitation*
	Guessing Task	Guessing Task
	Intervention 3: Strategy elicitation	*Intervention 3: Several draws*
	Guessing Task	Guessing Task

Table 2.1: Overview of the three treatments.

In Treatment 2, the three Interventions were a repetition of the Guessing Task where the balls were drawn without replacement, the Belief-Elicitation Task, and a Strategy-Elicitation stage. Intervention 1, the Guessing Task without replacement, was meant to draw the participants' attention to the fact that the balls were drawn from the same urn also in the focal Guessing Task, and that there was replacement in the focal task. This should have brought the participants' attention to the fact that drawing two balls of the same color was more likely than drawing two balls of different colors. In contrast to the Guessing Task where the participants mostly need to "go after their signals" or project, they have to "go against their signals" or under-project in Intervention 1.[9]

Intervention 2 of Treatment 2, the Belief-Elicitation Task, was meant to nudge the participants into realizing that the urns had different posterior probabilities after observing their specific signal.[10] In Intervention 3, we asked the participants to write down the reasoning behind their reported probabilities from Intervention 2. Intervention 3 should have prompted the participants to think again—and more deeply—about the task, giving them another opportunity to adjust their subsequent Guessing-Task choices.

We start Treatment 3 by adding a demonstration using physical urns to the instructions at the beginning of the treatment. The participants first read on-screen instructions about the Guessing Task and the experiment in general, similarly as in Treatments 1 and 2. Then right before the beginning of the first Guessing-Task iteration, the participants learn about the process in the Guessing Task again, only this time it is demonstrated with physical urns and commented by the experimenter. We use a number of dark balls and light balls in the size of table tennis balls for demonstration. Furthermore, we use intransparent bags or urns

[9]The participants perform Intervention 1 involving only without-placement draws 12 times: six times per prior, including counter-balancing.

[10]Note that in Treatment 1, we included the Belief-Elicitation Task only as a diagnostic tool after the Guessing Task. In Treatment 2, we were interested in what happens to the Guessing-Task choices *after* Belief-Elicitation.

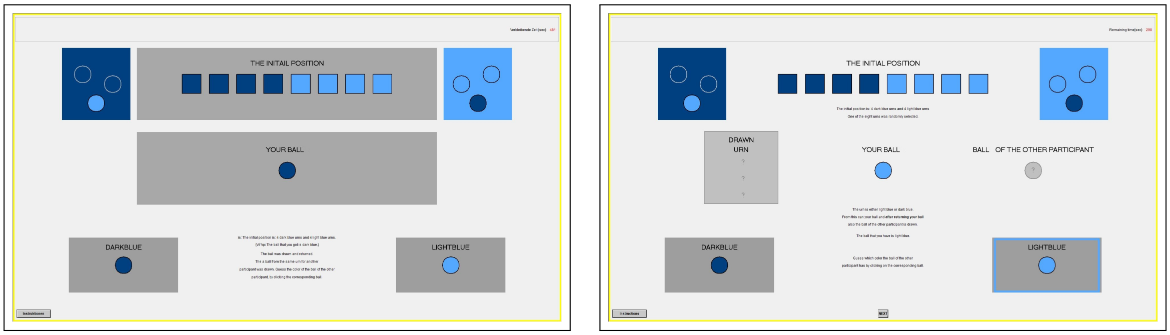
that are black on the outside and dark or light respectively on the inside. The appearance and numbers of the balls and bags are composed in a way that mimics the appearance and numbers described in the on-screen instructions right before.

In Treatment 3, the three in-between interventions were—again—a Belief-Elicitation Task, the Strategy-Elicitation Task, and a Guessing Task with several draws. The rationale for the first two interventions was two-fold. First, they were meant to capture any changes induced by our pre-play enacting of the situation. Second, the Belief-Elicitation Task this time used the L-adjusted scoring rule introduced by Offerman and Palley (2016), which should approximate a proper elicitation procedure without drawing the probability belief reports towards the 50% probability.[11]

Intervention 3 of Treatment 3, the Guessing Task with several draws, should have drawn participants' attention to the fact that signals are informative. In this task, participants saw three draws with replacement from the urn in addition to their signal. After seeing the first-drawn signal from the randomly drawn urn (as in the focal Guessing Task), we show the participants three additional with-replacement signals from this urn before they have to make their guess about the color of the other participant's ball. Seeing many balls of the same color should make participants realize that one of the urn types is more likely to be the randomly drawn one, much like facing a many-door Monty Hall problem helped participants make the right choice in the study by Page (1998). Thus, we may expect participants to transfer back this insight when returning to the single-signal Guessing Task afterwards.

Finally, note that we added an additional robustness check between Treatment 1 and Treatments 2 and 3. Namely, we provide more detailed decision screens and more elaborate instructions in Treatments 2 and 3, as depicted in Figure 2.3 below. In particular, the more elaborate screen in Figure 2.3 (b) stresses the intermediate step of one particular urn having been drawn, as well as the fact that both balls (being positioned on the same line as the urn) come from the same urn. Please see the instructions in Appendix B.

[11]This is a robustness check for the results from the earlier Belief-Elicitation Tasks. In those tasks, participants received a prize for a probability report that did not deviate from the true, objective probability by more than 2.5 percentage points. A loss-averse participant estimating the probability at, e.g., 52% with some uncertainty might in this case prefer to report 50%. The loss-aversion-adjusted choice list is designed to avoid probability belief reports at 50% whenever the participant's belief does not coincide with 50%.



(a) (b)

Fig. 2.3. Changes in the decision screen from Treatment 1 (a) to Treatments 2 and 3 (b).

2.3.3. Procedure

We programmed all treatments using z-Tree (Fischbacher, 2007) and recruited all participants using ORSEE (Greiner, 2015). We conducted eight experimental sessions at the Lakelab in Konstanz, Germany—four for Treatment 1 and two for Treatments 2 and 3 each. The results of Treatment 1 are based on the following sample composition: $n = 116$ participants over four sessions, on average 22 $\pm$ standard deviation (SD) of 2 years old, 38% male, average payment of € 18 $\pm$ SD 5 for about 90 minutes.[12] The results of Treatment 2 are based on the following sample composition: $n = 58$ participants over two sessions, on average 23 $\pm$ SD 2 years old, 48% male, average payment of € 13 $\pm$ SD 4 for about 70 minutes. Similarly, the results of Treatment 3 are based on the following sample composition: $n = 54$ participants over two sessions, 48% of the participants are male, on average 21.3 $\pm$ SD 2.2 years old, average payment of € 16 $\pm$ SD 5 for about 70 minutes.

2.4. Results

2.4.1. How much projection is there really?

2.4.1.1. Guessing Task

Starting with Treatment 1, the results show that participants update their prior beliefs notably less than expected by theories of type projection or Bayesian updating. The participants rather appear to be guided by the prior distribution. In the right-most column of Table 2.2 below, we show how much projection we would expect if participants learned from their

[12]Participants in the sessions for Treatment 1 participated in a completely unrelated pilot treatment for another study after Treatment 1. The instructions for the additional treatment were distributed only after Treatment 1 was finished.

signals (given the correct Bayesian posteriors in the parentheses). For comparison, we show the actual projection frequency in the middle column *Subjects' guesses projecting the signal*. The column shows how often the participants actually guessed the same color signal that they received.

Let us again focus on the simplest $4/8$-situation. The result of 76% differs strongly from the 100% benchmark. Keeping in mind that random choice would mean 50% same-color guesses—and that projection would mean that participants *always* go with the signal—76% same-color guesses is not a result to support projection.[13]

Prior	Signal	Subjects' guesses projecting the signal	Theoretical predictions	
4/8-situation	D or L	**76.3%**	100% projecting	(56%)
5/8-situation	D	94.5%	100% projecting	(59%)
	L	**36.1%**	100% projecting	(52%)
7/8-situation	D	94.7%	100% projecting	(64%)
	L	6.9%	0% projecting	(41%)

Table 2.2: Results of the Guessing Task in Treatment 1.

2.4.1.2. Belief-Elicitation Task

The results of Treatment 1 in the Belief-Elicitation Task after the Guessing Task show even stronger evidence for information neglect and against projection, especially if we again focus on the simplest $4/8$-situation as in Figure 2.4 below. More than 60% of the participants report probability beliefs of 50%, which coincides with the $4/8$ prior probability. A small minority of participants report probability beliefs of 67%. These participants seem to be 100% sure about the urn they are facing—or neglect that a 67% probability can only be correct if they were completely certain about the randomly drawn urn.

The sub-figure on the left shows (i) the participants' probability beliefs that the second ball is dark if the signal is dark and (ii) the participants' probability beliefs that the second ball is light if the signal is light (for both, the correct Bayesian posterior is 56% as indicated by the dashed line). The sub-figure on the right shows the reported answers to two other questions: (i) the participants' probability beliefs that the urn is dark if the signal is dark

[13] Another result from the non-50%-prior situations also indicates that the participants are biased towards the prior information. The 36% in the $5/8$-situation differ even more strongly from the 100% benchmark. The other percentages in the $5/8$-situation and $7/8$-situation on the contrary are quite close to the respective benchmarks.

and (ii) the participants' probability beliefs that the urn is light if the signal is light (both 67% correct Bayesian posterior, dashed line).

In all cases, the average (solid line) differs significantly at a 1% significance level from the correct (dashed line) in Wilcoxon signed-rank tests (p-values <0.001). Roughly 70% of the participants report their probability beliefs at 50% (for both urn and ball colors), that is, about two thirds of the participants appear to believe that they cannot learn anything about the unknown selected urn from their received signal.

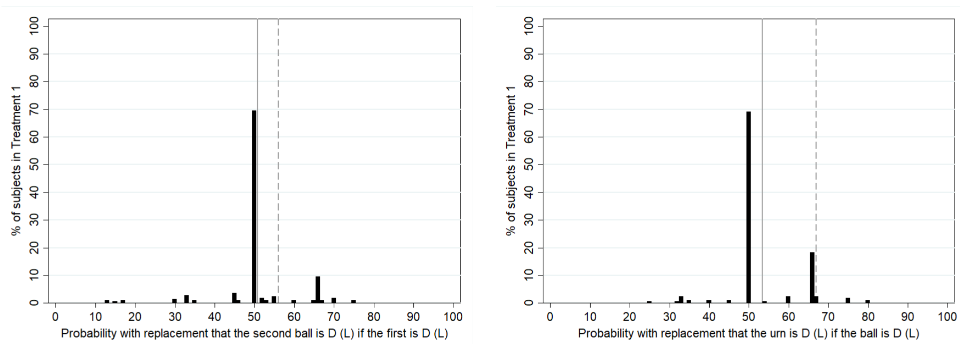


Fig. 2.4. Results of the Belief-Elicitation Task in Treatment 1 (with ball-probability beliefs on the left and urn-probability beliefs on the right); theoretical predictions dashed line, average beliefs solid line.

2.4.2. Are the nudging interventions effective?

The results indicate that the only intervention to somewhat help is the Belief-Elicitation Task. In both Treatments 2 and 3, it brings up percentages in the Guessing Task by around 5-8 percentage points. There are significant differences between behavior in the Guessing Tasks before and after the Belief-Elicitation intervention at a 5% significance level (that is, between the second and third iterations in Table 2.3, p-value=0.034, and the first and second iterations in Table 2.4, p-value=0.023). About 10-16% of the "information neglect" people seem to change their behavior from neglect to projection, and their performance does not deteriorate again afterwards. Otherwise, nothing helps (all p-values>0.352 and >0.696 in Treatments 2 and 3, respectively).

Remember that we would see "50%" in all the iteration columns if the participants were guessing light or dark completely randomly, while in accordance with the theoretical predictions, we would expect to see "100%" in all the iteration columns in Tables 2.3 and 2.4 below. What we see instead is that the share stays below 82% in all iterations of the 4/8-situation. We conclude that none of the interventions can substantially improve the participants' performance in the Guessing Task.

Prior	Signal	1^{st} iter.	*Interv. 1*	2^{nd} iter.	***Interv. 2***	3^{rd} iter.	*Interv. 3*	4^{th} iter.
4/8-situation	D or L	71.1%	*21.7%*	69.2%	*s.2.4.2.1*	76.2%	*s.2.4.2.2*	73.1%
5/8-situation	D	89.8%	*57.0%*	88.3%	*s.2.4.2.1*	94.8%	*s.2.4.2.2*	94.2%
	L	30.3%	*2.7%*	22.3%	*s.2.4.2.1*	28.2%	*s.2.4.2.2*	24.6%

Table 2.3: Results of the four Guessing-Task iterations and three Interventions in Treatment 2.

Prior	Signal	1^{st} iter.	***Interv. 1***	2^{nd} iter.	*Interv. 2*	3^{rd} iter.	*Interv. 3*	4^{th} iter.
4/8-situation	D or L	71.3%	*s.2.4.2.1*	79.0%	*s.2.4.2.2*	80.3%	*s.2.4.2.3*	81.4%
5/8-situation	D	91.2%	*s.2.4.2.1*	90.6%	*s.2.4.2.2*	96.8%	*s.2.4.2.3*	96.8%
	L	37.1%	*s.2.4.2.1*	31.4%	*s.2.4.2.2*	32.0%	*s.2.4.2.3*	28.9%

Table 2.4: Results of the four Guessing-Task iterations and three Interventions in Treatment 3, after the Intervention 0 with physical urns.

2.4.2.1. Belief-Elicitation interventions

The situation concerning the reported probability beliefs in Treatment 2 barely improved compared to Treatment 1. Let us again focus on the simplest $4/8$-situation, as in Figure 2.5 below. Roughly 60% of the participants still report probability beliefs of 50%. Again, a small minority of participants report probability beliefs of 66% or 67%.

However, the situation concerning the reported probability beliefs improved in Treatment 3 compared to the Treatments 1 and 2. Only 40.7% of the participants report probability beliefs of 50% in the Treatment 3 for the *ball probabilities*, as depicted in the left sub-figure of Figure 2.6 below. Again, a minority of participants (14.8% and 13.0% in D and L accordingly) report probability beliefs of 66% or 67%.

Note, however, that only less than 20% (18.5% and 16.7% in D and L accordingly, as depicted in the right sub-figure of Figure 2.6 below) report probability beliefs of 50% for the *urn probabilities* in Treatment 3. Instead, 25.9% of participants report probability beliefs of 66% or 67%. This substantial change in stated beliefs is hard to explain under the assumption that the distributions of the true beliefs are the same under both rules.[14] Instead, the results

[14]Under the elicitation rules in Treatment 1 (2), risk- or loss-aversion can maximally explain a shift toward 50% of 2.5 (0.5) percentage points. Hence, adjusting for these potential biases could only explain a shift of the same magnitude away from 50%. However, if the stated beliefs under the L-adjusted rule in Treatment 3 were true, they would imply far bigger shifts.

suggest that the participants' beliefs are influenced differently by the elicitation method itself. Potentially, the L-adjusted scoring rule nudges some participants' to think differently about the problem.

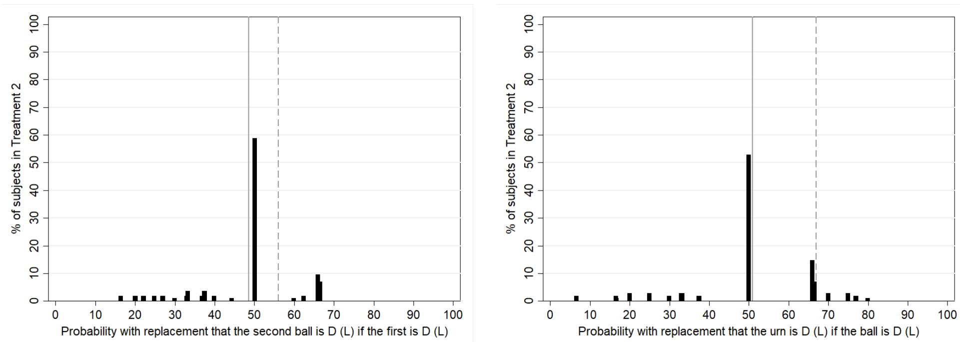


Fig. 2.5. Results of the Belief-Elicitation Task in Treatment 2 (with ball-probability beliefs on the left and urn-probability beliefs on the right); theoretical predictions dashed line, average beliefs solid line.

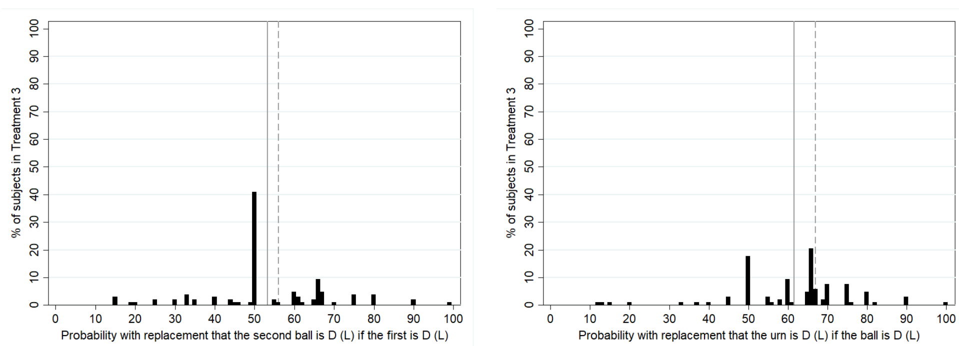


Fig. 2.6. Results of the Belief-Elicitation Task in Treatment 3 (with ball-probability beliefs on the left and urn-probability beliefs on the right); theoretical predictions dashed line, average beliefs solid line.

It is unclear why so many subjects still report 50% ball-probability beliefs in all three experiments. Related to this, it is also unclear why the participants do not sufficiently alter their behavior in the Guessing Task, in either of the three treatments, even after the helping interventions. The participants' elicited strategies (Intervention 3 in Treatment 2 and Intervention 2 in Treatment 3) gives further insights into both of these questions.

2.4.2.2. Strategy-Elicitation interventions

How do the participants explain their answers in the Belief-Elicitation Task? We asked the participants to describe their reasoning process behind the beliefs they provided in free text. Let us again focus on the simplest $4/8$-situation, for which we classified and reclassified the reported strategies into several categories. We can compare the answers of the participants who report 50% probability beliefs with those who report non-50% probability beliefs, as depicted in Tables 2.5 to 2.7 below. The reasoning strategies between these two groups differ substantially, as the participants who report non-50% probability beliefs mention learning from the signal most often in both Treatment 2 and Treatment 3 (e.g., "I assumed that a dark urn was chosen, as it contains more dark balls; so there are 3 balls, 2 of which are dark").

The participants who report 50% beliefs, however, most often mention the following three reasons behind their answers. Firstly, they argue that the probability of drawing a dark urn is the same as drawing a light urn (e.g., "Each urn has the same probability of being chosen, hence the uniform distribution 50-50"). Secondly, they count the colors of the balls in each of the two urns together (e.g., "A light and a dark urn constitute a total of 3 light balls and 3 dark balls; the probability is thus 50%"). Thirdly, they argue that there is no difference which signal has been drawn, given that the draws are with replacement (e.g., "The ball was put back in the urn, such that 6 balls were in the game all the time; of these, there were always 3 light balls and 3 dark balls"). Meanwhile, as usual, some of the free-text answers were difficult to categorize meaningfully, either due to scarce information or confusion, or other reasons.

Strategy categories	Subjects with 50% beliefs	with non-50% beliefs
1. Learn from the signal	-	51%
2a. Equal probabilities	55%	-
2b. Equal number of balls	20%	2%
2c. No difference after replacement	15%	2%
3a. Scarce information	5%	12%
3b. False or confused	5%	33%
N	65	51

Table 2.5: Results of the Intervention 3 (Strategy Elicitation) in Treatment 2.

Due to the similarities in the reported probability beliefs, we pool the strategies for the ball probabilities and the urn probabilities together for Treatment 2, as depicted in Table 2.5 above. For Treatment 3, however, the reported beliefs differ substantially between the ball probabilities and the urn probabilities. We thus depict the respective reported strategies separately in Tables 2.6 and 2.7 below.

We conclude that the strategies mentioned in Treatment 3 can be categorized somewhat similarly to the strategies in Treatment 2. However, the shares of the respective categories differ. Namely, the participants reporting non-50% beliefs mention learning from the signal more often in Treatment 3 than in Treatment 2 (e.g., "I acted as if the urn was [dark], given that the drawn ball in this case is [dark] with a 66% probability").

