Essays in Behavioral Economics

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von
Florian Zimmermann
aus Leonberg

Bonn 2013

Dekan:	Prof. Dr. Klaus Sandmann
Erstreferent:	Prof. Dr. Armin Falk
Zweitreferent:	Prof. Dr. Sebastian Kube

Tag der mündlichen Prüfung: 5. Juli 2013

Acknowledgments

I thank Armin Falk for great help, advice and encouragement through all stages of my dissertation. I am still amazed by, and very grateful for the amount of time and energy he was willing to invest in my own and our joint research. His way of doing research has been very inspiring.

I also thank Paul Heidhues, Sebastian Kube and Daniel Krähmer for the help and advice. Uri Gneezy has been a wonderful host during my stays at UC San Diego. I enjoyed a lot and benefited from many discussions with him. I also thank my coauthor Mara Ewers for a great collaboration.

Bonn Graduate School of Economics has been a wonderful place to study and do research, and I thank all people who help running and improving BGSE, especially Urs Schweizer, Silke Kinzig and Pamela Mertens.

The institute in empirical economics has always been a great place to hang out and I am thankful for all the great people I was allowed to meet there over the years. Special thanks are owed to Konstanze Albrecht, Steffen Altmann, Mirko Seithe, Hannah Schildberg-Hoerisch and Matthias Wibral.

Finally and most importantly, I thank my family! My parents and my sister for their love, encouragement and support.

to my parents

Contents

List of Figures

List of Tables

Introduction

Answers to virtually all questions of economic relevance require an understanding of how economic agents behave. The economic consequences of tax cuts can only be studied with a theory of how individuals, e.g., consumers, employees, employers, respond to the tax cuts. Likewise, effects of changing unemployment benefits cannot be understood without having an idea of how the unemployed, but also the employed and firms will respond to the changes. The traditional economic approach to decision-making has been to assume rational agents that possess well-defined preferences and, given beliefs that are formed through Bayesian updating and include all available information, select their preferred alternative. This approach has been very successful and provided a tractable and parsimonious workhorse to study economic behavior.

In the past decades, however, economists have started to incorporate insights from related disciplines, e.g., psychology, sociology, to develop a more precise and realistic model of economic behavior. Pathbreaking studies have been (just to name a few) Kahneman and Tversky (1979), who provide a model of expected utility that incorporates loss aversion as well as non-bayesian belief formation, Laibson (1997), formalizing the notion of hyperbolic discounting and Rabin (1993), proposing a way to include social preferences into game-theoretic analysis. This approach has helped to align empirical phenomena that are hard to reconcile with "standard" economic assumptions. It has also lead to the development of novel policy instruments that (for example) take into account cognitive limitations and misperceptions of agents. Examples are Thaler and Benartzi (2004) for savings behavior or Hastings and Weinstein (2008) on parents' school choice.

In the following chapters, both theory and controlled experiments are used to better understand the foundations of economic decision-making, and to derive novel economic implications. While the topics of the four chapters are rather diverse, the common theme is the attempt to contribute to a more realistic model of economic behavior. In chapter

1, we conduct a controlled lab experiment to test a key implication of a recent model developed in Kőszegi and Rabin (2009). Chapters 2 and 3 are similar in structure. Both propose simple behavioral models whose central implications are then tested experimentally. Chapter 4 uses insights from chapter 3 as well as from a related literature in psychology to provide an explanation for so-called anchoring effects, a phenomenon that is at odds with traditional models of economic decision-making. In the following, the four chapters of this dissertation are briefly summarized.

Chapter 1 focuses on individuals' attitudes towards the timing of information. We test a theoretical prediction by Kőszegi and Rabin (2009), that people prefer to get information "clumped together" rather than piecewise. We conduct a controlled lab experiment where subjects participate in a lottery and can choose between different resolutions of uncertainty (clumped or piecewise). In two treatments we analyze which kind of resolution is preferred. Two additional treatments allow us to get a quantitative measure of subjects' preferences over different information structures. Our data does not support the prediction that piecewise information is utility-decreasing.

In chapter 2, we ask if reports of private information about skills, abilities or achievements are affected by image concerns. We develop a simple model that illustrates how image utility can lead to misreporting of private information in contexts where truthful reports maximize monetary outcomes. In addition, we test the model's predictions in a controlled lab experiment. In the experiment, all subjects go through a series of quiz questions and subsequently report a performance measure. We vary if reports are made to an audience or not and find evidence for image effects. In the audience treatment, stated reports are significantly higher than in the private treatment. This suggests that overconfident appearance might be a consequence of social approval seeking. We also find that men state higher self-assessments than women. This gender difference seems to be driven by men responding more strongly to the presence of an audience.