Strategy categories	Subjects with 50% beliefs	with non-50% beliefs
1. Learn from the signal	-	62.5%
2a. Equal probabilities	36.4%	3.1%
2b. Equal number of balls	54.5%	3.1%
2c. No difference after replacement	-	-
3a. Scarce information	-	6.3%
3b. False or confused	-	15.6%
3c. Other	9.1%	9.4%
N	22	32

Table 2.6: Results of the Intervention 2 (Strategy Elicitation, ball probability) in Treatment 3.

Strategy categories	Subjects with 50% beliefs	with non-50% beliefs
1. Learn from the signal	-	77.3%
2a. Equal probabilities	20.0%	-
2b. Equal number of balls	30.0%	-
2c. No difference after replacement	-	-
3a. Scarce information	10.0%	13.6%
3b. False or confused	30.0%	2.3%
3c. Other	10.0%	6.8%
N	10	44

Table 2.7: Results of the Intervention 2 (Strategy Elicitation, urn probability) in Treatment 3.

2.4.2.3. Other interventions

Moreover, neither the *without-replacement* nor the *several-draws* interventions help. Also the *physical-urns* intervention cannot nudge the participants closer to the correct Bayesian beliefs or guesses. For the effects of the without-replacement intervention, please see the 1^{st} and 2^{nd} iteration columns of Treatment 2 and the without-replacement Intervention 1 in the respective middle column of the Table 2.3 above again. If in all the iteration columns we would expect to see 100%, then in the Intervention 1 column we would expect to see no projection at all, or 0%. We can conclude that the participants do change their behavior to a considerable extent in the without-replacement draws as compared to the with-replacement draws. That is, they do appear to understand that the same urn is used for them and the other participant in their group. However, in the 2^{nd} iteration of the Guessing Task afterwards compared to the 1^{st} iteration of the Guessing Task before, we see no significant learning at all.

For the effect of the several-draws intervention, please see the 3^{rd} and 4^{th} iteration columns of Treatment 3 in the Table 2.4 above again. We allow the participants to receive three additional with-replacement signals from this urn before making their guess about the color of the other participant's signal. The participants do indeed appear to project their signals more after receiving three or four balls of the same color than after receiving one ball in the single-draw Guessing Task (compare Tables 2.8 and 2.9 below with the Table 2.4 above). That is, they do appear to understand that the same urn is used for all the four draws—and that this urn is then also used for the other participant in their group. Note, however, that we do not see 100% projection rates even when participants get to see three or four balls of the same color.[15]

Finally, for the physical-urns intervention, please see the 1^{st} iteration column of Treatment 3 in the Table 2.4 above. The Intervention 0 with physical urns was performed already before the Guessing Task. For the purposes of assessing the effect of this intervention, we can compare the 1^{st} iteration of the Treatments 2 and 3, respectively, in the Tables 2.3 and 2.4 above. We conclude that the participants hardly change their behavior.

Signal	D	L	Prediction	Probability
4D, 0L	25	5	100%	65%
%	***83.3%***	*16.7%*		
3D, 1L	68	7	100%	60%
%	***90.7%***	*9.3%*		
2D, 2L	56	60	-	50%
%	*48.3%*	*51.7%*		
1D, 3L	8	70	0%	40%
%	*10.3%*	***89.7%***		
0D, 4L	2	23	0%	35%
%	*8.0%*	***92.0%***		

Table 2.8: Results of the Intervention 3 with four ball draws, 4/8-situation.

[15]Using this data, we can quantify the fraction of people who either do not pay any attention, fall prey to the gambler's fallacy, or have another weird quirk in their thinking to be around 10%—taking the number of people doing the wrong thing even in the "several-draws" task.

Signal	D	L	Prediction	Probability
4D, 0L	30	3	100%	65%
%	***90.9%***	*9.1%*		
3D, 1L	55	5	100%	62%
%	***91.7%***	*8.3%*		
2D, 2L	48	13	100%	54%
%	*78.7%*	*21.3%*		
1D, 3L	11	39	0%	43%
%	*22.0%*	***78.0%***		
0D, 4L	28	92	0%	36%
%	*23.3%*	***76.7%***		

Table 2.9: Results of the Intervention 3 with four ball draws, 5/8-situation.

2.5. Discussion and conclusions

The starting point of our investigation was the observation of two growing strands of the literature that make starkly contrasting predictions regarding the use of information. The first strand is the projection literature that suggests that people typically infer too much from their own type about the types of other players. Several papers in this literature appear to take projection as a pervasive ingredient of behavior and use it as a primitive of the model that they set up to explain certain real-world phenomena, such as bidding in auctions. In the first paragraph of his paper presenting a projection-based model to explain the winner's curse and over-bidding in auctions more generally, Breitmoser (2019, p.423), for example, states that "projection is of intuitive relevance in all choices under incomplete information".

The second strand of the literature that we dub "information neglect" literature starts from the opposite end and assumes some form of information neglect. Eyster and Rabin (2005, p.1624), who set out to explain the same phenomenon as Breitmoser (2019), for example, start as follows: "In this paper, we formally model a generalization of the winner's curse that assumes that players in a Bayesian game underestimate the extent to which other players' actions are correlated with their information".

There are a number of empirical studies, of different setups *notabene*, that support the basic assumptions of either camp. To find out, we designed a simple experimental task, which allows us to study how strongly experimental participants project their type onto another player. In this setting, which is very similar to the settings used in other experiments of games with imperfect information, we find no evidence for projection. Instead, we find strong evidence for information neglect. Focusing on the Guessing Task in the

4/8-situation, for example, we see around 70–75% projection where we would expect 100% projection. Assuming that those participants who do not deliberately choose to project their type randomize uniformly, this means that only around 40–50% take the information contained in their type into account. Judging by the stated probability beliefs, this share even drops to 30–40%.

These results raise the question of what the reason for so much information neglect might be, which leads us back to the literature on the Monty Hall problem. Burns (2017) argues that people follow a "no-change principle", under which probabilities remain the same after the host opens one of the doors. He calls the phenomenon "no-change principle" because the probabilities that he elicits from the participants suggest a reasoning that implies that "there has be no change to these doors, so they maintain their equal status" (Burns, 2017, p.1701).[16]

The participants in our experiment seem to commit the same error. Their answers to our questions regarding their reasoning clearly reflect a "no-change principle", too. *The ball has been put back, hence nothing has changed and the probabilities must be the same as the prior probabilities*, seems to be the faulty logic that many of our participants follow. While this might explain the absence of projection and presence of information neglect in our setup, further research is needed to understand when to expect which type of effect and, hence, when to apply a projection or an information-neglect theory to predict behavior. Our study clearly shows that projection bias is not a general phenomenon that can be expected to occur in any setup, as might be presumed.

[16]Interestingly, those who applied the relative version of the "no-change principle" and calculated new probabilities that simply reflect the prior probabilities of the remaining options showed a higher-than-average value in the cognitive-reflection test. Hence, it is not necessarily cognitive laziness or cognitive immobility that yields such behavior.

Appendix A. Theoretical predictions

Task	Situation	Signal	Prediction	Under-proj.	Rational proj.	Over-proj.
Guesses	4/8-situation	Dark	100%	[0%;97%)	[97%;100%]	-
		Light	100%	[0%;97%)	[97%;100%]	-
	5/8-situation	Dark	100%	[0%;97%)	[97%;100%]	-
		Light	100%	[0%;97%)	[97%;100%]	-
	7/8-situation	Dark	100%	[0%;97%)	[97%;100%]	-
		Light	0%	-	(0%;3%]	(3%;100%)

Table A2.1: Theoretical predictions in the Guessing Task.

Task	Situation	Signal	Prediction	Under-proj.	Rational proj.	Over-proj.
Beliefs (ball)	4/8-situation	Dark	56%	[0%;52%)	[53%;59%]	(59%;100%]
		Light	56%	[0%;52%)	[53%;59%]	(59%;100%]
	5/8-situation	Dark	59%	[0%;56%)	[56%;62%]	(62%;100%]
		Light	52%	[0%;50%]	(50%;55%]	(55%;100%]
	7/8-situation	Dark	64%	[0%;61%)	[61%;67%]	(67%;100%]
		Light	41%	[0%;39%)	[39%;50)	[50%;100%]
Beliefs (urn)	4/8-situation	Dark	67%	[0%;64%)	[64%;70%]	(70%;100%]
		Light	67%	[0%;64%)	[64%;70%]	(70%;100%]
	5/8-situation	Dark	77%	[0%;74%)	[74%;80%]	(80%;100%]
		Light	55%	[0%;52%)	[52%;58%]	(58%;100%]
	7/8-situation	Dark	93%	[0%;90%)	[90%;96%]	(96%;100%]
		Light	22%	[0%;19%)	[19%;25%]	(25%;100%]

Table A2.2: Theoretical predictions in the Belief-Elicitation Task (ball p and urn p).

Appendix B. Instructions

The following instructions (translated from German) correspond to the most frequently used text versions over the three treatments. In the meantime, there were a few adjustments between the treatments. Firstly, we changed the layout of the decision screen and the instruction texts before Treatment 2. Secondly, we changed the colors of the light blue and dark blue balls to white and blue balls, respectively, before Treatment 3.

B.1. *Instructions for the focal Guessing Task*

Overview

Welcome to this experiment. We ask you not to speak with the other participants and turn off your mobile phone and other mobile technical devices during the experiment. For taking part in today's experiments, you are paid in cash at the end. The amount of the payout depends partly on chance and partly on your decisions. It is thus important that you carefully read and understand the instructions before starting the experiment.

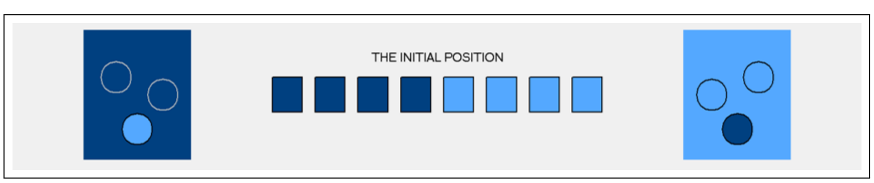
Today's experiment includes **six parts**, each comprising several rounds. In the end, several randomly drawn rounds will be paid out. From the parts 1, 3, 5 and 6 of the experiment, **one** round is randomly selected and paid out. In addition, **one** randomly drawn round from part 2 and **all** rounds of part 4 are paid out. The not-drawn rounds will not be paid out.

Your payout is based on the points you earned in the rounds, plus 6 euro for completing a subsequent questionnaire. The conversion of the points into euro happens as follows. Each point is worth 50 cents, such that: **10 points = 5.00 euro**. Each participant is privately paid so that other participants cannot see how much you have earned.

Composition of the experiment

This experiment includes six different parts. The 1^{st}, 2^{nd}, 3^{rd}, 5^{th} and 6^{th} parts each comprise 12 identical rounds. In each round, a decision that can be correct or incorrect has to be made. The tasks in these parts are similar in structure. The 4^{th} part comprises 8 identical rounds. In each round, questions that can be correct or incorrect have to be answered. **Before starting each part of the experiment, you will learn the instructions for the respective part.** On the next two pages, you will find the instructions for **part 1**.

The initial position



In each round, there are a total of 8 urns of balls. Each urn is either light blue and contains two light blue and one dark blue ball or dark blue and contains two dark blue and one light blue ball.

There can be three different initial positions. Either 3, 4 or 5 of the 8 urns can be dark blue. The remaining urns are light blue. The initial position is displayed in each round at the top of the screen.

In the example shown above, you see a dark blue urn on the left, a light blue urn on the right and an initial position in the center with 4 light blue and 4 dark blue urns.

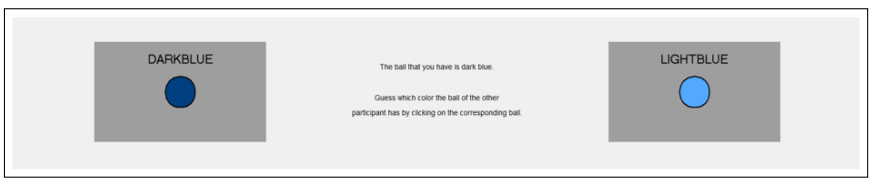
The drawn urn

As you already know, there are two types of urns: dark blue urns and light blue urns. The dark blue urn contains 1 light blue and 2 dark blue balls, the light blue urn 2 light blue and 1 dark blue. In each round, one of the 8 urns from the initial position is randomly drawn by the computer. Each urn has the same drawing probability. However, you will not know whether the drawn urn is light blue or dark blue.

Draw of the ball(s)

In each round, one or more balls are drawn randomly from the selected urn. You will see the color of the randomly drawn ball(s) on the screen. Whenever several balls are drawn, they are drawn one after the other and immediately returned. Another player will also receive one or more randomly drawn balls from the selected urn. Each ball in the drawn urns always has the same selection probability. This applies to the other participant as well as to you. Each participant only sees the color of his or her ball(s), not that of the other participant.

The tasks



In the first part of the experiment, exactly one ball is displayed to each participant. You then have to guess the color of the ball that was drawn for the other participant. Before making your decision, you will be informed of the (for both identical) initial position and the color of your drawn ball. In the shown example, this is dark blue.

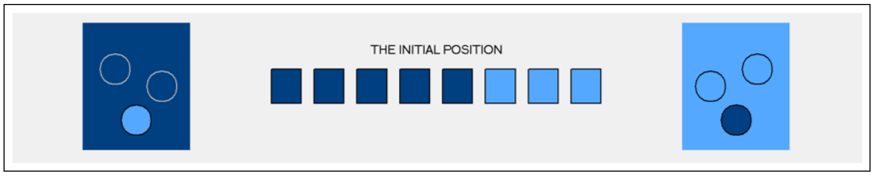
In addition to the normal rounds described so far, there are situations in which neither the urn nor the ball is randomly drawn. The predetermined situations amount for less than 1/6 of all rounds. You cannot distinguish these rounds from the normal rounds. For your payout, however, only those rounds are relevant, as described, for which chance decides which urn is drawn and which ball is drawn from this urn.

The payout

From the 12 rounds of the second part, one round is randomly selected. One decision is randomly selected from the 48 rounds of parts 1, 3, 5 and 6 of the experiments. The predefined situations cannot be selected. If you have given the correct answer, that is, the correct color of the other participant's ball, you receive 12 points. The details on the payment of the 4^{th} part can be found at the beginning of part 4.

Questions?

Take the time to look carefully at the instructions. If you have any questions, please lift your hand. An experimenter will then come to your place.



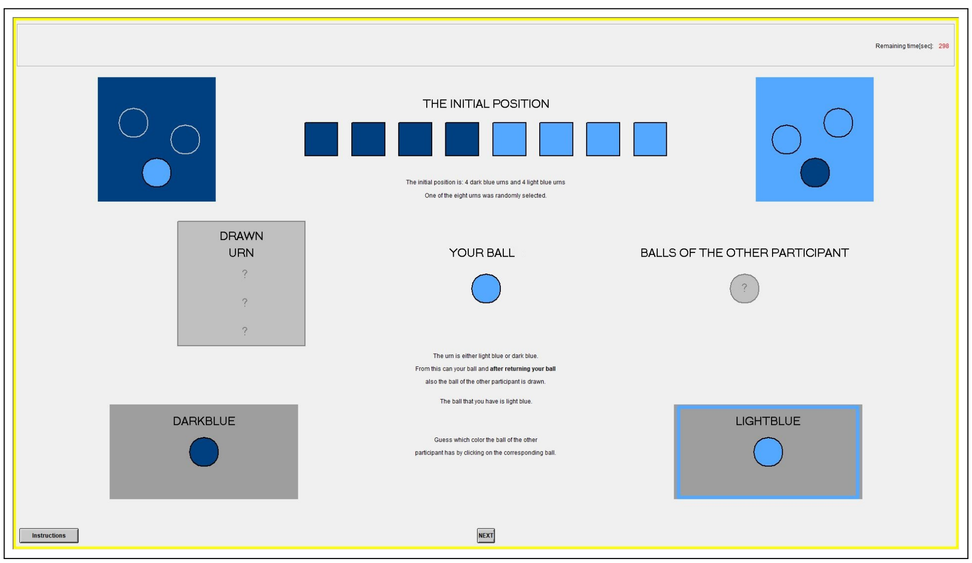
Please answer the following questions about the experiment.

1. How many light blue balls are there in a light blue urn? [...]
2. In the above example, with what probability is a light blue urn drawn (in %)? [...]
3. How many rounds will be paid out in the second part? [...]

Part 1 [Part 3 / Part 5 / Part 6] of the experiment

In each round of the 1^{st} [3^{rd}, 5^{th}, 6^{th}] part, a ball is randomly drawn from the selected urn. You will learn the color of the randomly drawn ball on the screen. The ball is drawn and returned immediately. Another player also receives a randomly drawn ball from the same selected urn after your ball has been returned.

In the [first] part, exactly one ball is displayed to each participant. You then have to guess the color of the ball that was drawn for the other participant. Before making your decision, you will be informed of the (for both identical) initial position and the color of your drawn ball. [...]



Payout [at the end of the experiment]

Participant [x] will now determine the payout-relevant decision by a dice throw. Please wait at your place while the payout-relevant initial positions and situations are diced out.

The first dice roll determines the payout-relevant initial position in the [first] part of the experiment: ”1” for the initial position with 4, ”2” for the initial position with 5 dark blue urns, also ”4” for the initial position with 4, ”5” for the initial position with 5 light blue urns. If a ”3” or ”6” is diced, the throw has to be repeated.

The second dice roll determines the payout-relevant part of the 1^{st}, 3^{rd}, 5^{th} 6^{th} parts: ”1” for the first part, ”3” for the third part, ”5” for the fifth part, ”6” for the sixth part. If a ”2” or ”4” is diced, the throw has to be repeated.

The third dice roll determines the payout-relevant initial position in the [second] part of the experiment: ”1” for the initial position with 4, ”2” for the initial position with 5 dark blue urns, also ”4” for the initial position with 4, ”5” for the initial position with 5 light blue urns. If a ”3” or ”6” is diced, the throw has to be repeated.

The fourth dice roll determines the payout-relevant situation under the previously diced initial position in the [first] part of the experiment: ”1” and ”4” for situation 1, ”2” and ”5” for situation 2, ”3” and ”6” for situation 3. And the fifth dice throw determines the payout-relevant situation under the previously diced initial position in the [second] part of the experiment: ”1” and ”4” for situation 1, ”2” and ”5” for situation 2, ”3” and ”6” for situation 3.

In the unlikely event of a predefined situation being diced, the computer randomly selects one of the other situations.

B.2. Interventions

B.2.1. Intervention 1 of Treatment 2

Part 2 of the experiment

In each round of the **2nd part**, a ball is randomly drawn from the selected urn. You will learn the color of the randomly drawn ball on the screen. The ball is drawn and **NOT returned**. Another player also receives a randomly drawn ball after your ball has been returned to the same selected urn.

Each participant thus sees exactly one ball. You then have to guess the color of **the next** ball that was drawn for the other participant. Before making your decision, you will be informed of the (for both identical) initial position and the color of your drawn ball.

B.2.2. Intervention 2 of Treatment 2

Part 4 of the experiment

In the **4th part**, your task is to answer questions. The questions relate to different scenarios. The scenarios will be similar to those in the first, second and thirds parts.

Your task is to specify the probability of a dark blue or a light blue ball being drawn in each scenario, or the probability that the urn is dark blue or light blue in each scenario.

In the **4th part**, you get 1 point for each correct answer. An answer is considered to be correct if the probability that you give is a maximum of 2.5 percentage points away from the correct probability. That is, if the correct probability is x (in %), you get 1 point for each response that is not greater than x+2.5 and not smaller than x-2.5. In addition, you get 1 point for each answer that is not greater than x+0.5 and not smaller than x-0.5. Your income is then converted into euro and paid out privately.

In the first and third part of the experiment, your task was to guess the color of the ball drawn for the other participant. Your ball was drawn and immediately [returned / NOT returned] before the other participant received a ball. Now you must indicate the probability that the ball of the other participant has a certain color and the probability that the urn has a certain color.

The initial position is: 4 [5] dark blue urns and 4 [3] light blue urns. If you saw a dark blue ball under the above initial position, with what probability (in %) was the ball of the other participant also dark blue? [...] If you saw a light blue ball under the above initial position, with what probability (in %) was the ball of the other participant also light blue? [...] If you saw a dark blue ball under the above initial position, with what probability (in %) was the selected urn dark blue? [...] If you saw a light blue ball under the above initial position, with what probability (in %) was the selected urn light blue? [...]

B.2.3. Intervention 3 of Treatment 2

Thank you, you have almost finished part 5 of the experiment. A few questions about your choices in the past parts of the experiment will now follow.

In the 1^{st}, 3^{rd} and 5^{th} part, your task was to guess the color of the ball drawn for the other participant. The ball was drawn and returned immediately before the other participant had received a ball.

Then in the fourth part, you had to specify the probability that the ball had a certain color and the probability that the urn had a certain color. Now we ask you to answer two question about your decisions. The initial position was: 4 dark blue urns and 4 light blue urns.

In Part 4, you answered the following question: If you saw a dark blue ball given the above initial position, what was the probability (in %) that the other player's ball drawn after replacing your ball was also dark blue? Your answer was (in %): [...] Please describe what you considered when answering this question: (please press "Enter" after entering the text) [...]

In Part 4, you also answered the following question: If you saw a dark blue ball under the above initial position, with what probability (in %) was the drawn urn dark blue? Your answer was (in %): [...] Please describe what you considered when answering this question: (please press "Enter" after entering the text) [...]

B.2.4. Intervention 1 of Treatment 3

Part 2 of the experiment

In Part 2, your task is to answer questions. The questions relate to different scenarios. The scenarios will be similar to those in the first part. In the first part of the experiment, your task was to guess what color the ball drawn for the other participant had.

Your task is to indicate the probability with which a blue or white ball is drawn in each scenario, or with what probability the drawn urn is blue or white in each scenario. Your payout will then depend on how well you choose your estimate.

The entering of your probability estimate is done in two steps: (1) First enter your estimate in the input field. (2) After that, another menu will be displayed on the right to reflect your input. Here you can also see how high your payout is if the other participant has indeed drawn a ball of the same color or a different color or if the urn has indeed the same color or a different color.

You can change your estimate in two ways. Firstly, you can enter a new estimate in the input field. Alternatively, you can directly click on a line in the menu on the right and update your estimate by clicking on "Customize". When you are satisfied with your decision, press the red "Confirm" button to proceed to the next task. The button will appear on the right as soon as you have entered your first estimate. The initial position is: 4 [5] blue urns and 4 [3] white urns.

Instructions

In the upper part of the screen, you can see the corresponding initial position. Your task: Indicate how likely you think it is that the other participant's ball is also blue [white] if you have seen a blue [white] ball, given the initial position shown above. Your payout will then depend on how well you choose your estimate.

Entering your probability estimate is done in two steps: (1) First enter your estimate in the input field. (2) After that, another menu will be displayed on the right to reflect your input. Here you can also see how high your payout is if the other participant has indeed drawn a same-colored or different-colored ball.

At this point, you can still change your estimate in two ways: Firstly, you can enter a new estimate in the input box. Alternatively, you can directly click on a line in the menu on the right and update your estimate by clicking on "Customize". When you are satisfied with your decision, press the red "Confirm" button to proceed to the next task. This will appear on the right as soon as you have entered your first estimate.

If you saw a blue ball given the initial position shown above, with what probability (in %) was the other player's ball blue as well? [...] If you saw a blue ball given the initial position shown above, with what probability (in %) was the urn also blue? [...] Give your estimate as an integer between 0 and 100.

B.2.5. Intervention 2 of Treatment 3

In the 1^{st} and 3^{rd} part, your task was to guess what color was the ball drawn for the other participant was. The ball was drawn and returned immediately before the other player got a ball. In the 2^{nd} part, you had to specify the probability that this ball had a certain color and with what probability the drawn urn had a certain color. Now we ask you to answer two questions about your decisions. The initial position: 4 blue urns and 4 white urns.

In the 2^{nd} part, you answered the following question: If you saw a blue ball under the initial position shown above, with what probability (in %) was the other player's ball also blue? Your answer was (in %): [...] Please describe what you considered when answering this question: (after entering your answer, please press "Enter") [...]

In the 2^{nd} part, you also answered the following question: If you saw a blue ball under the initial position shown above, what was the probability (in %) that the drawn urn was blue? Your answer was (in %): [...] Please describe what you considered when answering this question: (after entering your answer, please press "Enter") [...]

B.2.6. Intervention 3 of Treatment 3

Part 5 of the experiment

In each round of the 5^{th} part, a ball is randomly drawn four times from the drawn urn and returned back again. You will see the color of the randomly drawn balls on the screen. Each ball is drawn, displayed and returned immediately. Another participant will then also see four randomly drawn balls from the same drawn urn.

You then have to guess the color of the first drawn ball for the other participant. Before making your decision, you will be informed about the (for both identical) initial position and the colors of your drawn balls.

Chapter 3

Beliefs about oneself: Overconfidence and group composition in investment decisions

Baiba Renerte, Jan Hausfeld & Torsten Twardawski

Abstract

The focus of overconfidence studies has mostly been on individual decision making, even though important economic decisions are often made by groups rather than individuals. To examine whether groups make less optimal investments in risky prospects when more overconfident individuals are involved in the decision, we design a laboratory experiment with two types of investment situations—with fixed chances of success and with performance-dependent chances of success. Our design thus allows to disentangle the overconfident group members' perceived ability to "beat the odds" of the investment situations (the overplacement component of overconfidence). We show that group investment levels increase when the group is composed of more overconfident members, but the two types of situations are treated differently. The group members decide less optimally when a possibility to "beat the odds" of success is provided, and this mindset, facilitated by communication, spills over to cases where such a possibility is no longer provided. Our results thus allow drawing policy implications for board composition with respect to overconfidence.

JEL classification: C91, C92, G41.

Keywords: motivated beliefs, overconfidence, group decisions, risky decisions, communication, laboratory experiment, behavioral finance, experimental finance.

3.1. Introduction

Overconfidence is a commonly observed motivated belief that can have real-life material consequences. For example, consider Boeing's best-selling product, the 737 Max jetliner, that recently had to be grounded worldwide (Egan, 2019). After the first fatal 737 Max crash in October 2018, Boeing severely underestimated the underlying risks and failed to invest in emergency actions. After the second crash in March 2019, the global community imposed an overall prohibition on the use of these planes (Shefrin, 2019). Now, Boeing's management teams and board are under immense pressure, not only to get the 737 Max back in the air, but also to reassure the global community about the soundness of the company's decision-making processes.

Important economic decisions are often made by groups rather than individuals. Nonetheless, the focus in the literature has mostly been on individual decisions. We address this topic by examining how overconfidence is related to both individual and group investment decisions in various types of investment situations. In particular, we examine how the individual overconfidence of one or more group members affects the collective decision making of the whole group.

There are several ways how overconfidence in a group environment can differ from individual overconfidence. On the one hand, the group could mitigate individual overconfidence due to, for example, countering or compromising in the negotiation process. On the other hand, several group decision-making phenomena, such as groupthink, irrational exuberance or risky shift could aggravate the effects of individual overconfidence (Bénabou, 2012, 2015). It is therefore unclear *ex ante*, how the effects of overconfidence would differ in individual compared to group decision making. Throughout this study, we define overconfidence as an excessive belief in one's own judgment or abilities, namely as a difference between one's confidence and actual performance.

In order to better understand the role of overconfidence in investment decision making, we attempt to bridge the psychology literature on motivated beliefs and the finance literature on investment decision making. According to the psychology literature, people's preferences can motivate and affect their beliefs, leading to beliefs that feel objective but are in fact biased (Epley and Gilovich, 2016). Motivated beliefs are biased beliefs that fulfill one's psychological and functional needs, and individual overconfidence is "perhaps the most common" economically relevant manifestation of the motivated-beliefs phenomenon (Bénabou and Tirole, 2016).

According to the finance literature, individual investment decisions can be affected by individual overconfidence (Roll, 1986; Malmendier and Tate, 2005, 2008; Billett and Qian, 2008, among others). We extend this line of research to decision making in groups, such as boards of directors, and find that groups make less optimal investments in risky prospects when a higher group share of overconfident individuals are involved in the decision. Our results provide potential implications for board composition policies to improve decision making.

In addition, we systematically examine the intersection of these two branches of literature—psychology and finance—by examining not only individual but also group decision making *and* by examining not only standard but also performance-dependent investment situations. We argue that the latter type of investment situations includes a crucial over-placement component (namely, the perceived ability to "beat the odds" of the market) to reflect real-life investment situations more precisely than the commonly used standard investment situations.

To illustrate, consider a standard lottery with a 50% chance of a positive outcome, zero otherwise. In comparison, consider an equivalent lottery with a positive outcome if and only if one successfully solves a task that the general population are able to solve in 50% of cases. Would the subjective chances of success be evaluated as the same in both situations? Possibly, but likely not, due to the over-placement property of the overconfidence phenomenon.

Due to this feature of our experimental design, we can thus distinguish whether overconfident experimental board members see themselves as overly able to choose the optimal investment in terms of its success odds (the overstatement component of overconfidence) or whether they even see themselves as able to "beat the odds" of the investment (the over-placement component of overconfidence). Indeed, we find that overconfidence has a stronger effect in the latter performance-dependent situations, leading to more above-optimum investment levels than in the former standard situations.

Importantly, we examine potential channels for this effect and conclude that it is particularly pronounced in group decisions (as compared to individual decisions) and can be facilitated by within-group pre-decision communication (as compared to situations without communication). These results allow us to draw further policy implications for board composition, for avoiding bias spillovers both between the decisions and within the groups themselves. These include (i) attempting to frame decisions with objective odds of success as much as possible and even (ii) avoiding to vote on investment levels in the post-communication stages of the board meetings.

The rest of this chapter is structured as follows. We start by introducing related studies in section 3.2 and our experimental design in section 3.3, where we juxtapose individual and group investment decisions as well as objective and performance-dependent investment situations. We further introduce our results in three parts. In section 3.4, we introduce our main *group* investment results with respect to group overconfidence. In section 3.5, we compare these to the *individual* investment results with respect to individual overconfidence. We conclude with treatment analysis in section 3.6 as well as final remarks in section 3.7.

3.2. Related literature

A long-established finding in the literature is that most individuals are overconfident about their own relative judgments and abilities (Lichtenstein and Fischhoff, 1977; Weinstein, 1980; Taylor and Brown, 1988). For example, early experimental studies show that almost 80% of respondents rate themselves in the top 50% of car drivers (Svenson, 1981). A substantial body of research in economics and finance demonstrates a link between overconfidence of individuals and faulty individual investment decision making. This phenomenon has been observed in chief executives (Roll, 1986; Malmendier and Tate, 2005, 2008; Billett and Qian, 2008) as well as private and institutional traders (Barber and Odean, 2001; Deaves et al., 2009), security analysts (Hilary and Menzly, 2006) and experiment participants (Camerer and Lovallo, 1999; Biais et al., 2005; Dittrich et al., 2005, among others).

On the other hand, group dynamics in the presence of overconfident group members remain unclear. While there have been a few studies investigating group decision making and overconfidence (Sniezek, 1992; Zarnoth and Sniezek, 1997; Kerr and Tindale, 2004), these do not consider risky investment decisions. In contrast, Kocher and Sutter (2005) and Viscusi et al. (2011), among others, do consider risky group investment decisions, while neglecting the respective overconfidence of the group members. An unanswered question in economics and finance thus remains: How are group investment levels in risky prospects linked to overconfidence of one or more group members? We aim to answer this question by examining whether and how *individual overconfidence* influences *group decision making* when it comes to two distinct types of *investment decisions*.

In terms of the research question, the empirical study by Kind and Twardawski (2016) appears to be most closely related to our experimental study. The authors examine group decision making by boards of directors. Several research articles suggest that the directors discuss details of mergers and acquisitions (M&A) investment decisions in interactive board meetings (Fama and Jensen, 1983; Hillman and Dalziel, 2003; Schwartz-Ziv and Weisbach, 2013) and show that various board characteristics (e.g., gender composition) affect the performance of the approved deals (Khorana et al., 2007; Kolasinski and Li, 2013). Accordingly, Kind and Twardawski (2016) investigate another board characteristic: overconfidence.[1] They find that the directors' individual overconfidence might be negatively related to the board's M&A investment success, although the exact mechanisms at place are not yet clear.

Note, however, that empirical studies tend to have certain drawbacks, such as a lack of important control variables, that can easily be addressed in experimental studies. In particular, for answering our research question, these control variables include: (i) individual risk preferences that might be driving the overconfidence effect (Johnson and Fowler, 2011) and (ii) various individual personality traits such as optimism that are known to correlate with overconfidence (Schaefer et al., 2004; Trevelyan, 2008).

Importantly, empirical studies also cannot disentangle and control for domain-specific overconfidence. These studies must take the types of investment situations as given, while in an experimental study, we can consider a noteworthy distinction—whether the investment decision at hand requires *judgment* about the optimal prospect or rather *abilities* to carry out the prospect. That is, we can distinguish whether the overconfident experimental board members see themselves as overly able to choose the optimal investment in terms of its success odds (the overstatement component of overconfidence, with its directional effect implied by correlates of overconfidence such as the illusion of control; see Langer, 1975, among others) or whether they even see themselves as able to "beat the odds" of the investment (the over-placement component of overconfidence; see Moore and Healy, 2008).

3.3. Experimental design

In this study, we extend the previous empirical work on overconfidence in group investment decisions by using the respective controls and testing the proposed underlying mechanisms experimentally. We construct controlled experimental settings that mimic the main elements of an interactive meeting of a board of directors. We consider two types of investment situations and two potential channels for the overconfidence effects. We hypothesize that in our controlled settings, the group composition with respect to overconfidence will be positively linked to above-optimum investment levels, such that groups with more overconfident group members decide in favor of above-optimum investments.

Hypothesis 1. *Group overconfidence and group composition with respect to overconfidence are positively linked to above-optimum investment levels.*

[1] Based on a sample of 468 deals performed by U.S. publicly-traded companies, Kind and Twardawski (2016) show that their experience-based proxy measure of board overconfidence is negatively related to abnormal stock returns upon M&A announcements and positively related to the premiums paid in such transactions.

To test Hypothesis 1, we first need to define two terms: *overconfidence* and *above-optimum investment levels*; we do so in sub-sections 3.3.1 and 3.3.2 respectively. The general structure of our experimental design is as follows. Each subject faces ten investment situations. Each situation is faced twice, first individually and then in a two-member group using the unanimity rule (i.e., using renegotiation until a unanimous decision can be reached). Five of the investment situations are standard *Objective* situations (see sub-section 3.3.2.1) and the other five are performance-dependent *Ability* situations (see sub-section 3.3.2.2). Finally, in addition to these two within-subject comparisons, we introduce two between-subjects comparisons (see sub-section 3.3.3) by varying whether subjects first face the standard or performance-dependent situations and whether they are allowed to communicate before each of the group decisions.

3.3.1. *Definition and measurement of overconfidence*

We define overconfidence as an excessive belief in one's own judgment or abilities, namely as a difference between confidence (in the said judgment or abilities) and actual performance (Klayman et al., 1999, among others). We use an established multiple-choice general-knowledge task with 18 questions, which are adjusted for neutrality to hard-easy effects, to extract a *bias score* for each subject (Michailova and Katter, 2014). For example, a question can be as follows: "Who is the author of the opera *Tosca*?" The subjects provide one of the three possible answers: G. Puccini, G. Verdi or A. Vivaldi. After choosing one of the answers, the subjects report their certainty that their answer was correct, between 33% (absolute guessing, chance level) and 100% (absolute certainty).

The bias score for each subject is calculated as the difference between the average confidence level across all questions and the proportion of the correct answers. A positive bias score represents overconfidence, a negative bias score represents under-confidence, and a bias score of zero indicates an accurately calibrated (neutral) subject.

$$bias = mean\%confidence - mean\%correct$$

We measure individual overconfidence using the bias score and afterwards check the robustness of our measure using two core properties of the overconfidence phenomenon: (i) overestimation of one's actual performance and (ii) over-placement of one's performance relative to others. Namely, we ask for the subjects' estimates of how many of the items they answered correctly (0 to 18 items) and their estimates of what performance rank they have in the session (1 to n, where n is the number of subjects in a given session).

3.3.2. *Definition and measurement of above-optimum investment levels*

We construct an investment spectrum that allows measuring the closeness to a risk-neutral optimum of the chosen investment level in each investment situation.[2] We examine risky investment situations that share two important characteristics. Firstly, the investment options in each situation are distributed non-linearly, with concave and single-peaked expected value functions. Secondly, the choices between the investment options require non-trivial reasoning. That is, the optimal choice in terms of the expected value lies either in the higher or lower middle sections of the spectrum or in one of the extremes of the spectrum.

For simplicity, we can group these situations into two categories: high-optimum situations (where it is optimal to invest a high amount) and low-optimum situations (where it is optimal to invest little or nothing). For each of the investment situations, the subjects receive an endowment of 50 monetary units (MU) and can decide how much of it to invest.

3.3.2.1. Standard Objective investment situations

Table 3.1 below illustrates exemplary investment situations. The left-hand side refers to the standard investment situations (denoted *Objective*). In these situations, the listed probabilities are fixed and correspond to the actual probabilities of success in an "objective" sense. The column (a) of Table 3.1 depicts a list of investment levels that a subject can choose from after receiving the 50 MU endowment. Each investment level corresponds to an investment in a lottery.

The lotteries' High final outcomes and Low final outcomes in columns (b) and (d) respectively are always distributed in a certain way. Namely, the Low outcome equals the endowment of 50 MU minus the investment level, and the High outcome equals the endowment of 50 MU plus some premium.

Also the corresponding high probabilities p(High) and low probabilities p(Low) in columns (c) and (e) respectively are always distributed in a certain way. If the chosen investment level is positive, the subject receives the high lottery outcome with a probability p(High) and the low lottery outcome with a probability $1 - p$(High). If the chosen investment level is zero, the subject can keep the endowment of 50 MU with certainty.

[2] In this study, we mainly focus on examining the closeness to a risk-neutral optimum instead of a risk-preference adjusted optimum. We argue that this leads to a cleaner comparison between the investment situations, given that it is unclear whether risk preferences are expressed in the same way in the *Objective* situations compared to the *Ability* situations.

Objective investment situations					*Ability* investment situations				
Invest	High	*p*(High)	Low	*p*(Low)	Invest	High	Low	*p*(Q)	EV
0			50		0		50		50
5	55	90%	45	10%	5	55	45	90%	54
10	60	80%	40	20%	10	60	40	80%	56
15	65	70%	35	30%	15	65	35	70%	56
20	70	60%	30	40%	20	70	30	60%	54
25	75	50%	25	50%	25	75	25	50%	50
30	80	40%	20	60%	30	80	20	40%	44
35	85	30%	15	70%	35	85	15	30%	36
40	90	20%	10	80%	40	90	10	20%	26
45	95	10%	5	90%	45	95	5	10%	14
(a)	(b)	(c)	(d)	(e)	(a)	(b)	(d)	(f)	(g)

Table 3.1: An example of comparable investment situations in *Objective* and *Ability* treatments, where the columns (a), (b), (d) and (g) are all measured in monetary units. Note that the columns (b) and (d) denote the respective High final outcomes and Low final outcomes. The information in column (g) was not visible to the subjects.

By keeping these distribution rules fixed but changing the outcomes and probabilities, we create the concave expected value functions and the low-optimum and high-optimum situations. In the example investment situation depicted in Table 3.1, the optimum investment in terms of the expected value (EV) in column (g) is 10 or 15 MU, such that this situation is classified as low-optimum.[3]

We compare how such investment decisions are made individually with how they are made in groups with various overconfidence compositions. For the group investment decisions, we keep the group size of two and the (renegotiated) unanimity decision rule fixed.

3.3.2.2. *Performance-dependent Ability investment situations*

The subjects play half of the rounds with the described *Objective* type of investment situations with fixed (objective) odds of success. The other half of the rounds, however, the subjects play with a second type of investment situations that differ in the way how the odds of success are portrayed. The performance-dependent investment situations (denoted *Ability*) are depicted on the right-hand side of Table 3.1 above. In these investment situa-

[3]Appendix A includes a list of all investment situations and an overview of general task comprehension results. In addition, Appendix F includes translated instructions with accompanying example screenshots from the experiment sessions.

tions, we allow the subjects to potentially "beat the odds" of the lotteries. By doing so, we mimic the willingness and perceived ability of the board directors to "beat the market" in real-life investment decisions.