Chapter 3 studies the role of consistency as a signaling device. We propose a two-period model that highlights the informativeness of consistency as a signal of skills and allows to analyze consequences for behavior. In a simple principal-agent experiment we test the basic intuition of the model, that consistency is valued by others, inducing people to act consistently. In the second part of the chapter we study the consequence of early commitment for behavior. In the context of an estimation task we demonstrate that commitment leads to a neglect of valuable information. Furthermore, the potential of consistency as a device of social influence is studied in the context of surveys.

In Chapter 4, we provide an explanation for so-called anchoring effects. Random anchors have been shown to systematically affect judgments and valuations. This has called into question the rationality of judgments as well as the existence of stable preference relations. Instead this evidence suggests that both judgments and valuations are to a large degree arbitrary. This chapter is an attempt to reconcile evidence from anchoring manipulations with a model where decision-makers are rational and have stable preferences or judgments.

A final remark concerning the use of the first person plural throughout this dissertation: it is owed to the fact that chapter 2 was developed in a collaboration with Mara Ewers, and chapters 3 and 4 are the product of joint work with Armin Falk. For reasons of consistency, the plural is also used in this introduction and chapter 1.[1] The next four chapters are each presented as self-contained units.

[1]See chapter 3 for both theoretical and empirical evidence for the informativeness of consistency as a signal of skills.

Chapter 1

Clumped or Piecewise? - Evidence on Preferences for Information

1.1 Introduction

The selection and processing of information is a key element in virtually all areas of economic decision-making. Individuals facing economic choices, e.g., investing in education, choosing an optimal health insurance plan, buying a house or deciding how much to save for the future, need to choose sources of potentially helpful information and process this information to be able to make an informed decision. Likewise, economic choices affect the kind, structure and timing of information decision-makers will receive. A decision to participate in a risky enterprise implies, that the decision-maker will receive news about the success or failure of the enterprise in the future. Therefore, attitudes or preferences towards information structures can be an important factor influencing choices and behavior.

Furthermore, the structuring of information can serve as a policy or managerial instrument. Policy-makers, when providing information on, e.g, the current state of political reform or consequences from a natural disaster, need to take the impact of the timing of information provision into account. Likewise, employers providing feedback to their employees can structure the feedback to their own advantage. The traditional economic approach to decision-making, however, neglects that the information an individual receives might have direct utility consequences.

Recent theories, e.g., Kőszegi and Rabin (2009), have tried to incorporate attitudes towards information into models of decision-making.[1] A key prediction of Kőszegi and Rabin (2009) is that individuals are averse to piecewise information. Thus, they should prefer to receive information in one piece rather than piece by piece.[2] Their model provides explanations for various phenomena such as loss aversion over wealth, overconsumption or precautionary savings. Empirically, however, little is known about preferences for clumped or piecewise information. In this chapter, we use a controlled lab experiment to test the implication that people have a preference for information in one piece. As a whole, we find no support for this prediction.

Kőszegi and Rabin (2009) develop a dynamic model of reference-dependent preferences. A central assumption of the model is that utility depends on anticipated changes in beliefs about current and future consumption. Beliefs are rational and people are loss averse with regard to changes in their beliefs.[3] Thus bad news decrease utility more than good news increase it. Furthermore it is assumed that people care less about changes in beliefs, the further away the time of belief change lies from the actual point of consumption. In other words, a person is assumed to be less sensitive to changes in beliefs, the more time lies in between news and the time of consumption. The model gives rise to informational preferences, i.e., preferences towards the timing of non-instrumental information. Loss aversion in belief changes leads to a preference for clumped information. Since bad news decrease utility more than good news increase it, decision-makers are averse to fluctuations in their beliefs. Consequently piecewise information is utility-decreasing.

In this chapter we test the prediction that piecewise information is utility-decreasing. In the experiment, subjects can choose how they want to be informed about the outcome of a lottery. They have two options: Either they learn the outcome of the lottery in

[1]Caplin and Leahy (2001) incorporate anticipatory emotions towards uncertainty resolution into an expected utility framework and analyze consequences, for example on portfolio choice. In another paper, Caplin and Leahy (2004), use an expected utility framework with anticipatory emotions to analyze how much information an expert should transmit to a poorly informed person.