The *Ability* treatment with performance-dependent probabilities is a unique feature of our design. We add a task to each of the investment levels, and the probabilities in column (f) indicate how easy or difficult the associated task will be. In other words, instead of receiving the High final outcome with the respective given probability as in the *Objective* investment situations, the subjects in the *Ability* treatment receive the High final outcome if they successfully do a task for which we know from a large sample of general population what the task success chances in fact are.[4]

For example, the task associated with the 5 MU investment was successfully answered by 90% of the large general population sample, as indicated by the column (f) in Table 3.1. To receive the High final outcome associated with this investment level, our subjects need to answer the same question as well. If the subjects answer correctly, they receive the High outcome. If not, they receive the Low outcome. The probabilities in the column (f) of the *Ability* treatment thus correspond to the probabilities in the column (c) of the *Objective* treatment, as it shows how often, on average, the High outcome is achieved. Note that in the individual decisions, each subject does the task herself. In the group decisions, however, a Volunteer (one of the two group members) does the task.[5]

This experimental design allows us to create directly comparable investment situations in the *Objective* and *Ability* treatments. The probabilities are equal for neutrally calibrated subjects who think that they are neither better nor worse than the general population. However, given the over-placement property of overconfidence, the perceived probabilities in the *Objective* and *Ability* treatments might differ.

We can thus add another within-subject hypothesis. While we anticipate over-investment in line with the group composition with respect to overconfidence (the more overconfidence, the higher the investments) in both *Objective* and *Ability* treatments, we expect the *Ability* treatment to show the effect more strongly. This follows from the differences in the investment situations: While the *Objective* situations reveal the overstatement component of overconfidence, the *Ability* situations also also trigger the over-placement component.

[4]We use difficulty data from a popular TV show in Germany called "QuizDuell", where everyone watching the show can vote on their preferred answers to the displayed multiple-choice questions in a mobile application to compete with the guest in the studio. After the vote, the viewers can see how many percent of the viewers voted for each of the multiple-choice options, including the right answer. Note that our student sample appears to perform similarly as (or even slightly worse than) this general population sample, both in individual tasks and in group tasks performed by the Volunteers, thus ruling out comparability or self-selection concerns.

[5]The Volunteer for each two-person group is chosen in a preference-consistent way before the first group decision of each group, according to ranked self-reports of the willingness to be the Volunteer.

Hypothesis 2. *The effect in Hypothesis 1 is more pronounced in the Ability treatment than in the Objective treatment.*

3.3.3. Selected channels for the overconfidence effects

Let us recap what within-subject comparisons have been introduced up to now. Half of the investment decisions are with *Objective* probabilities and the other half (counter-balanced) with *Ability*-based probabilities. Of these halves respectively, each investment decision is made twice, first individually and then in a two-member group using the unanimity rule. In addition to these within-subject comparisons, we make the following 2×2 between-subjects comparisons.

We examine two channels for how individual overconfidence might affect group decisions. Firstly, we add a treatment with pre-decision communication in a free-text chat format (*Comm* treatment) to consider the persuasion effects due to the "behavioral signature" of the overconfident group members. Previous studies have shown that overconfident persons exhibit characteristics that appear like competence to others (Anderson et al., 2012). We thus give the subjects a chance to exhibit such influence in a chat environment.

Secondly, by changing the order of the *Objective* and *Ability* investment decision blocks (*Order* treatment), we consider spillover effects of the subjects' mindsets associated with the respective investment types. For example, if the subjects start with the *Ability* investment situations, they might possibly continue with the perceived ability to beat the odds also in the *Objective* investment situations and invest higher above the optimum than otherwise. Spillovers are thus defined as a type of inertia in this context.

These treatments allow us to add a third hypothesis about the described spillover effects. Given that communication allows the more overconfident group members to reveal their "behavioral signature" and exert more influence on the group decision-making process, we expect the spillover effects to be more pronounced in interactions with pre-decision communication.

Hypothesis 3. *The Ability treatment creates spillover effects on the Objective treatment. Due to the influence of the more overconfident group members, these effects are facilitated by pre-decision communication.*

3.3.4. Summary of the experimental design

Figure 3.1 below provides a summary of the experimental design with the between-subjects treatments <u>Comm</u> and <u>Order</u> underlined. We start by measuring the subjects' overconfidence (using the bias score by Michailova and Katter (2014), as discussed in sub-section

3.3.1 above) and risk preferences using the multiple price list by Holt and Laury (2002). Afterwards, the subjects perform the main experiment stages with the *Objective* and *Ability* investment tasks (as discussed in sub-section 3.3.2 above) in two separate treatment blocks. In the first block, each investment situation is faced first individually, then in a group, then the next again individually, then again in the same group, and so on. Afterwards, the same procedure follows in the second block, after random group rematching between the blocks.

Note that the group changes between the *Objective* and *Ability* treatments, and the outcome values are jittered[6] in the second block. Depending on the treatment (as discussed in sub-section 3.3.3 above), the subjects begin either with the *Objective* or the *Ability* block and, depending on the treatment, the subjects either do or do not have a pre-decision communication stage before each group decision.

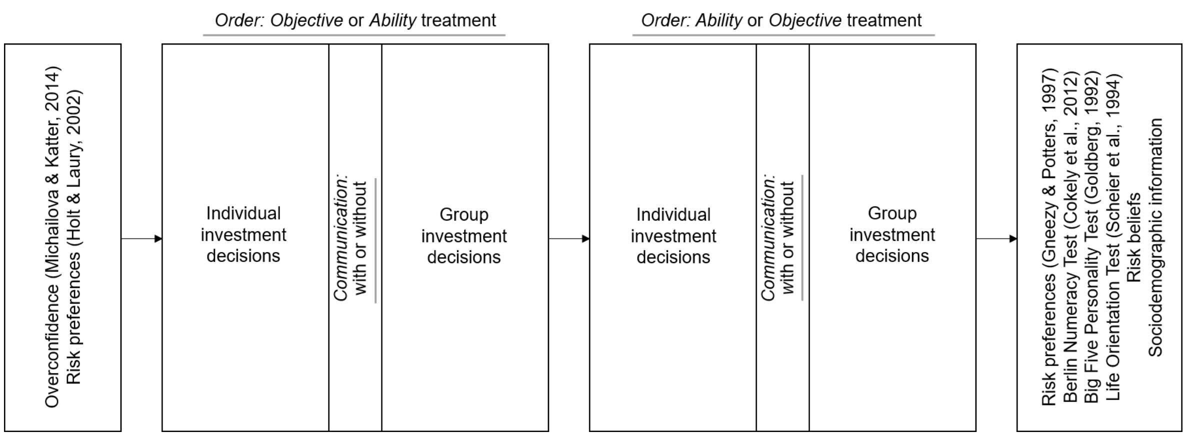


Fig. 3.1. Summary of the experimental design, with the between-subjects treatments underlined.

After the main stages, we again measure the subjects' risk preferences using a single-item investment game measure (Gneezy and Potters, 1997; Charness and Gneezy, 2010). In the post-experiment questionnaire, we consider the following tests for control purposes: the four-item Berlin Numeracy Test (Cokely et al., 2012) that we use in addition to the self-reported school grades to measure competence, the ten-item Life Orientation Test to measure optimism that is known to correlate with overconfidence (Scheier et al., 1994) and the ten-item Big Five Personality Test (Goldberg, 1992; Gosling et al., 2003) to measure other personality traits that can be related to overconfidence. We also consider the subjects' self-reported perceived success chances and demographic variables.

The subjects are incentivized as follows. In the overconfidence measurement, they are paid 30 MU if they answered a randomly selected item correctly. In the investment decisions,

[6]Namely, to avoid direct repetition between the blocks for the subjects, all values in the decision situations of the second block are 1 to 2 MU higher or lower.

the subjects are paid up to 140 MU for one randomly drawn investment decision, which can be drawn either from the *Objective* or *Ability* block and can be either an individual or group decision. It is then played out in accordance to the provided probabilities or the provided answers to the performance-dependent task. In the two risk preference measurements, the subjects are paid up to 40 MU and 25 MU respectively for their chosen investments. Each MU is worth €0.12 such that 100 MU are equal to €12.00. The subjects are additionally paid a show-up fee of €5.00.

3.3.5. Procedure

We used z-Tree software (Fischbacher, 2007) and ORSEE recruitment platform (Greiner, 2015) for the experiment with student subjects at the Lakelab in Konstanz, Germany. We gathered a dataset on $n = 168$ subjects over six experiment sessions.[7] 42.5% of the subjects were male, with an average age of 21.7 $\pm$ SD 2.5 years and an average payment of €17.0 $\pm$ SD 4.2 for an approximately 90 minute session.

3.4. Group results

3.4.1. Confidence bias score

In preparation for the analysis of investment behavior and overconfidence, we first present the results of the overconfidence measures. As mentioned, we started the experiment by measuring the overconfidence of our subjects using the bias score. Figure 3.2 below depicts the distribution of the scores. Note that a positive bias score represents overconfidence and a negative bias score represents under-confidence. The average and median scores among our subjects are larger than zero, in line with the previous literature: The majority of the subjects are overconfident.

Our analysis shows that not only do the highly overconfident subjects in our sample report higher confidence, they also perform worse (Wilcoxon-Mann-Whitney rank sum tests, $p < 0.001$) such that their overconfidence is "unjustified", as depicted in Figure 3.3 below. For the purposes of illustrative tests and figures, we divide the subjects at median to create two categories: subjects with high overconfidence (HOC) and subjects with very low overconfidence or under-confidence (LOC).[8]

[7]Note that due to technical difficulties, we lack crucial questionnaire data for eight subjects. The main results are robust to including these subjects in the dataset.

[8]Note that our main results largely hold also if we alternatively divide the sample at zero overconfidence, with only strictly overconfident or strictly under-confident subjects.

Furthermore, remember that we also ask the subjects two robustness questions concerning their estimates of how many of the items they answered correctly and their estimates of what performance rank they have in the session. As depicted in Figure 3.4 below, these overconfidence robustness measures that check the two main properties of the overconfidence phenomenon—overestimation and over-placement—strongly correlate with the bias score (Spearman's rank-order correlations, $p < 0.001$). The HOC subjects report beliefs about significantly higher numbers of correct answers and place themselves in significantly higher ranks (Wilcoxon-Mann-Whitney rank sum tests, $p < 0.001$) than the LOC subjects.

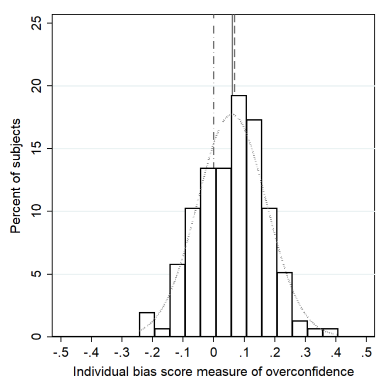


Fig. 3.2. Distribution of the confidence bias scores. The solid and dashed lines denote the mean and median scores, respectively; the dot-dashed line is at zero, and the dotted line shows a normal distribution.

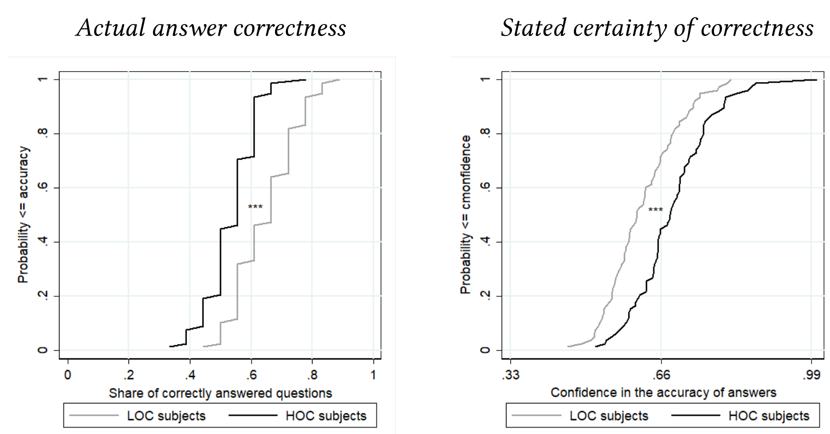


Fig. 3.3. Actual share of correct answers in the overconfidence bias measure divided by overconfidence (left) and reported certainty about the answers in the overconfidence bias measure divided by overconfidence (right). The asterisks indicate Wilcoxon-Mann-Whitney rank sum tests. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

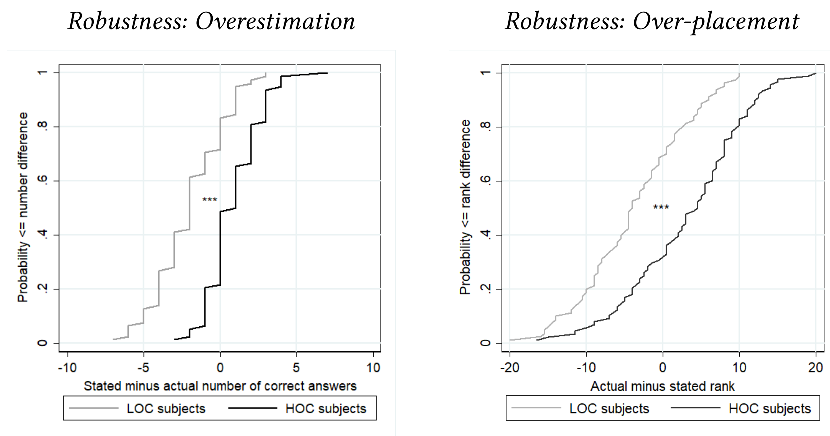


Fig. 3.4. Overconfidence robustness questions that check the two main properties of the overconfidence phenomenon—overestimation (left) and over-placement (right)—divided by overconfidence. The asterisks indicate Wilcoxon-Mann-Whitney rank sum tests. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Using the HOC and LOC categories, we can distinguish three different possible group compositions with respect to overconfidence: "LOC + LOC" and "HOC + HOC" for groups with both members exhibiting similar levels of overconfidence and "LOC + HOC" for groups that are mixed in terms of overconfidence.[9]

3.4.2. *Group confidence and investments*

Using the above measure of overconfidence, we investigate how overconfidence shapes group decision making. We find an upward trend in group decisions: The group investment levels are directly related to the group overconfidence bias score, as depicted in Figure 3.5 below, and to group composition with respect to overconfidence, as depicted in Figure 3.6 below. On average, groups with more HOC members make higher investments.

Note that there appear to be different thresholds for *Objective* and *Ability* investment situations in Figure 3.6—as we would expect given that the *Ability* group decisions are influenced by the Volunteer who is not necessarily the HOC subject in the mixed groups.[10] Accordingly, a group seems to need both HOC subjects to agree on doing a task in the *Ability* treatment, but only one HOC subject suffices to shift the group decision in the *Objective* treatment.

[9]For the purposes of further regressions, we can also use a more telling continuous average bias score per group. Details follow in sub-section 3.4.2 below.

[10]Our analysis shows that willingness to become the Volunteer for the group is not significantly related to individual overconfidence. While the subjects indicating higher willingness are more likely to be female (Wilcoxon-Mann-Whitney rank sum test, $p < 0.001$), younger ($p = 0.016$) and less competent (in terms of the bias score question correctness, $p = 0.018$, and the Berlin Numeracy Test performance, $p = 0.004$), these or other measured characteristics do not correlate with the subjects actually becoming Volunteers.

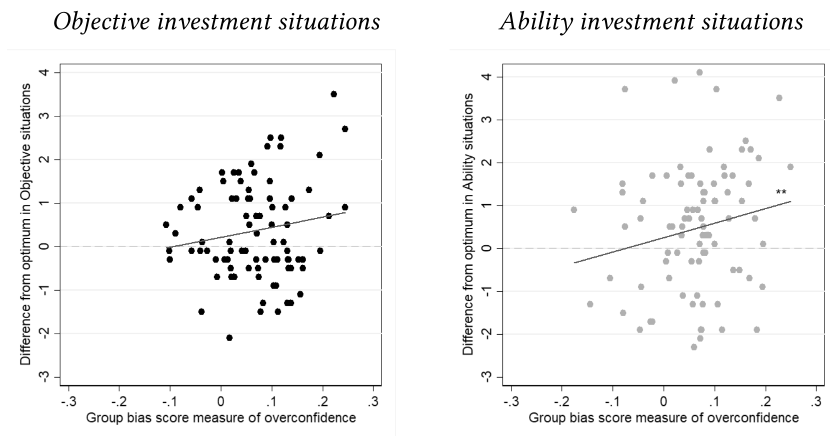


Fig. 3.5. Group investment difference from the optimum in the *Objective* (left) and *Ability* (right) treatments, by group overconfidence bias score and compared to zero (dashed). The line shows linear prediction, and the asterisks indicate Spearman's rank-ordered correlations. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

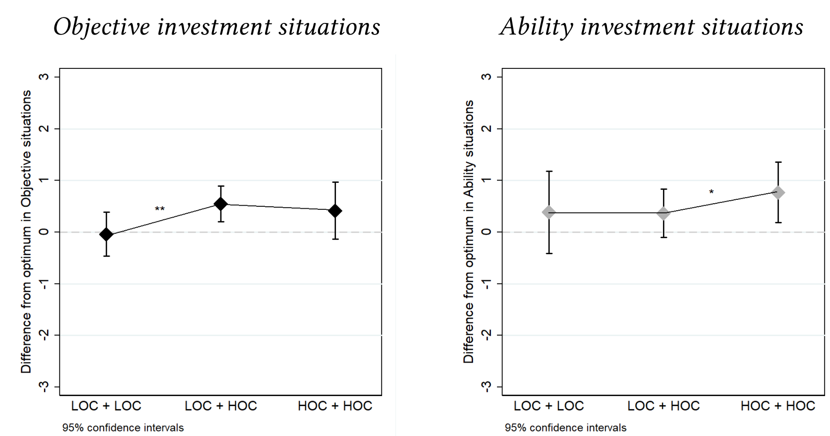


Fig. 3.6. Group investment difference from the optimum in the *Objective* (left) and *Ability* (right) treatments, by group composition with respect to overconfidence and compared to zero (dashed). The asterisks indicate Wilcoxon-Mann-Whitney signed-rank tests. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3.2 below additionally depicts regression models on group overconfidence, with the deviation from the optimum investment levels as the dependent variable. The models control for important additional factors including risk preferences (as outlined in subsection 3.3.4) and show that there is still a significant upward trend with respect to group overconfidence in the *Ability* treatment (columns 5 and 6) and a weaker trend in the *Objective* treatment (columns 1 and 2).

We perform linear ordinary least squares (OLS) regressions and correct for experiment data dependencies using robust clustered errors at the matching-group level. We use the average continuous overconfidence bias score per group as the independent variable for group overconfidence.[11] We find *some* support for Hypothesis 1: The higher the group overconfidence, the higher the average group investment levels.

Result 1. *On average, group investments tend to increase with overconfidence in the Ability treatment, while the result is inconclusive for the Objective treatment.*

	gObject	*gObject*	*gObjectH*	*gObjectL*	*gAbility*	*gAbility*	*gAbilityH*	*gAbilityL*
OCGroup	2.559	2.880	-0.947	5.432**	4.146**	3.729**	4.199**	3.415*
	(1.56)	(1.91)	(2.08)	(2.42)	(1.81)	(1.79)	(1.90)	(2.02)
Order		-0.228	-0.454	-0.078		0.547	-0.122	0.993**
		(0.34)	(0.42)	(0.40)		(0.33)	(0.41)	(0.40)
Comm		-0.475	0.105	-0.862**		0.188	-0.093	0.375
		(0.30)	(0.39)	(0.36)		(0.36)	(0.50)	(0.37)
OrderComm		0.835*	0.428	1.106*		-0.208	0.438	-0.639
		(0.47)	(0.63)	(0.56)		(0.56)	(0.61)	(0.66)
RiskHL	-0.541***	-0.542***	-0.612***	-0.496**	-0.465*	-0.456**	-0.173	-0.644***
	(0.17)	(0.16)	(0.17)	(0.22)	(0.24)	(0.20)	(0.26)	(0.22)
Male	0.076	0.073	0.131	0.035	0.757**	0.693**	0.721**	0.674*
	(0.22)	(0.29)	(0.26)	(0.37)	(0.28)	(0.31)	(0.32)	(0.34)
Constant	0.905	0.628	0.940	0.419	0.525	-0.274	-2.697*	1.341
	(0.94)	(1.21)	(1.48)	(1.48)	(1.35)	(1.40)	(1.43)	(1.68)
$Controls$	Yes*	Yes	Yes	Yes	Yes*	Yes	Yes	Yes
R^2	0.105	0.169	0.240	0.206	0.178	0.253	0.216	0.251
N	160	160	160	160	160	160	160	160
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)

Table 3.2: Regressions of group overconfidence bias score on average difference from the optimum in group investments (*gObjective*, *gAbility*) as the dependent variable. Columns 1-4 concern the *Objective* situations, while columns 5-8 concern the *Ability* situations. We consider two models and two sub-models for each. Columns 1 and 5 include the baseline model, and columns 2 and 6 additionally include the treatment dummies and their interaction. Columns 3 and 7 include only low-optimum (*L*) situations, and columns 4 and 8 include only high-optimum (*H*) situations. Note: OLS regressions, robust standard errors in parentheses clustered on matching-group level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

[11] Appendix B includes full regression models including various compositions of control variables (e.g., individual characteristics and individual competence indicators) and various optimum investment levels. Note that we find gender effects in line with previous literature (Barber and Odean, 2001, among others): Male subjects in our sample tend to be more overconfident than female subjects.

The overconfidence effect is more pronounced in the *Ability* situations than in the *Objective* situations. The reason appears to be related to how differently high-optimum situations (where it is optimal to invest a high amount) and low-optimum situations (where it is optimal to invest little or nothing) are being evaluated. While higher group overconfidence leads to higher investments in both high-optimum and low-optimum investment cases in the *Ability* situations (columns 7 and 8 in Table 3.2 above), the relation between overconfidence and investment levels is less straight-forward in the *Objective* situations. Namely, in the *Objective* situations, higher group overconfidence leads to higher investment levels only in low-optimum situations, and higher group overconfidence leads to somewhat lower investment levels in high-optimum situations (columns 3 and 4).

In other words, more overconfidence generally leads to less optimal investments, and there appear to be only selected situations (e.g., high-optimum *Ability*-type of situations) where overconfidence could actually have a positive effect on investment optimality. In Figures 3.7 and 3.8 below, we divide the average investment levels in Figures 3.5 and 3.6 by the optimum investment levels, showing fitted trends for low-optimum investment situations compared to high-optimum investment situations. We conclude that these conflicting trends explain why Result 1 is more pronounced for the *Ability* situations than for the *Objective* situations. We thus find support for Hypothesis 2 and conclude that the *Ability* situations allow more pronounced overconfidence effects.

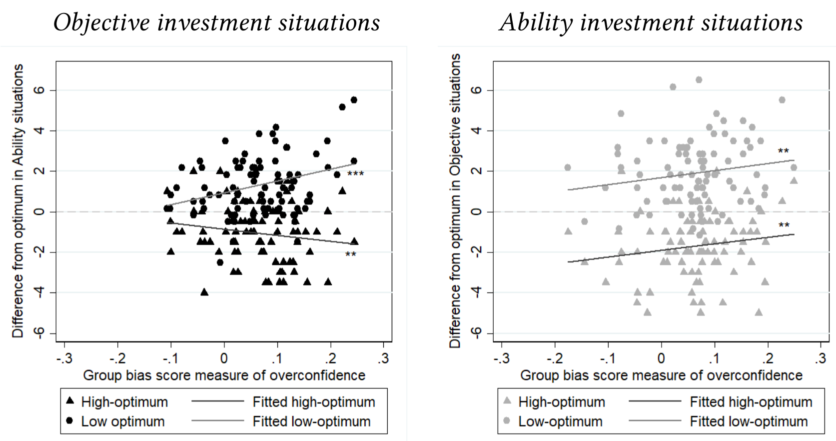


Fig. 3.7. Group investment difference from the optimum in the *Objective* (left) and *Ability* (right) treatments with respect to group overconfidence *and* by optimum investment levels (high-optimum triangle, low-optimum circle), compared to zero (dashed). Note that the subjects invest rather too little in all high-optimum situations, possibly due to ceiling effects or interactions with hard-easy effects. The lines show linear predictions, and the asterisks indicate Spearman's rank-ordered correlations. $^{***}\ p < 0.01$, $^{**}\ p < 0.05$, $^{*}\ p < 0.1$.

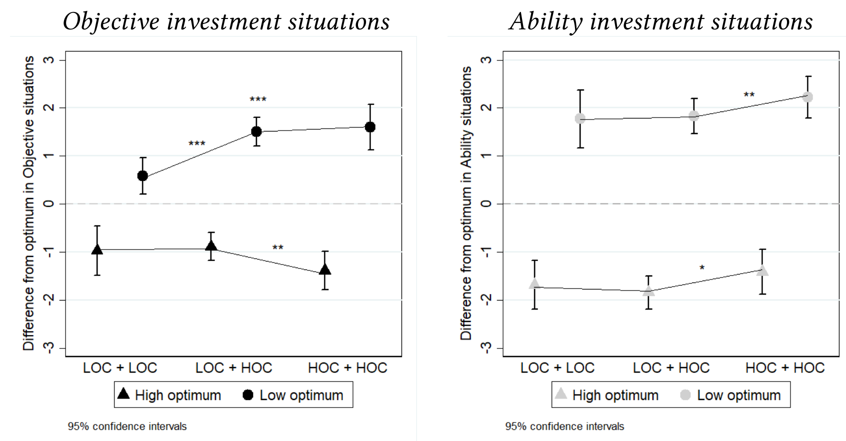


Fig. 3.8. Group investment difference from the optimum in the *Objective* (left) and *Ability* (right) treatments, divided by group composition with respect to overconfidence *and* by optimum investment levels (high-optimum triangle, low-optimum circle), compared to zero (dashed). The asterisks indicate Wilcoxon-Mann-Whitney signed-rank tests. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Result 2. *The overconfidence effect that drives investment levels upwards is more pronounced in the Ability treatment than in the Objective treatment, as it is driven only by the low-optimum situations in the Objective treatment and by all situations in the Ability treatment.*

3.5. Comparison results

3.5.1. Individual confidence and investments

In order to obtain a more comprehensive view on how overconfidence influences investment decision making, we additionally examine how *individual* overconfidence influences *individual* investment decisions. Regression results in Table 3.3 below show that higher individual overconfidence is associated with higher investment levels in both the *Objective* and *Ability* treatments. We again perform linear ordinary least squares (OLS) regressions with the difference from optimum investment levels as the dependent variable and correct for experiment data dependencies using robust clustered errors at the matching-group level. This time, we use the individual continuous bias score as the independent variable for *individual* overconfidence. Our regression models that control for the important additional factors as outlined in sub-section 3.3.4 show that the HOC subjects invest significantly higher than the LOC subjects (columns 1, 2, 5 and 6), especially when it is optimal to invest little or nothing at all (columns 4 and 8).[12]

[12] Appendix C includes full regression models including various compositions of control variables and various optimum investment levels.

We conclude that overconfidence does indeed drive higher individual investments, and in contrast to the *group* decisions in sub-section 3.4.2 where only *Ability* situations show a clear overconfidence effect, *individual* decisions show this effect in both the *Objective* and *Ability* situations. In addition to the results from the group decisions in section 3.4, we can thus also replicate some of the previous findings regarding faulty individual decision making of e.g., chief executives (Roll, 1986; Malmendier and Tate, 2005, for example).

	iObjective	*iObject*	*iObjectH*	*iObjectL*	*iAbility*	*iAbility*	*iAbilityH*	*iAbilityL*
IndivOC	1.880**	1.584*	1.227	1.822*	2.988***	2.663**	2.162	2.997***
	(0.86)	(0.82)	(1.07)	(1.00)	(0.93)	(1.04)	(1.32)	(1.10)
Order		-0.161	-0.373	-0.019		0.392	0.186	0.530*
		(0.30)	(0.36)	(0.34)		(0.25)	(0.31)	(0.28)
Comm		-0.496*	-0.148	-0.727**		-0.107	-0.043	-0.149
		(0.28)	(0.34)	(0.32)		(0.30)	(0.40)	(0.32)
OrderComm		0.546	0.471	0.596		-0.179	-0.158	-0.193
		(0.42)	(0.54)	(0.47)		(0.43)	(0.55)	(0.46)
RiskHL	-0.811***	-0.856***	-0.857***	-0.855***	-0.408*	-0.423*	-0.317	-0.494**
	(0.22)	(0.23)	(0.28)	(0.27)	(0.23)	(0.22)	(0.28)	(0.24)
Male	0.556*	0.678**	0.900**	0.531	1.022***	0.968**	1.083**	0.892**
	(0.31)	(0.31)	(0.36)	(0.37)	(0.31)	(0.39)	(0.43)	(0.41)
Constant	1.870	0.267	0.611	0.037	0.232	-1.047	-3.414*	0.531
	(1.17)	(1.55)	(1.93)	(1.63)	(1.23)	(1.39)	(1.74)	(1.43)
Controls	Yes*	Yes	Yes	Yes	Yes*	Yes	Yes	Yes
R^2	0.233	0.294	0.260	0.273	0.287	0.331	0.296	0.279
N	160	160	160	160	160	160	160	160
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)

Table 3.3: Regressions of individual overconfidence bias score on average difference from the optimum in *individual* investments (*iObjective*, *iAbility*) as the dependent variable. Columns 1-4 concern the *Objective* situations, while columns 5-8 concern the *Ability* situations. We consider two models and two sub-models for each. Columns 1 and 5 include the baseline model, and columns 2 and 6 additionally include the treatment dummies and their interaction. Columns 3 and 7 include only low-optimum (*L*) situations, and columns 4 and 8 include only high-optimum (*H*) situations. The differences between columns 1 and 2 as well as 5 and 6 are equivalent to those in Table 3.2. Note: OLS regressions, robust standard errors in parentheses clustered on matching-group level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Result 3. *Individual investments tend to increase with overconfidence in both Objective and Ability treatments.*

Result 4. *The effect of overconfidence for individual investments is more pronounced in the Ability treatment than in the Objective treatment.*

3.5.2. Individual and group confidence and investments

After examining also individual investment decisions alone, Figure 3.9 below compares the average individual and group investment levels. We find "risky shift" differences in the *Objective* treatment (Wilcoxon-Mann-Whitney signed-rank tests, $p = 0.003$), in line with some previous findings in the literature (Collins and Guetzkow, 1964, for example), while the results are inconclusive in the *Ability* treatment. The risky shift phenomenon is a tendency for people to make riskier decisions when they are in groups than when they are alone.

Interestingly, for the individual decisions depicted in Figure 3.9, the risky shift effect appears to be of equivalent magnitude to the Ability-versus-Objective treatment effect. Namely, there is no significant difference between average *group* investment levels in the *Objective* treatment and average *individual* investment levels in the *Ability* treatment. As depicted in Appendix D, the results of this sub-section are consistent for both the HOC and LOC subjects similarly.

All investment situations

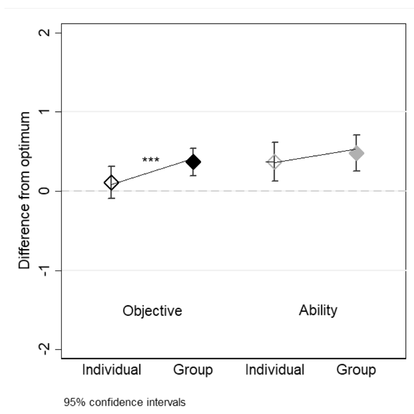


Fig. 3.9. Average individual and group investment difference from the optimum in *Objective* and *Ability* treatments, compared to the average optimum investment level (dashed). The asterisks indicate Wilcoxon-Mann-Whitney signed-rank tests. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

In Figure 3.10 below, we divide the average investment levels in Figure 3.9 by the optimum investment levels: low-optimum situations compared to high-optimum situations. Notice that, compared to Figure 3.9 above, significant differences between the *Objective* and *Ability* treatments appear. Namely, the *Ability* treatment drives both individual and group investment levels further away from the optimum compared to the *Objective* treatment. This is the case both for low-optimum investment situations (*Ability* investments are higher above optimum than *Objective* investments) and high-optimum investment situations (*Ability* investments are lower below optimum than *Objective* investments).

In contrast, notice also that Figure 3.10 shows the risky shift phenomenon remaining consistently *upwards* for both low-optimum situations (further from the optimum) and high-optimum situations (*closer* to the optimum). We can conclude that the risky shift effect in Figure 3.9 is thus not due to mere compromising, as this would bring the investment levels closer to median in both cases.

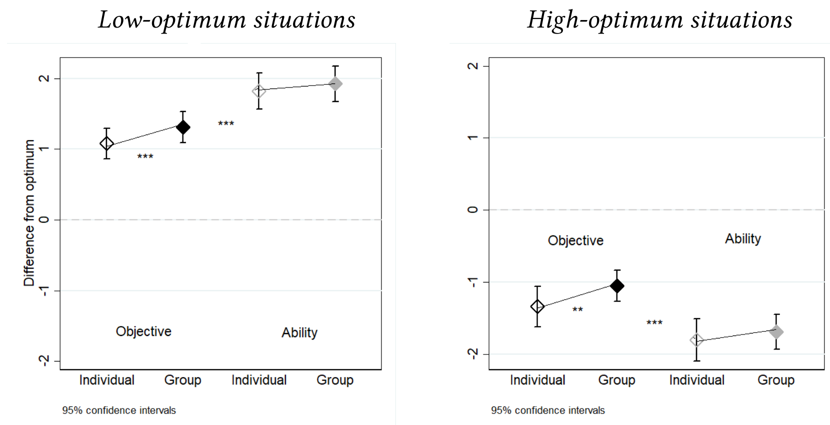


Fig. 3.10. Difference from the optimum of individual and group investments in *Objective* and *Ability* treatments compared to the optimum average investment level (dashed), divided by optimum investment levels: low-optimum investment situations in the left sub-figure, high-optimum investment situations in the right sub-figure. The asterisks indicate Wilcoxon-Mann-Whitney signed-rank tests. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Result 5. *The Ability investment situations drive the investment levels further away from the optimum than the Objective investment situations.*

3.6. Intervention results

3.6.1. Order and Communication channels

After summarizing the results of the within-subject comparisons (*Objective* versus *Ability* situations, and individual versus group investments), we now turn to the between-subjects comparisons. As outlined in sub-section 3.3.3, we check for differences in response to (i) the main channel variation, which alternates whether the subjects start with the *Ability* or the *Objective* investment situations (*Order* treatment), as well as (ii) pre-decision communication in a chat format (*Comm* treatment).

As illustrated in Figure 3.11 below, we find that group investment levels throughout the session are on average higher whenever the *Ability* treatment is the first one in a session.

Importantly, as depicted in Appendix E, the effect appears to be driven particularly by the "HOC + HOC" groups consisting of two highly overconfident group members.

If we additionally divide the results in Figure 3.11 by pre-decision communication, as in Figure 3.12 below, we see that communication is a mediator for this effect. There is a clear upward shift that we see in the figures (as well as in the regressions with clustered robust standard errors and various controls in the Table 3.2 above). We thus find some support for Hypothesis 3 and conclude that we can identify spillover effects for the "beat the odds" mindset on group investment decisions.

Investment levels by Order treatment

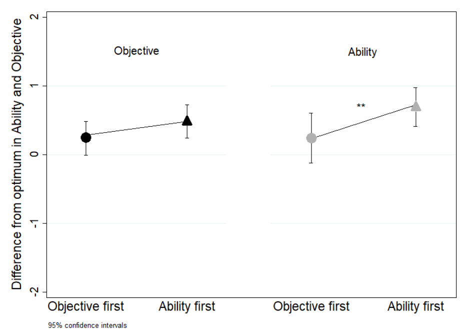


Fig. 3.11. Group investment difference from the optimum in the *Objective* and *Ability* treatments, divided by the *Order* treatment (circles for the sessions with *Objective* block first, triangles for the sessions with *Ability* block first). The asterisks indicate Wilcoxon-Mann-Whitney signed-rank tests. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Investment levels by Order and Comm treatments

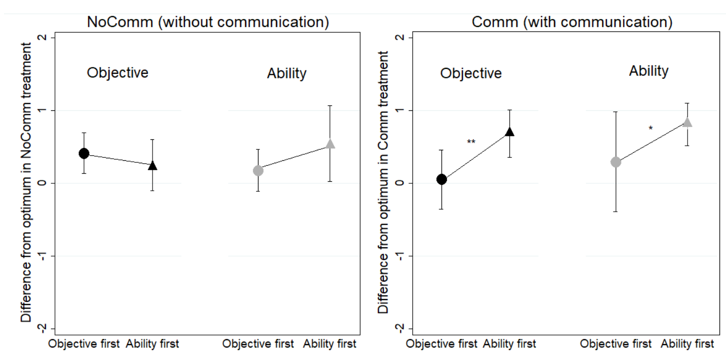


Fig. 3.12. Group investment difference from the optimum in the *Objective* and *Ability* treatments, divided by the *Order* and *Comm* treatments (circles for the sessions with *Objective* block first, triangles for the sessions with *Ability* block first). The asterisks indicate Wilcoxon-Mann-Whitney signed-rank tests. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Result 6. *The Ability investment situations at the beginning of the session create spillovers: in the subsequent decisions of the session, the Objective investment levels are comparatively higher.*

3.6.2. Communication analysis

Indeed, a detailed analysis of the pre-decision communication provides further insights in the underlying mechanisms behind the Result 6 above.[13] Firstly, we find objective differences in how the communication stage is used: the differences prevail both between the HOC and LOC subjects and between the *Objective* and *Ability* treatments. In Figure 3.13 below, we show that the HOC subjects are the first ones to talk significantly more often than the LOC subjects in the pre-decision communication stages of the mixed HOC + LOC groups in the *Ability* situations (Wilcoxon-Mann-Whitney rank sum test, $p = 0.003$), while we find no such difference in the *Objective* situations.

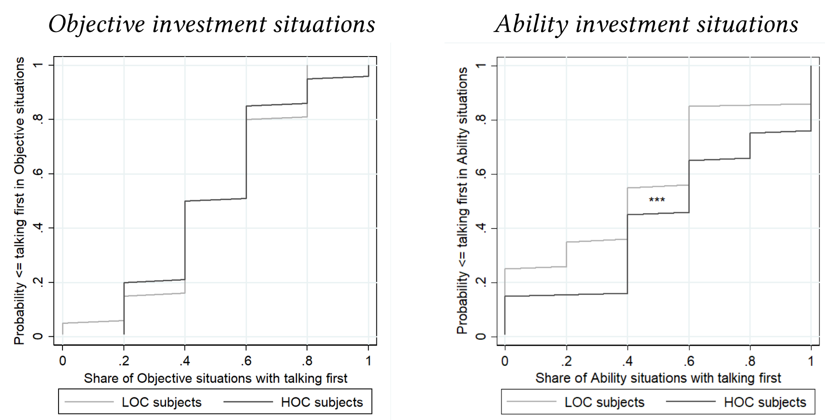


Fig. 3.13. Comparison of which group member (the HOC or the LOC subject) speaks first in the pre-decision communication stages in the *Objective* (left) and *Ability* (right) decision situations. The asterisks indicate Wilcoxon-Mann-Whitney rank sum tests. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

We can thus conclude that in certain situations, HOC and LOC subjects do indeed tend to have somewhat different "behavioral signatures". Our analysis also shows that individual overconfidence is positively correlated with individual optimism or Life Orientation ($r = 0.165, p = 0.037$) and slightly less strongly related to certain Big Five Personality traits, like conscientiousness ($r = 0.155, p = 0.051$) and openness ($r = 0.139, p = 0.080$).[14]

[13]Note that due to the concise nature of most of the observed pre-decision conversations, we mostly focus on objective measures of communication in this sub-section, e.g., we examine whether the HOC or LOC group members are the first ones to chat and the first ones to propose a certain investment level. We generally find only slight differences for the latter.

[14]Other Big Five Personality traits do not appear to be significantly related to overconfidence in our sample.

Secondly, we find significant between-subjects differences in how the other group members are evaluated in the *Comm* and *NoComm* treatments. In the post-experiment questionnaire, we asked the subjects directly to evaluate their peers' competence and likability on a Likert scale.[15] In Figure 3.14 below, we show that the evaluations are consistently more positive in the *Comm* treatment than in the *NoComm* treatment (Wilcoxon-Mann-Whitney rank sum test, $p < 0.001$). Importantly, when comparing the group compositions, this significant difference comes mainly from the mixed LOC + HOC groups (Wilcoxon-Mann-Whitney rank sum tests, $p < 0.001$ in both *Objective* and *Ability* situations) and less from the balanced LOC + LOC ($p = 0.365$ in *Objective*, $p = 0.181$ in *Ability*) and HOC + HOC ($p = 0.571$ in *Objective*, $p = 0.406$ in *Ability*) groups.