[2]Similar implications are derived in theoretical work by Palacios-Huerta (1999) and Dillenberger (2010). Palacios-Huerta (1999) develops an argument why people might prefer clumped information based on an example of the model of disappointment aversion by Gul (1991). Dillenberger (2010) considers a general class of recursive, non-expected preferences over compound lotteries. He shows equivalence between a preference for information in one piece and the so-called "certainty effect" by Kahneman and Tversky (1979). A related intuition can also be found in Kőszegi and Rabin (2009).

[3]The idea that reference points are determined by rational expectations has been developed in Kőszegi and Rabin (2006, 2007). Similar approaches can be found in the disappointment aversion models of Bell (1985), Loomes and Sugden (1986), and Gul (1991). Several recent empirical studies provide support for expectation-based reference points. See for example Abeler et al. (2011), Crawford and Meng (2011), Gill and Prowse (forthcoming) and Ericson and Fuster (forthcoming).

one piece, or they are sequentially informed about it. Information in this setting is non-instrumental since the lottery is an exogenous event which cannot be influenced by the subjects. Subjects' choices allow us to analyze which information structure is preferred. Two additional treatments allow us to specify a willingness to pay, i.e., a quantitative measure. In these treatments, subjects cannot choose between clumped or piecewise information but are exposed to either one of the two. A subject's choice in these treatments is to state a willingness to pay for participating in the lottery. Comparison of the average willingness to pay between the two treatments provides a quantitative measure for preferences over different information structures.

Summarizing our results, we find no evidence that subjects are averse to piecewise information. When subjects can directly choose between the two information conditions, only slightly more than 50 percent prefer to receive information in one piece. This is only compatible with a preference for clumped information if one is willing to allow for very high error rates. The average willingness to pay for the lottery is more than 2 Euro higher when subjects are sequentially informed about the outcome of the lottery. We can reject the null hypothesis that subjects' willingness to pay for the lottery is higher in the clumped information condition.

Our study is the first to provide a direct experimental test of whether individuals are averse to piecewise information. Moreover, our findings are important, as the assumptions that lead to the prediction we test have several implications for behavior. Kőszegi and Rabin (2009) show that loss aversion in belief changes provides a foundation for loss aversion over total wealth as is assumed for example in prospect theory (Kahneman and Tversky (1979)). The intuition is simple. Wealth gains and wealth losses are news about current and future consumption. Consequently, loss aversion over consumption news induces gain-loss utility over wealth. In a two-period application, Kőszegi and Rabin (2009) show how their model can generate a novel type of overconsumption. For example, in contexts where wealth is deterministic, people might deviate from ex-ante optimal consumption plans and overconsume relative to the plan, because good news about increased consumption now might outweigh bad news about future consumption due to decreasing sensitivity towards belief changes. Consequently the ex-ante optimal plan is not credible. Actual consumption in period 1 will be above the ex-ante optimal level to account for the lack of credibility of the ex-ante optimal plan. Loss aversion in belief changes also generates a new type of precautionary savings motive. In their two-period application, Kőszegi and Rabin (2009) analyze how decision-makers respond to future wealth uncertainty (resolved in period 2). They show that decision makers respond to higher uncertainty by reducing consumption in period 1. Intuitively, future wealth uncertainty exposes decision makers to (potentially) negative belief shocks which are felt heavily due to loss aversion, but can be dampened by higher savings in period 1.

In addition, our results contribute to the experimental literature on myopic loss aversion (see Benartzi and Thaler (1995) and Gneezy and Potters (1997)). Gneezy and Potters (1997) let subjects repeatedly go through risky investment choices and vary the frequency with which they received feedback regarding the outcome and with which they could make their choices. They find that investments in the risky asset are higher when the frequency of feedback and choices is low. Haigh and List (2005) replicate this result with professional traders. One question that arises is whether these results are due to the frequency of choices or the frequency of feedback. Our results suggest that myopic loss aversion is most likely not driven by a direct preference for a clumped timing structure in the resolution of risk. Note that Bellemare et al. (2005) provide evidence in the opposite direction. They conduct an experiment similar to Gneezy and Potters (1997), with the additional twist that it allows to disentangle effects of frequency of feedback from frequency of choices. They find that manipulating feedback is sufficient to generate myopic loss aversion. This finding is compatible with a preference for clumped information. Langer and Weber (2008), however, document the opposite. They identify frequency of choices as the relevant factor that drives myopic loss aversion.[4]