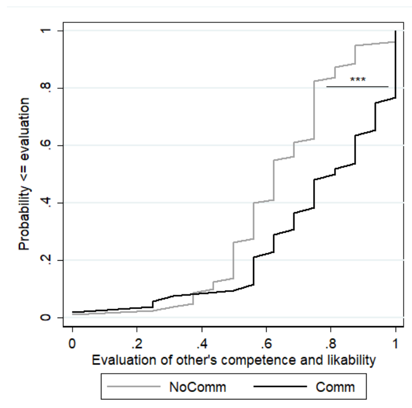


Fig. 3.14. Evaluation of the other group member's competence and likability in the post-experiment questionnaire, by whether there were pre-decision communication stages in the given session. The asterisks indicate Wilcoxon-Mann-Whitney rank sum tests. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Result 7. *The HOC subjects cast their influence on the group decisions via a qualitatively-different communication style than that of the LOC subjects. This effect is stronger in the pre-decision communication rounds before the Ability situations than before the Objective situations.*

Note, however, that this study uses very concise test versions and can generally provide little insight in the underlying processes behind the decisions of HOC and LOC group members or behind the processing modes used in the evaluation of *Objective* and *Ability* investment situations. More extensive testing as well as complementary research approaches, such as eye-tracking or mouse-tracking analyses, could thus be beneficial for further research.

[15]Note that we present aggregate results of these evaluations for conciseness purposes; the separate results of these evaluations follow the same pattern.

3.7. Discussion and conclusions

3.7.1. Interpretation of the results

What is it about the *Ability* investment situations that is so different from the *Objective* investment situations? And what is it about the HOC subjects that is so different from the LOC subjects? As a potential answer, consider the following self-reports of perceived success chances that we gathered in the post-experiment questionnaire. We asked the subjects two questions about their individual decisions that perhaps appear redundant at first: "In your opinion, how successful will your 50%-success-chance investment be in (i) the *Objective* treatment and (ii) the *Ability* treatment?" The answers were provided on an 11-point Likert scale, where the middle point is the answer we would expect in both cases.

But we find differences in the results instead. In the *Objective* treatment, as depicted in the left sub-figures of Figure 3.15 below, both HOC and LOC subjects provide answers that on average do not significantly differ from the expected 50% answer. In contrast, in the *Ability* treatment, as depicted in the right sub-figures of Figure 3.15 below, the HOC subjects are the only ones who provide an answer that significantly differs from (exceeds) the 50% answer, and the HOC subjects also provide answers that are significantly higher than the LOC subjects' answers (Wilcoxon-Mann-Whitney signed-rank tests, $p < 0.001$).

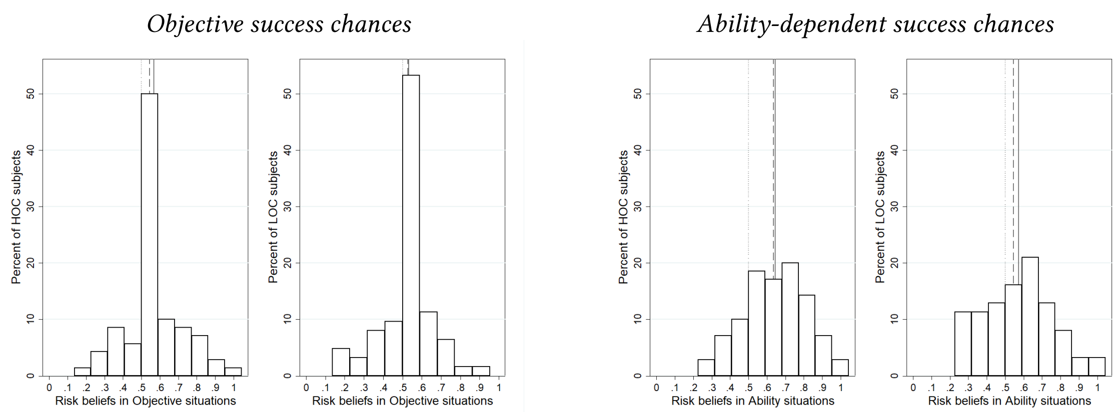


Fig. 3.15. Perceived success chances in the *Objective* (left) and *Ability* (right) treatments, divided by individual overconfidence. The solid and dashed lines denote the mean and median perceived success chances, respectively; the dotted line denotes the theoretical prediction.

In general, we can thus conclude that many aspects of the decision situations (beyond their potential outcomes and objective probabilities) should be carefully taken into account when evaluating the effects of overconfidence on investment decision making. The way how success chances of an investment opportunity are perceived can play a significant role for how much overconfidence is exhibited. Although the *Objective* and *Ability* investment situations are by construction equivalent, the decision makers tend to treat them quite differently—both in terms of how much overconfidence manifests itself in the chosen investment levels and how far from the optimum the chosen investment levels lie (as outlined in Results 1 to 5).

But the differences are not in the investment behavior alone. We find that the decision makers also learn from these situations differently and communicate about these situations differently. On average, all decision makers (both HOC and LOC types) learn to invest more "overconfidently" if first primed with the *Ability* investment situations—at least in part because the overconfident decision makers communicate more proactively in these *Ability* investment situations than in the *Objective* situations (as outlined in Results 6 to 7).

3.7.2. Concluding remarks

We have found support for our main hypotheses. Firstly, we show that the investment levels do indeed depend on the average group overconfidence or the group composition with respect to overconfidence. Groups make less optimal investments in risky prospects when, on average, more overconfident individuals or a larger group share of overconfident individuals are involved in the decisions. Overconfidence thus plays a significant role not only in individual investment decisions but also in group decisions.

Given that many important economic decisions are made by groups (such as boards of directors) not just individuals (such as chief executives), the experiment results of our study jointly with the empirical results of Kind and Twardawski (2016), among others, provide potential implications for board composition policies with respect to overconfidence. We find that higher overconfidence generally leads to less optimal investment levels, and there are indeed only limited types of investment situations (e.g., high-optimum *Ability*-type of situations) where overconfidence could actually have a positive effect on investment optimality. The first policy recommendation thus implies attempting to limit individual overconfidence in management teams and boards as much as possible.

Secondly, we also show that groups—as well as individuals—decide less optimally when a possibility to "beat the odds" of success is provided. This "beat the odds" mindset appears to spill over to cases where such possibility is no longer provided, especially in interaction with pre-decision communication. These results in combination with the result above thus emphasize the potential hazard of bias spillovers both between the decisions and within the groups themselves. Further policy implications thus include (i) attempting to frame board investment decisions with objective odds of success as much as possible or even (ii) avoiding to vote on investment levels in the post-communication stages of the board meetings.

Meanwhile, note that this study provides only partial insight into how individual overconfidence can influence group decisions, and there still remain many directions to add to this line of research. The design of this study itself could be expanded in several ways, for example, by varying the group size and the voting rule for the group decisions. Further extensions could also include different categories of tasks that the subjects need to solve in the performance-dependent situations.

Appendix A. Investment situations

Invest	High	p(High)	Low	p(Low)	EV
0			50		50
5	55	90%	45	10%	54
10	60	80%	40	20%	56
15	65	70%	35	30%	56
20	70	60%	30	40%	54
25	75	50%	25	50%	50
30	80	40%	20	60%	44
35	85	30%	15	70%	36
40	90	20%	10	80%	26
45	95	10%	5	90%	14
0			50		50
5	55	45%	45	55%	50
10	60	40%	40	60%	48
15	65	35%	35	65%	46
20	70	30%	30	70%	42
25	75	25%	25	75%	38
30	80	20%	20	80%	32
35	85	15%	15	85%	26
40	90	10%	10	90%	18
45	95	5%	5	95%	10
0			50		50
5	55	95%	45	5%	55
10	60	90%	40	10%	58
15	65	85%	35	15%	61
20	70	80%	30	20%	62
25	75	75%	25	25%	63
30	80	70%	20	30%	62
35	85	65%	15	35%	61
40	90	60%	10	40%	58
45	95	55%	5	45%	55

Invest	High	p(High)	Low	p(Low)	EV
0			50		50
1	55	80%	49	20%	54
2	60	75%	48	25%	57
3	65	70%	47	30%	60
4	70	65%	46	35%	62
5	75	60%	45	40%	63
6	80	55%	44	45%	64
7	85	50%	43	50%	64
8	90	45%	42	55%	64
9	95	40%	41	60%	63
0			50		50
5	60	95%	45	5%	59
10	70	90%	40	10%	67
15	80	85%	35	15%	73
20	90	80%	30	20%	78
25	100	75%	25	25%	81
30	110	70%	20	30%	83
35	120	65%	15	35%	83
40	130	60%	10	40%	82
45	140	55%	5	45%	79

Table A3.1: All investment situations in the (equivalent) *Objective* and *Ability* treatments: low-optimum situations in the left sub-table, high-optimum situations in the right sub-table.

Appendix B. Extended group investment regressions

	gObject	*gObject*	*gObjectH*	*gObjectL*	*gAbility*	*gAbility*	*gAbilityH*	*gAbilityL*
OCGroup	2.559	2.880	-0.947	5.432**	4.146**	3.729**	4.199**	3.415*
	(1.56)	(1.91)	(2.08)	(2.42)	(1.81)	(1.79)	(1.90)	(2.02)
Order		-0.228	-0.454	-0.078		0.547	-0.122	0.993**
		(0.34)	(0.42)	(0.40)		(0.33)	(0.41)	(0.40)
Comm		-0.475	0.105	-0.862**		0.188	-0.093	0.375
		(0.30)	(0.39)	(0.36)		(0.36)	(0.50)	(0.37)
OrderComm		0.835*	0.428	1.106*		-0.208	0.438	-0.639
		(0.47)	(0.63)	(0.56)		(0.56)	(0.61)	(0.66)
RiskHL	-0.541***	-0.542***	-0.612***	-0.496**	-0.465*	-0.456**	-0.173	-0.644***
	(0.17)	(0.16)	(0.17)	(0.22)	(0.24)	(0.20)	(0.26)	(0.22)
RiskGP	-0.126	-0.011	-0.356	0.219	-0.313	-0.313	-0.481	-0.201
	(0.43)	(0.46)	(0.45)	(0.54)	(0.51)	(0.50)	(0.57)	(0.50)
Sequence	-0.256	-0.302	-0.550*	-0.137	-0.164	-0.207	-0.157	-0.241
	(0.24)	(0.24)	(0.29)	(0.30)	(0.27)	(0.27)	(0.31)	(0.32)
Male	0.076	0.073	0.131	0.035	0.757**	0.693**	0.721**	0.674*
	(0.22)	(0.29)	(0.26)	(0.37)	(0.28)	(0.31)	(0.32)	(0.34)
Age	-0.019	-0.010	-0.042	0.012	-0.016	-0.008	0.000	-0.014
	(0.04)	(0.04)	(0.04)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)
AverageGrade	-0.093	-0.561	-1.327	-0.050	-0.387	-0.831	-1.425	-0.435
	(0.80)	(0.88)	(0.92)	(1.14)	(1.15)	(1.06)	(1.14)	(1.16)
Numeracy	0.440	0.417	0.670*	0.249	0.872*	0.729	0.963*	0.574
	(0.31)	(0.31)	(0.35)	(0.42)	(0.45)	(0.44)	(0.48)	(0.48)
Optimism		-0.834	-1.929**	-0.104		-0.986	-1.163	-0.868
		(0.73)	(0.84)	(0.88)		(0.80)	(0.80)	(0.88)
Big5Extrovert		0.061	0.526	-0.249		-1.055*	-0.830	-1.205*
		(0.53)	(0.58)	(0.66)		(0.56)	(0.65)	(0.61)
Big5Agreeable		0.726	1.630**	0.123		-0.825	-0.250	-1.208
		(0.69)	(0.75)	(0.95)		(0.91)	(0.87)	(1.05)
Big5Conscient		-0.514	-1.165**	-0.079		0.201	0.539	-0.024
		(0.44)	(0.53)	(0.49)		(0.51)	(0.55)	(0.59)
Big5Stable		0.661	0.778	0.583		1.895***	1.669**	2.046***
		(0.57)	(0.65)	(0.72)		(0.64)	(0.72)	(0.70)
Big5Open		0.396	0.196	0.529		1.513**	1.376	1.605*
		(0.55)	(0.88)	(0.64)		(0.73)	(0.83)	(0.83)
Constant	0.905	0.628	0.940	0.419	0.525	-0.274	-2.697*	1.341
	(0.94)	(1.21)	(1.48)	(1.48)	(1.35)	(1.40)	(1.43)	(1.68)
R^2	0.105	0.169	0.240	0.206	0.178	0.253	0.216	0.251
N	160	160	160	160	160	160	160	160

Table A3.2: Full regressions of group overconfidence on the average difference from optimum in group investment levels, two models and two sub-models. Note: OLS regressions, robust standard errors in parentheses, clustered on matching-group level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix C. Extended individual investment regressions

	iObject	*iObject*	*iObjectH*	*iObjectL*	*iAbility*	*iAbility*	*iAbilityH*	*iAbilityL*
IndivOC	1.880**	1.584*	1.227	1.822*	2.988***	2.663**	2.162	2.997***
	(0.86)	(0.82)	(1.07)	(1.00)	(0.93)	(1.04)	(1.32)	(1.10)
Order		-0.161	-0.373	-0.019		0.392	0.186	0.530*
		(0.30)	(0.36)	(0.34)		(0.25)	(0.31)	(0.28)
Comm		-0.496*	-0.148	-0.727**		-0.107	-0.043	-0.149
		(0.28)	(0.34)	(0.32)		(0.30)	(0.40)	(0.32)
OrderComm		0.546	0.471	0.596		-0.179	-0.158	-0.193
		(0.42)	(0.54)	(0.47)		(0.43)	(0.55)	(0.46)
RiskHL	-0.811***	-0.856***	-0.857***	-0.855***	-0.408*	-0.423*	-0.317	-0.494**
	(0.22)	(0.23)	(0.28)	(0.27)	(0.23)	(0.22)	(0.28)	(0.24)
RiskGP	-0.648	-0.477	-1.065*	-0.085	-0.828*	-0.617	-0.925	-0.412
	(0.48)	(0.47)	(0.56)	(0.51)	(0.43)	(0.46)	(0.60)	(0.47)
Numeracy	0.181	0.222	1.231**	-0.451	1.112*	1.091*	1.741***	0.658
	(0.34)	(0.33)	(0.48)	(0.34)	(0.56)	(0.57)	(0.63)	(0.58)
Sequence	-0.204	-0.146	-0.170	-0.130	-0.036	-0.019	0.295	-0.228
	(0.20)	(0.20)	(0.26)	(0.23)	(0.22)	(0.22)	(0.29)	(0.23)
Male	0.556*	0.678**	0.900**	0.531	1.022***	0.968**	1.083**	0.892**
	(0.31)	(0.31)	(0.36)	(0.37)	(0.31)	(0.39)	(0.43)	(0.41)
Age	-0.039	-0.043	-0.093	-0.010	0.047	0.042	0.011	0.062
	(0.05)	(0.05)	(0.07)	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)
AverageGrade	-1.026	-1.160	-1.333	-1.045	-2.917***	-2.912**	-2.386*	-3.262***
	(0.88)	(0.87)	(0.97)	(1.02)	(0.99)	(1.09)	(1.29)	(1.20)
Optimism		0.981	0.485	1.311		1.432	1.701	1.253
		(0.88)	(1.12)	(0.92)		(1.23)	(1.61)	(1.21)
Big5Extrovert		0.037	0.308	-0.144		-0.378	0.297	-0.828
		(0.54)	(0.73)	(0.60)		(0.68)	(0.94)	(0.67)
Big5Agreeable		1.032	2.080**	0.333		-0.660	-0.437	-0.808
		(0.66)	(0.95)	(0.70)		(0.93)	(1.08)	(1.01)
Big5Conscient		-0.313	-0.561	-0.147		-0.786	-1.208	-0.504
		(0.55)	(0.76)	(0.57)		(0.73)	(0.86)	(0.81)
Big5Stable		-0.099	-0.983	0.490		0.731	0.449	0.919
		(0.62)	(0.67)	(0.73)		(0.75)	(0.97)	(0.87)
Big5Open		0.863	0.049	1.406**		1.349*	1.560	1.208
		(0.59)	(0.87)	(0.58)		(0.76)	(0.96)	(0.81)
Constant	1.870	0.267	0.611	0.037	0.232	-1.047	-3.414*	0.531
	(1.17)	(1.55)	(1.93)	(1.63)	(1.23)	(1.39)	(1.74)	(1.43)
R^2	0.233	0.294	0.260	0.273	0.287	0.331	0.296	0.279
N	160	160	160	160	160	160	160	160

Table A3.3: Full regressions of individual overconfidence on average difference from optimum in individual investment levels, two models and two sub-models. Note: OLS regressions, robust standard errors in parentheses, clustered on matching-group level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix D. Comparison results for HOC and LOC

The comparison results as outlined in sub-section 3.5.2 appear to be largely similar for HOC and LOC subjects. Note that Figure A3.1 corresponds to the Figure 3.9 in the main text and the sub-figures in Figure A3.2 correspond to the sub-figures of Figure 3.10 in the main text.

All investment situations

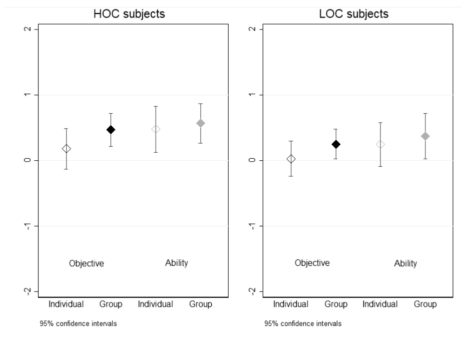
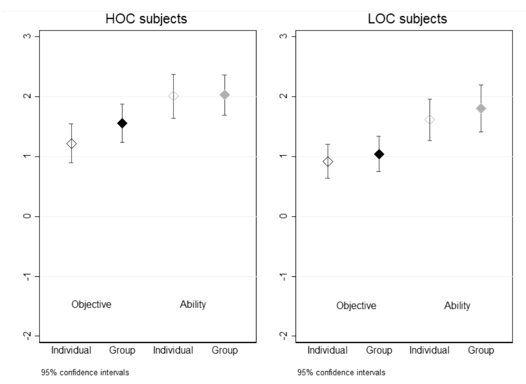


Fig. A3.1. Individual and group investment difference from the optimum in the *Objective* and *Ability* investment decisions, divided by individual overconfidence.

Low-optimum situations



High-optimum situations

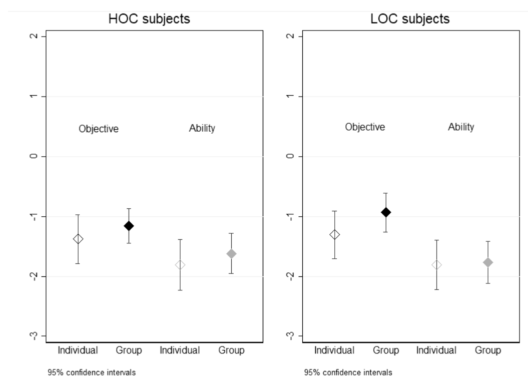


Fig. A3.2. Individual and group investment difference from the optimum in the *Objective* and *Ability* investment decisions, divided by individual overconfidence and by optimum investment levels.

Appendix E. Order treatment and group composition

The *Order* spillover effects, as outlined in sub-section 3.4.2, appear to be partly driven by the HOC + HOC groups in particular.

Objective investment situations

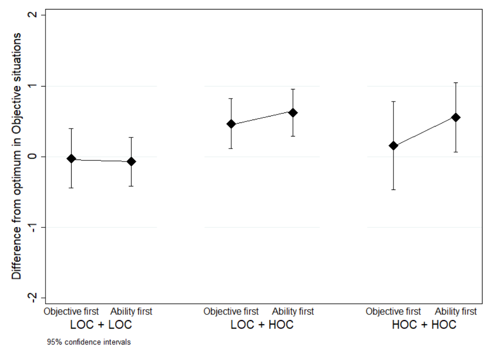


Fig. A3.3. *Order* differences in the *Objective* investment decisions, divided by group composition.

Ability investment situations

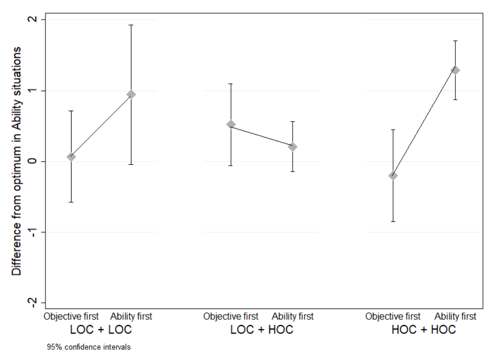


Fig. A3.4. *Order* differences in the *Ability* investment decisions, divided by group composition.

Appendix F. Instructions

The following instructions (translated from German) correspond to the treatment that (i) includes pre-decision communication (as opposed to no communication) and (ii) presents the investment situations in an order that shows the *Objective* situations before the *Ability* situations (as opposed to vice versa).

Overview

Welcome to this experiment. Please do not speak with other participants during the experiment and turn off your mobile phones and other mobile electronic devices.

To participate in today's experiment, you will be paid in cash at the end. The amount of the payout depends partly on chance and partly on your decisions. It is therefore important that you carefully read and understand the instructions.

Today's experiment consists of four parts, each comprising several rounds. At the end, several randomly drawn rounds are paid out. From Part 1, two rounds (one round from Part 1a and one round from Part 1b) will be randomly drawn and paid out. From Parts 2 and 3, one round will be randomly drawn and paid out. From Part 4, one round will be randomly drawn and paid out.

Your payout will result from the earned points in the drawn rounds. These points will be converted into euro, and you will receive additional 5 euro to complete the subsequent questionnaire. The conversion of the points into euro is done as follows. Each point is worth 12 cents, so the following applies: 100 points = 12.00 euro. Each participant is paid privately so that other participants cannot see how many points you have earned.

Setup of the experiment

This experiment consists of four different parts. Part 1 consists of Part 1a and Part 1b. Part 1a consists of 18 identical rounds. In each round, you will answer questions with three choice options, with answers that can be right or wrong. Then, in Part 1b, you will make 10 decisions in a table.

The Parts 2 and 3 have 5 + 5 identical rounds each. An investment decision has to be made in every round. The tasks in these parts are similar in structure. In Part 2, individual investment decisions and group investment decisions are made five times, in an alternating manner. Similarly, in Part 3, five individual investment decisions and five group investment decisions are made alternately. That is, you first make a first individual investment decision on a first investment situation, then a two-member group investment decision on the same investment situation, then a second individual investment decision on a different investment situation, then a second two-member group investment decision on the same second investment situation, then a third etc.

Parts 2 and 3 differ in the following way. In Part 2, the mentioned probabilities will determine the chances of success of the investment. In Part 3, certain tasks with an appropriate level of ease (easiness of task) will determine the success of the investment. The detailed instructions for Part 2 will be shown after Part 1, and detailed instructions for Part 3 will be shown after Part 2. The instructions for Part 4 (4a and 4b) will be shown after Part 3.

In summary, the sequence of this experiment is as follows: 1a, 1b, five times alternating investment decisions and group decisions in Part 2, five times alternating individual investment decisions and group decisions in Part 3, and finally 4a, 4b.

Payout

From Part 1, two rounds will be randomly drawn (one round from Part 1a and one round from Part 1b). In Part 1a, you get 20 points if you answered the drawn round (1 out of 18 answers) correctly or 0 points if you answered the drawn round incorrectly. Also in Part 1b, you will be paid out the respective drawn round (1 out of 10 decisions). The exact number of points you get in Part 1b depends partly on your decisions and partly on chance.

From the Parts 2 and 3, a total of one round (1 of 20 decisions: either an individual decision or a group decision) will be randomly drawn and paid out. You can get up to 140 points in this round. The exact number of points you get depends partly on your decisions and partly on chance. Then, from Part 4, one round is randomly drawn and paid out. Each point is worth 12 cents, so the following applies: 100 points = 12.00 euro.

Questions?

Take your time to review the instructions thoroughly. If you have questions, please raise your hand. An experimenter will come to your place.

[Quiz]

Please answer the following understanding-question about the experiment. How many rounds are paid out of Parts 2 and 3 in total? You did [not] answer the question correctly. Your answer: [...] The correct answer: [...] If you wish, go back to the instructions or ask an experimenter for explanations. Otherwise, click “Next”.

Part 1a

Welcome to Part 1a of the experiment. It includes some general-knowledge questions. Imagine that you are involved in a game like “Trivial Pursuit” or “Who Wants to Be a Millionaire?” and you have to choose the right answer from the three alternatives.

1. Please click on one of the three given answers. Only one answer is correct. You will not receive any feedback as to which answer is correct.

2. Once you have made your choice and clicked on your answer, we would like to know how confident you are that your answer is correct. Since there are three alternative answers and only one of them is correct, you have a 33% chance to respond correctly. Thus, 33% means that you are guessing and do not know the right answer. Correspondingly, 100% means that you are absolutely sure which answer is correct.

You can use any number between 33% and 100% to indicate your confidence that your answer is correct. Fill in this confidence for each answer in the gap after each question: How sure are you that your answer is correct? (33% to 100%).

Note: Please answer all questions one by one in the order in which they appear in the questionnaire. If you do not know the answers, you can guess. Please do not jump back and forth between the questions and do not go back to questions that have already been answered to change your answers. We are interested in your first answer.

Please answer the questions below and assess how confident you are that your answer is correct. What is the name of an instant camera? a) Canon camera, b) Polaroid camera, c) Minolta camera. [...] How sure are you that your answer is correct? [...] (33% to 100%) [etc., 18 questions, see Michailova and Katter (2014)]

Please estimate: In your opinion, how many questions, out of the total of 18 questions in Part 1a, did you answer correctly? Please enter a number between 0 and 18. [...] In this experiment, [n] participants are taking part. We ask you to estimate now how many questions you have answered correctly compared to the other participants. If you estimate that you have answered more questions correctly than any other participant, enter rank 1. If you estimate that you answered the least questions correctly, enter rank [n]. Please enter a number between 1 and [n]. [...] Thank you, you have completed Part 1a of the experiment. Now Part 1b follows.

Part 1b

Welcome to Part 1b of the experiment. You will see the instructions on the next screen. In this Part (1b), you will make a total of 10 decisions. In every decision, you have the option to choose either Option A or Option B. All 10 decisions are shown in a table below. Each row is one of the ten decisions. At the end of the experiment, we will randomly determine which of your 10 decisions is relevant for the payout.

Here is an example of a decision. If you choose option A in the first line, you can win 20.0 points with a 10% probability. With the remaining probability of 90%, you win 16.0 points. If you choose option B in the first line, you can win 38.5 points with a probability of 10%. With the remaining probability of 90%, you win 1.0 point. Note that you can only make one selection per line. You decide by either checking the box at A or B. [...] [10 lottery choices, see Holt and Laury (2002)]

Thank you, you have completed the entire Part 1 of the experiment. Now Part 2 follows. In Part 2 you will take five individual turns in making investment decisions and group decisions. Please wait until all participants have completed Part 1b. After that, you will receive the detailed instructions for Part 2.

Part 2

Welcome to Part 2 of the experiment.

Individual decisions

In Part 2, the shown probabilities will determine the chances of success of the investment. Your task is to select an investment. Each turn gives you 50 points of capital, and you can decide how much you want to invest.

Group decisions

Immediately after making the first individual decision, you make the same decision in a group that consists of you and one other participant. After that, you make the second individual decision and the same second decision in the same group, then the third, and so on. Before each group decision, you have 45 seconds to communicate with the other participant in your group in a chat. Both you and the other participant remain anonymous.

The result of each group decision will count for both participants in the group. You have to choose the investment in such a way that you and the other participant have selected the same option. You have a maximum of 90 seconds per round for the group decision. If you need more than 60 seconds, you will be penalized with 15 minus points from the total account. After 90 seconds, option A is selected for both participants, but with 15 minus points, and you automatically move to the next round. The timer on the top right will start at 90 seconds.

General notes

You can only select one investment option in each round. In the whole Part 2, you make individual decisions and group decisions five times in an alternating manner.

Part 2: Individual decisions

Option	Investition	Hohes Ergebnis	Hohe Wahrscheinlichkeit	Niedriges Ergebnis	Niedrige Wahrscheinlichkeit
A	0			50	
B	5	55	90%	45	10%
C	10	60	80%	40	20%
D	15	65	70%	35	30%
E	20	70	60%	30	40%
F	25	75	50%	25	50%
G	30	80	40%	20	60%
H	35	85	30%	15	70%
I	40	90	20%	10	80%
J	45	95	10%	5	90%

In this part, your task is to select an investment. Each turn gives you 50 points of capital, and you can decide how much you want to invest. In the situation shown above, you have received 50 points and can invest from 0 to 45 points (options A to J). If you decide not to make an investment (Option A, 0 points), you keep the capital endowment. For example, if you invest 45 points (Option J), you receive 95 points with 10% probability and keep only the remaining not-invested 5 points with 90% probability.

Part 2: Group decisions

Ihre Optionsauswahl	Investition	Vorherige Optionsauswahl des anderen Teilnehmers
A	0	
B	5	
C	10	
D	15	
E	20	E
F	25	
G	30	
H	35	
I	40	
J	45	

In each round, you will again have a capital of 50 points and decide how much you want to invest. You must choose in such a way that you and the other participant have selected the same option. In the situation shown above, the other participant in your group has selected option E. You have a maximum of 90 seconds per round for the group decision. If you need more than 60 seconds, you will be penalized with 15 minus points from the total account. After 90 seconds, option A is selected for both participants, but with 15 minus points, and you automatically move to the next round.

[Quiz]

Option	Investition	Hohes Ergebnis	Hohe Wahrscheinlichkeit	Niedriges Ergebnis	Niedrige Wahrscheinlichkeit
A	0			50	
B	5	55	90%	45	10%
C	10	60	80%	40	20%
D	15	65	70%	35	30%
E	20	70	60%	30	40%
F	25	75	50%	25	50%
G	30	80	40%	20	60%
H	35	85	30%	15	70%
I	40	90	20%	10	80%
J	45	95	10%	5	90%

Please answer the following understanding-questions about the experiment. 1. In the example shown above, what amount of investment (in points) corresponds to option E? [...] 2. If you chose option E, at most how many points could you get paid? [...] 3. How much time (in seconds) do you have in each group decision until you are penalized with 15 minus points from the total account? [...]

You have not answered all the questions correctly.

1. In the example shown above, which investment (in points) corresponds to the option E? Your answer: [...] The correct answer: 20.

2. If you chose option E, at most how many points could you get paid? Your answer: [...] The correct answer: 70.

3. How much time (in seconds) do you have in each group decision until you are penalized with 15 minus points from the total account? Your answer: [...] The correct answer: 60.

If you wish, go back to the instructions or ask an experimenter for explanations. Otherwise, click "Next".

[First individual decision]
Capital for the [first] individual decision: 50 points. Please consider the following investment opportunity. How much do you want to invest? [...]

[First communication stage]
Before each group decision, you have 45 seconds to communicate with the other participant in your group. Now you can discuss the [first] group decision of Part 2. Please press "Enter" to send messages. [...]

[First group decision]
Capital for the [first] group decision: 50 points. Please note the following investment opportunity again. How much do you want to invest? [...]
[either] No agreement was found. Your suggestion: [...] Suggestion of the other participant: [...] You are entering a new round of negotiations.
[or] The proposal was accepted. Your suggestion: [...] Suggestion of the other participant: [...] [etc.]

Thank you, you've completed the entire Part 2 of the experiment. Now Part 3 of the experiment follows. Please note that new groups for the group decisions in Part 3 are created. That is, in Part 3, you make the group decisions in a group with a different second participant than in Part 2.

Part 3

Welcome to Part 3 of the experiment. The process in Part 3 is similar to Part 2. Overall, in Part 3, you make five individual investment decisions and group decisions in an alternating manner. In Part 2, the shown probabilities determined the chances of success of the investment. In the following Part 3, instead, certain tasks with a corresponding degree of ease (easiness of the task) will determine the success of the investment.

The easiness of the task indicates what percentage of a sample of German general population has answered the task correctly. The type of tasks is similar to Part 1: A task consists of a question with four possible answers. You will not receive any feedback as to which answer is correct.

Individual decisions

Your task is to select an investment. Each turn gives you 50 points of capital and you can decide how much you want to invest. After selecting the investment option, you will answer a task. If you answer the corresponding task correctly, you will receive the respective high result. If not, you will receive the respective low result. The tasks never repeat. The task should not take more than 10-15 seconds. After fifteen seconds, you automatically enter the next round.

Group decisions

Immediately after making the first individual decision, you make the same decision in a group that consists of you and one other participant. Then you make the second individual decision and the same second decision in a group, then the third, and so on. Before each group decision, you have 45 seconds to communicate with the other participant in your group in a chat. Both you and the other participant remain anonymous.

The result of each group decision will count for both participants in the group. You have to vote for a decision in such a way that you and the other participant have selected the same option. You have a maximum of 90 seconds per round for the group decision. If you need more than 60 seconds, you will be penalized with 15 minus points from the total account. After 90 seconds, option A (without a task) is selected for both participants, but with the 15 minus points, and you automatically move to the next round. The timer on the top right will start at 90 seconds.

Part 3: Individual decisions

Option	Investition	Hohes Ergebnis	Niedriges Ergebnis	Leichtigkeit der Aufgabe
A	0		50	
B	5	55	45	90%
C	10	60	40	80%
D	15	65	35	70%
E	20	70	30	60%
F	25	75	25	50%
G	30	80	20	40%
H	35	85	15	30%
I	40	90	10	20%
J	45	95	5	10%

In this part, your task is to select an investment. Each turn gives you 50 points of capital and you can decide how much you want to invest. In the situation shown above, you have received 50 points and can invest 0 to 45 (options A to J). For example, if you invest 45 points, you have to answer a task that has been correctly answered with a probability of 10% by a sample of general population. If you answer this task correctly, you receive the respective high result (95 points). If not, you receive the respective low result (5 points, meaning you will only retain the remaining not-invested points).

Part 3: Group decisions

In each round, you again have a capital of 50 points and can decide how much you want to invest. You must vote so that you and the other participant have selected the same option. In the shown situation, the other participant in your group has selected option E. You have a maximum of 90 seconds per round for the group decision. If you need more than 60 seconds, you will be penalized with 15 minus points from the total account. After 90 seconds, option A will be selected for both participants, but with the 15 minus points, and you will automatically move to the next round.

Ihre Optionsauswahl	Investition	Vorherige Optionsauswahl des anderen Teilnehmers
A	0	
B	5	
C	10	
D	15	
E	20	E
F	25	
G	30	
H	35	
I	40	
J	45	

Volunteers in group decisions

For individual decisions, you yourself answer the corresponding tasks. For the group decisions, a participant in the group (called the volunteer) answers the tasks. Before the first group decision, you will decide who of the two participants in your group will be the volunteer. Only the volunteer will answer all five tasks for the group. The result will count for both participants in the group. If the volunteer answers the task correctly, both group members receive the respective high result. If not, both receive the low result. Important note again: For Part 3, the groups for the group decisions are newly created. In Part 3, you make the group decisions in a group with a different second participant than in Part 2.

[Quiz]

Option	Investition	Hohes Ergebnis	Niedriges Ergebnis	Leichtigkeit der Aufgabe
A	0		50	
B	5	55	45	90%
C	10	60	40	80%
D	15	65	35	70%
E	20	70	30	60%
F	25	75	25	50%
G	30	80	20	40%
H	35	85	15	30%
I	40	90	10	20%
J	45	95	5	10%

Please answer the following understanding-questions about the experiment. 1. If you selected option E and answered the question correctly, how many points could you get paid (if this round was randomly drawn at the end?) [...] 2. How many tasks will the volunteer be answering for the group decisions in Part 3? [...]

You have [not] answered all the questions correctly.

1. If you selected option E and answered the question correctly, how many points could you get paid (if this round was randomly drawn at the end)? Your answer: [...] The correct answer: 70.

2. How many tasks will the volunteer be answering for the group decisions in Part 3? Your answer: [...] The correct answer: 5.

If you wish, go back to the instructions or ask an experimenter for explanations. Otherwise, click "Next".

[First individual decision]
Capital for the [first] individual decision: 50 points. Please note the following investment opportunity. How much do you want to invest? [...]

After selecting the investment option, you will answer a task. If you answer the question correctly, you will get the high result. If not, you'll get the low result. "Easiness of task" indicates the probability that the general population could answer this task correctly.

[e.g.]
Easiness of task: 90%
What do you turn on when you press the "A/C" button in your car? a) Air conditioning, b) Fog lights, c) Windscreen wiper, d) Reverse gear. [...]

[Determining the volunteer]
You will now decide which of you in the newly created group will answer all five tasks for the following group decisions in Part 3. Please indicate if you would like to be the one to answer the tasks for your group (volunteer). If you and the other participant choose the same option, the roles (volunteer and not volunteer) will be randomly drawn.

1) Yes, I definitely want to answer the tasks for the group. 2) Yes, I want to answer the tasks for the group, unless the other participant in my group has selected option 1. 3) No, I do not want to answer the tasks for the group unless the other participant in my group has selected option 4. 4) No, I certainly do not want to answer the tasks for the group. [...]

Your role is: [volunteer / not volunteer]. The other participant in your group is: [not volunteer / volunteer].

[First communication stage]
Before each group decision, you have 45 seconds to communicate with the other participant in your group. Now you can discuss the [first] group decision of Part 3. Please press "Enter" to send messages. [...]

[First group decision]
Capital for the [first] group decision: 50 points. Please note the following investment opportunity. How much do you want to invest? [...]

After selecting the investment option, the volunteer will answer a task for the group. If the volunteer answers the question correctly, both participants in the group receive the respective high result. If not, both participants in the group receive the respective low result. “Easiness of task” indicates the probability that the general population could answer this task correctly.

[either] No agreement was found. Your suggestion: [...] Suggestion of the other participant: [...] You are entering a new round of negotiations.
[or] The proposal was accepted. Your suggestion: [...] Suggestion of the other participant: [...] Now the volunteer does the corresponding task. [etc.]

[e.g.]
Easiness of task: 90%
Which catchphrase often comes up when it comes to recharging batteries? a) Monopoly dilemma, b) Mikado effect, c) Memory effect, d) Mau Mau syndrome.

Thank you, you have completed the entire Part 3 of the experiment. Before you learn the payout information, another decision-making situation in Part 4a and four questions in Part 4b follow. How many points you get in Part 4a is randomly drawn at the end of the experiment. The instructions for these parts can be seen after clicking “Next”.

[Part 4a]
Welcome to Part 4a of the experiment. You will receive 10.0 points from us. You can use this money to invest an amount into a lottery. With the lottery you have a profit probability of 50%.

How is the payment made? It is randomly decided by chance whether your investment is successful or not. If your investment is successful, the amount you spend will be multiplied by 2.5 and paid out to you. If your investment is unsuccessful, you lose your stake. You can always keep the amount that you did not use.

Example 1: You invest 0.0 points. Thus, you will surely get 10.0 points. Example 2: You invest 1.25 points and your investment is successful. Thus you win and receive 8.75 + 1.25 * 2.5 = 11.88 points. Example 3: You invest 1.25 points and your investment is unsuccessful. Thus you lose and receive 10.0 − 1.25 = 8.75 points.

Click on the amount you want to place in the lottery (1 click). With your click on the points, you will get on to the next part immediately. [...] Thank you, you have completed Part 4a of the experiment.

[Part 4b]
Welcome to Part 4b of the experiment. The questions below are about decision making. You will face statistical and numerical questions. There are 4 questions with 4 answers each. You must work on the questions in the given order. You have a total of 4 minutes for this part. After 4 minutes you will automatically start the next part. If you need paper and pen for calculations, please approach us. Click “Next” to go to the questions.

Question 1: Imagine we throw a five-sided die 50 times. Of these 50 throws, how often would this five-sided die show an odd number (1, 3 or 5) on average? [...]

Question 2: Of 1000 people in a small town, 500 are members of the singing club. Of these 500 members of the singing club, 100 are men. Of the 500 inhabitants who are not in the singing club, 300 are men. What is the probability that a randomly selected man is a member of the singing club? Please specify the probability in percent. [...]

Question 3: Imagine throwing a marked die (6 sides). The probability that the die shows a 6 is twice the probability of each of the other numbers. Out of 70 throws, in how many of these 70 throws would this die show a 6, in expectation? [...]

Question 4: In a forest, 20% of the mushrooms are red, 50% are brown and 30% are white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is toxic with a 5% chance. What is the probability that a poisonous mushroom in this forest is red? Please specify the probability in percent. [...]

Thank you, you have completed Part 4b of the experiment. Now comes the payout information and finally the questionnaire.

Payout: Parts 1a and 1b

The randomly drawn round in Part 1a: [...] Your payout in Part 1a (points): [...] Your payout in Part 1b can be found on the next page.
The randomly generated number in Part 1b: [...] Your payout in Part 1b (points): [...]

Payout: Parts 2 and 3

The randomly drawn round from Parts 2 and 3: [...] Your payout in this part (points): [...]

Payout: Parts 4a and 4b

Your payout in Part 4a (points): [...] Your payout in Part 4b (points): [...]

Payout: Summary

Part 1a: [...] Part 1b: [...] Parts 2 and 3: [...] Part 4a: [...] The total payout (points): [...] The total payout (euro): [...] The total payout, including questionnaire (euro): [...]

Abgrenzung

Das erste Kapitel "Motive in ökonomischen Interaktionen: Eine (interaktive) Eye-Tracking-Studie" war eine gemeinsame Arbeit mit Urs Fischbacher und Jan Hausfeld. Die erste Idee für die Forschungsfrage kam von Urs, und von da an haben wir alle gemeinsam das Studiendesign entwickelt. Jan übernahm die Programmierung der verschiedenen Experimente, und wir zwei führten die Sitzungen zur Datenerfassung zusammen. Ich war für den Großteil der Datenanalyse und das Schreiben des Forschungspapiers verantwortlich, Jan und Urs betreuten den Prozess. Mehrere Diskussionsrunden und Überarbeitungen des Forschungspapiers haben wir alle gemeinsam durchgeführt.

Das zweite Kapitel "Überzeugungen über Andere: Vernachlässigung von Informationen anstatt von Typprojektion" war eine Zusammenarbeit mit Sebastian Fehrler und Irenaeus Wolff. Die erste Idee stammt von Sebastian und Irenaeus, und ich beteiligte mich an der Programmierung der ersten Intervention. Ich habe die Datenanalyse der ersten Daten durchgeführt, und wir drei haben gemeinsam die Ideen für die folgenden Interventionen entwickelt. Ich übernahm wieder die Programmierung und den größten Teil der Datenanalyse, während Sebastian und Irenaeus den Prozess betreuten. Ich habe den ersten Entwurf des Forschungspapiers geschrieben, und Sebastian und Irenaeus haben wichtige Beiträge hinzugefügt. Schließlich diskutierten und überarbeiteten wir alle drei gemeinsam das Forschungspapier.

Das dritte Kapitel "Überzeugungen von sich selbst: Selbstüberschätzung und Gruppenzusammensetzung in Investitionsentscheidungen" ist meine Erstautorenstudie in Zusammenarbeit mit Jan Hausfeld und Torsten Twardawski. In dieser Studie entwickelten wir die ursprüngliche Idee zusammen, danach übernahm ich die Verantwortung für das Projekt. Ich habe das Experiment programmiert, und Jan schlug Verbesserungen vor. Ich habe die meisten Sitzungen zum Sammeln von Daten durchgeführt, und Jan hat Unterstützung geleistet. Ich war dann für die meiste Datenanalyse und das Schreiben des Forschungspapiers verantwortlich, und Jan lieferte den größten Teil der Aufsicht.

Complete bibliography

Agranov, M., Goeree, J. K., Romero, J., and Yariv, L. (2017). What makes voters turn out: The effects of polls and beliefs. *Journal of the European Economic Association*, 16(3):825–856.

Aksoy, O. and Weesie, J. (2012). Beliefs about the social orientations of others: A parametric test of the triangle, false consensus, and cone hypotheses. *Journal of Experimental Social Psychology*, 48(1):45–54.

Al-Nowaihi, A. and Dhami, S. (2015). Evidential equilibria: Heuristics and biases in static games of complete information. *Games*, 6(4):637–676.

Anderson, C., Brion, S., Moore, D., and Kennedy, J. (2012). A status-enhancement account of overconfidence. *Journal of Personality and Social Psychology*, 103(4):718.

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica*, 44(3):211–233.

Barber, B. and Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, 116(1):261–292.

Bénabou, R. (2012). Groupthink: Collective delusions in organizations and markets. *Review of Economic Studies*, 80(2):429–462.

Bénabou, R. (2015). The economics of motivated beliefs. *Revue d'économie politique*, 125(5):665–685.

Bénabou, R. and Tirole, J. (2016). Mindful economics: The production, consumption, and value of beliefs. *The Journal of Economic Perspectives*, 30(3):141–164.

Biais, B., Hilton, D., Mazurier, K., and Pouget, S. (2005). Judgemental overconfidence, self-monitoring, and trading performance in an experimental financial market. *The Review of Economic Studies*, 72(2):287–312.

Billett, M. and Qian, Y. (2008). Are overconfident CEOs born or made? Evidence of self-attribution bias from frequent acquirers. *Management Science*, 54(6):1037–1051.

Blanco, M., Engelmann, D., Koch, A., and Normann, H. (2014). Preferences and beliefs in a sequential social dilemma: A within-subjects analysis. *Games and Economic Behavior*, 87:122–135.

Bouton, L., Castanheira, M., and Llorente-Saguer, A. (2016). Divided majority and information aggregation: Theory and experiment. *Journal of Public Economics*, 134:114–128.

Breitmoser, Y. (2019). Knowing me, imagining you: Projection and overbidding in auctions. *Games and Economic Behavior*, 113:423–447.

Brennan, S., Chen, X., Dickinson, C., Neider, M., and Zelinsky, G. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, 106(3):1465–1477.

Brusco, S. (1998). Unique implementation of the full surplus extraction outcome in auctions with correlated types. *Journal of Economic Theory*, 80(2):185–200.

Burns, B. D. (2017). Equiprobability principle principle or "no change" principle ? Examining reasoning in the Monty Hall Dilemma using unequal probabilities. *39th Annual Conference of the Cognitive Science Society*, pages 1697–1702.

Camerer, C. and Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89(1):306–318.

Charness, G. and Gneezy, U. (2010). Portfolio choice and risk attitudes: An experiment. *Economic Inquiry*, 48(1):133–146.

Cokely, E., Galesic, M., Schulz, E.and Ghazal, S., and Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin numeracy test. *Judgment and Decision Making*, 7(1):25.

Collins, B. and Guetzkow, H. (1964). *A social psychology of group processes for decision-making*. New York: John Wiley & Sons.

Conlin, M., O'Donoghue, T., and Vogelsang, T. (2007). Projection bias in catalog orders. *American Economic Review*, 97(4):1217–1249.

Cremer, J. and McLean, R. P. (1985). Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53(2):345–361.

Cremer, J. and McLean, R. P. (1988). Full extraction of the surplus in Bayesian and dominant strategy auctions. *Econometrica*, 56(6):1247.

Danz, D., Kübler, D., Mechtenberg, L., and Schmid, J. (2015). On the failure of hindsight-biased principals to delegate optimally. *Management Science*, 61(8):1938–1958.

Danz, D., Madarasz, K., and Wang, S. (2018). The biases of others: Anticipating information projection in an agency setting.

Dawes, R. (1989). Statistical criteria for establishing a truly false consensus effect. *Journal of Experimental Social Psychology*, 25(1):1–17.

Deaves, R., Lueders, E., and Luo, G. (2009). An experimental test of the impact of overconfidence and gender on trading activity. *Review of Finance*, 13(3):555–575.

Dittrich, D., Güth, W., and Maciejovsky, B. (2005). Overconfidence in investment decisions: An experimental approach. *The European Journal of Finance*, 11(6):471–491.

Duchowski, A. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers*, 34(4):455–470.

Egan, M. (2019). Boeing has a 'crisis of confidence.' It's time for the board to step up. https://edition.cnn.com/2019/05/16/business/boeing-737-max-crisis-faa/index.html. Accessed: 2019-05-22.

Engelmann, D. and Strobel, M. (2012). Deconstruction and reconstruction of an anomaly. *Games and Economic Behavior*, 76(2):678–689.

Enke, B. and Zimmermann, F. (2017). Correlation neglect in belief formation. *The Review of Economic Studies*, 86(1):313–332.

Epley, N. and Gilovich, T. (2016). The mechanics of motivated reasoning. *The Journal of Economic Perspectives*, 30(3):133–140.

Esponda, I. and Vespa, E. (2014). Hypothetical thinking and information extraction in the laboratory. *American Economic Journal: Microeconomics*, 6(4):180–202.

Eyster, E. and Rabin, M. (2005). Cursed equilibrium. *Econometrica*, 73(5):1623–1672.

Eyster, E. and Rabin, M. (2010). Naïve herding in rich-information settings. *American Economic Journal: Microeconomics*, 2(4):221–243.

Fama, E. and Jensen, M. (1983). Separation of ownership and control. *Journal of Law and Economics*, 26(2):301–325.

Fehrler, S. and Hughes, N. (2018). How transparency kills information aggregation: Theory and experiment. *American Economic Journal: Microeconomics*, 10(1):181–209.

Fiedler, S., Glöckner, A., Nicklisch, A., and Dickert, S. (2013). Social value orientation and information search in social dilemmas: An eye-tracking analysis. *Organizational Behavior and Human Decision Processes*, 120(2):272–284.

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

Foulsham, T. and Lock, M. (2015). How the eyes tell lies: Social gaze during a preference task. *Cognitive Science*, 39(7):1704–1726.

Friedman, D. (1998). Monty Hall's three doors: Construction and deconstruction of a choice anomaly. *American Economic Review*, 88(4):933–946.

Gibaldi, A., Vanegas, M., Bex, P. J., and Maiello, G. (2017). Evaluation of the Tobii EyeX Eye tracking controller and Matlab toolkit for research. *Behavior Research Methods*, 49(3):923–946.

Gilbert, D., Pinel, E., Wilson, T., Blumberg, S., and Wheatley, T. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, 75(3):617.

Gillman, L. (1992). The car and the goats. *American Mathematical Monthly*, 99(1):3–7.

Gneezy, U. and Potters, J. (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics*, 112(2):631–645.

Goeree, J. K. and Großer, J. (2007). Welfare reducing polls. *Economic Theory*, 31(1):51–68.

Goldberg, L. (1992). The development of markers for the big-five factor structure. *Psychological Assessment*, 4(1):26.

Gosling, S., Rentfrow, P., and Swann, W. (2003). A very brief measure of the big-five personality domains. *Journal of Research in Personality*, 37(6):504–528.

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

Guarnaschelli, S., McKelvey, R. D., and Palfrey, T. R. (2000). An experimental study of jury decision rules. *American Political Science Review*, 94(2):407.

Hanna, R., Mullainathan, S., and Schwartzstein, J. (2014). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics*, 129(3):1311–1353.

Hausfeld, J., Hesler, K., and Goldlücke, S. (2018). Strategic gaze: An interactive eye-tracking study.

Hilary, G. and Menzly, L. (2006). Does past success lead analysts to become overconfident? *Management Science*, 52(4):489–500.

Hillman, A. and Dalziel, T. (2003). Boards of directors and firm performance: Integrating agency and resource dependence perspectives. *Academy of Management Review*, 28(3):383–396.

Holmes, K., Ayzenberg, V., and Lourenco, S. (2016). Gamble on gaze: Eye movements reflect the numerical value of blackjack hands. *Psychonomic Bulletin & Review*, 23(6):1974–1981.

Holt, C. and Laury, S. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5):1644–1655.

Holt, C. A. and Smith, A. M. (2009). An update on Bayesian updating. *Journal of Economic Behavior and Organization*, 69(2):125–134.

Holt, C. A. and Smith, A. M. (2016). Belief elicitation with a synchronized lottery choice menu that is invariant to risk attitudes. *American Economic Journal: Microeconomics*, 8(1):110–139.

Johnson, D. and Fowler, J. (2011). The evolution of overconfidence. *Nature*, 477(7364):317–320.

Kerr, N. and Tindale, R. (2004). Group performance and decision making. *Annual Review of Psychology*, 55:623–655.

Khorana, A., Tufano, P., and Wedge, L. (2007). Board structure, mergers, and shareholder wealth: A study of the mutual fund industry. *Journal of Financial Economics*, 85(2):571–598.

Kind, A. and Twardawski, T. (2016). Board overconfidence in mergers & acquisitions: A self-attribution bias. *Academy of Management Journal*.

Klayman, J., Soll, J., González-Vallejo, C., and Barlas, S. (1999). Overconfidence: It depends on how, what, and whom you ask. *Organizational Behavior and Human Decision Processes*, 79(3):216–247.

Kluger, B. D. and Wyatt, S. B. (2004). Are judgment errors reflected in market prices and allocations? Experimental evidence based on the Monty Hall problem. *Journal of Finance*, 59(3):969–997.

Knoepfle, D., Wang, J., and Camerer, C. (2009). Studying learning in games using eye-tracking. *Journal of the European Economic Association*, 7(2-3):388–398.

Kocher, M. and Sutter, M. (2005). The decision maker matters: Individual versus group behaviour in experimental beauty-contest games. *The Economic Journal*, 115(500):200–223.

Kolasinski, A. and Li, X. (2013). Can strong boards and trading their own firm's stock help CEOs make better decisions? Evidence from acquisitions by overconfident CEOs. *Journal of Financial and Quantitative Analysis*, 48(04):1173–1206.

Lahey, J. and Oxley, D. (2016). The power of eye tracking in economics experiments. *American Economic Review*, 106(5):309–13.

Langer, E. (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32(2):311.

Levy, G. and Razin, R. (2015). Correlation neglect, voting behavior, and information aggregation. *American Economic Review*, 105(4):1634–1645.

Lichtenstein, S. and Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance*, 20(2):159–183.

Litchfield, D. and Ball, L. (2011). Rapid communication: Using another's gaze as an explicit aid to insight problem solving. *Quarterly Journal of Experimental Psychology*, 64(4):649–656.

Litchfield, D., Ball, L., Donovan, T., Manning, D., and Crawford, T. (2010). Viewing another person's eye movements improves identification of pulmonary nodules in chest x-ray inspection. *Journal of Experimental Psychology: Applied*, 16(3):251.

Madarász, K. (2012). Information projection: Model and applications. *The Review of Economic Studies*, 79(3):961–985.

Madarász, K. (2015). Projection equilibrium: Definition and applications to social investment and persuasion. *Mimeo*, (January):1–48.

Malmendier, U. and Tate, G. (2005). CEO overconfidence and corporate investment. *Journal of Finance*, 60(6):2661–2700.

Malmendier, U. and Tate, G. (2008). Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics*, 89(1):20–43.

Messé, L. and Sivacek, J. (1979). Predictions of others' responses in a mixed-motive game: Self-justification or false consensus? *Journal of Personality and Social Psychology*, 37(4):602.

Michailova, J. and Katter, J. (2014). Quantifying overconfidence in experimental finance. *International Journal of Behavioural Accounting and Finance*, 4(4):351–367.

Moore, D. and Healy, P. (2008). The trouble with overconfidence. *Psychological Review*, 115(2):502.

Müller, R., Helmert, J., and Pannasch, S. (2014). Limitations of gaze transfer: Without visual context, eye movements do not to help to coordinate joint action, whereas mouse movements do. *Acta Psychologica*, 152:19–28.

Müller, R., Helmert, J., Pannasch, S., and Velichkovsky, B. (2013). Gaze transfer in remote cooperation: Is it always helpful to see what your partner is attending to? *Quarterly Journal of Experimental Psychology*, 66(7):1302–1316.

Murphy, R., Ackermann, K., and Handgraaf, M. (2011). Measuring social value orientation. *Judgment and Decision Making*, 6(8):771–781.

Nalebuff, B. (1987). Puzzles. *Journal of Economic Perspectives*, 1(1):157–63.

Neider, M., Chen, X., Dickinson, C., Brennan, S., and Zelinsky, G. (2010). Coordinating spatial referencing using shared gaze. *Psychonomic Bulletin & Review*, 17(5):718–724.

Offerman, T. and Palley, A. (2016). Lossed in translation: An off-the-shelf method to recover probabilistic beliefs from loss-averse agents. *Experimental Economics*, 19(1):1–30.

Offerman, T., Sonnemans, J., and Schram, A. (1996). Value orientations, expectations and voluntary contributions in public goods. *Economic Journal*, pages 817–845.

Page, S. E. (1998). Let's make a deal. *Economics Letters*, 61:175–180.

Pajares, F. (2008). Motivational role of self-efficacy beliefs in self-regulated learning. *Motivation and self-regulated learning: Theory, research, and applications*, pages 111–139.

Palacios-Huerta, I. (2003). Learning to open Monty Hall's doors. *Experimental Economics*, 6(3):235–251.

Polonio, L., Di Guida, S., and Coricelli, G. (2015). Strategic sophistication and attention in games: An eye-tracking study. *Games and Economic Behavior*, 94:80–96.

Read, D. and Van Leeuwen, B. (1998). Predicting hunger: The effects of appetite and delay on choice. *Organizational Behavior and Human Decision Processes*, 76(2):189–205.

Roll, R. (1986). The hubris hypothesis of corporate takeovers. *Journal of Business*, pages 197–216.

Ross, L., Greene, D., and House, P. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13(3):279–301.

Schaefer, P., Williams, C., Goodie, A., and Campbell, W. (2004). Overconfidence and the Big Five. *Journal of Research in Personality*, 38(5):473–480.

Scheier, M., Carver, C., and Bridges, M. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the life orientation test. *Journal of Personality and Social Psychology*, 67(6):1063.

Schwartz-Ziv, M. and Weisbach, M. (2013). What do boards really do? Evidence from minutes of board meetings. *Journal of Financial Economics*, 108(2):349–366.

Selvin, S. (1975a). A problem in probability [Letter to the editor]. *The American Statistician*, 29(1):67.

Selvin, S. (1975b). On the Monty Hall problem [Letter to the editor]. *The American Statistician*, 29(3):134.

Shefrin, H. (2019). Psychological Issues Loom Large in Wake of Boeing 737 Max Crashes. https://www.forbes.com/sites/hershshefrin/2019/03/17/psychological-issues-loom-large-in-wake-of-boeing-737-max-crashes/. Accessed: 2019-05-20.

Siddiqi, H. (2009). Is the lure of choice reflected in market prices? Experimental evidence based on the 4-door Monty Hall problem. *Journal of Economic Psychology*, 30(2):203–215.

Slembeck, T. and Tyran, J.-R. (2004). Do institutions promote rationality? An experimental study of the three-door anomaly. *Journal of Economic Behavior & Organization*, 54(3):337–350.

Sniezek, J. (1992). Groups under uncertainty: An examination of confidence in group decision making. *Organizational Behavior and Human Decision Processes*, 52(1):124–155.

Stein, R. and Brennan, S. (2004). Another person's eye gaze as a cue in solving programming problems. In *Proceedings of the 6th international conference on Multimodal interfaces*, pages 9–15. ACM.

Stewart, N., Gächter, S., Noguchi, T., and Mullett, T. (2016). Eye movements in strategic choice. *Journal of Behavioral Decision Making*, 29(2-3):137–156.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, 47(2):143–148.

Taylor, C. R. and Yildirim, H. (2010). Public information and electoral bias. *Games and Economic Behavior*, 68(1):353–375.

Taylor, S. E. and Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103(2):193.

Tolvanen, J. (2017). Political ambiguity in a world of heterogeneous beliefs. *Mimeo*.

Tor, A. and Bazerman, M. H. (2003). Focusing failures in competitive environments: Explaining decision errors in the Monty Hall game, the acquiring a company problem, and multiparty ultimatums. *Journal of Behavioral Decision Making*, 16(5):353–374.

Trevelyan, R. (2008). Optimism, overconfidence and entrepreneurial activity. *Management Decision*, 46(7):986–1001.

Velichkovsky, B. (1995). Communicating attention: Gaze position transfer in cooperative problem solving. *Pragmatics & Cognition*, 3(2):199–223.

Viscusi, W., Phillips, O., and Kroll, S. (2011). Risky investment decisions: How are individuals influenced by their groups? *Journal of Risk and Uncertainty*, 43(2):81.

Vrij, A. (2004). Why professionals fail to catch liars and how they can improve. *Legal and Criminological Psychology*, 9(2):159–181.

Wang, J., Spezio, M., and Camerer, C. (2010). Pinocchio's pupil: Using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. *American Economic Review*, 100(3):984–1007.

Weinstein, N. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, 39(5):806.

Zarnoth, P. and Sniezek, J. (1997). The social influence of confidence in group decision making. *Journal of Experimental Social Psychology*, 33(4):345–366.