There exists a small empirical literature on informational preferences, but no incentivized study addresses the question if subjects prefer clumped information over piecewise information. Chew and Ho (1994) and Ahlbrecht and Weber (1996) are early examples. Both use questionnaire formats to examine preferences for different resolutions of uncertainty. More recently several incentivized experiments were conducted. Eliaz and Schotter (2007) find that subjects are willing to pay for earlier reception of non-instrumental information. Eliaz and Schotter (2010) show evidence for a demand for non-instrumental information about the likelihood that a risky choice was optimal. Van Winden et al. (2011) examine how investment decisions are affected by a delay in the resolution of risk. They find a significant impact of the delay of non-instrumental information and show that

[4]Fellner and Sutter (2009) find that both factors (frequency of feedback and frequency of choices) are important for myopic loss aversion.

emotions play a central role in explaining their results. Kocher et al. (2009) find that subjects holding a lottery ticket have a preference for delayed resolution of risk and that this preference is driven by positive anticipatory emotions.

The remainder of the chapter is organized as follows. The next section describes the experimental design and states hypotheses. Section 1.3 shows results and section 1.4 concludes.

1.2 Experimental Design and Hypotheses

An environment where preferences towards the timing of information can be studied needs the following features:

1. Non-instrumentality of information: information needs to be on a predetermined event that can not be affected by subjects. For this kind of information, “standard” expected utility theory predicts indifference towards the timing of information.

2. Meaningful time delays: Kőszegi and Rabin (2009) characterize differences in the timing of information by signals arriving in different time periods, leaving open the length of a time period. In principle time periods could be seconds, minutes, days or months. When testing their predictions we need to create an environment where the variation in the timing structure involves different time periods in the sense of Kőszegi and Rabin (2009). In particular very small variations might be problematic. Say for example that we would vary the timing structure by having signals arrive every 10 seconds. Then it could well be that subjects integrate signals that follow each other that closely into one signal, thereby perceiving piecewise as clumped information. Note however that while leaving the length of a time period open, Kőszegi and Rabin (2009) also do not exclude any specifications.

3. Absorption of information: to implement different timing structures, we need to make sure that subjects absorb information at the moment they receive it. If subjects have the possibility to delay absorption, for example by not reading information provided on a computer screen or a sheet of paper, we loose control over the timing structure.

1.2.1 Experimental Design

We designed an experiment that captures all features discussed above. We studied four treatments in total. In **treatments 1** and **2**, subjects were endowed with a lottery ticket. A central characteristic of the lottery was that it contained a natural sequence of three signals about the outcome of the lottery. Each of the three signals served as a piece of information. Since the lottery outcome could not be affected by subjects, information was non-instrumental. Subjects' choices were about how they wanted to be informed about the outcome of the lottery. We offered two possibilities: information in one piece or sequential information. Given our goal to make variations in the resolution of uncertainty meaningful we decided to run the experiment over days. The information conditions and the different steps of the experiment are illustrated in *Figure 1.1.* If subjects preferred to receive information clumped, the three signals were collapsed into one. Subjects were informed in one piece about the final outcome of the lottery on day 2 of the experiment. If subjects chose to receive information piecewise, they were sequentially provided with the three pieces of information. They learned the first piece on the second day of the experiment. One day after they received the second signal. On day 4 they learned the third and final piece of information regarding the lottery outcome.[5] In order to make sure that subjects absorbed information by the time we revealed it, we informed them via phone calls. Via telephone we achieved full control on the timing of resolution of uncertainty about the lottery outcome.[6]

The only difference between treatments 1 and 2 was the lottery. In **treatment 1**, part of the lottery was a starting endowment of 30 Euro (one Euro was worth 1.45 US-Dollar at the time). A fair dice was thrown three times and the numbers thrown were added up. If the total sum after three throws was larger than or equal to 13, subjects won 50 Euro which were added to their starting endowment of 30 Euro. In case the total sum was smaller than 13, subjects lost 15 Euro which were deducted from their starting capital. The lottery has an expected value of about 32 Euro and a standard deviation of 28.5. Each of the three dice throws represented a piece of information, allowing subjects to

[5] Note that in the clumped condition signals are collapsed into 1 signal that is received at day 2. Thus when comparing the clumped and the piecewise condition, no signals were delayed through clumping. This is important, because in Kőszegi and Rabin (2009) people only strictly prefer clumped to piecewise information if the clumped condition does not involve any delay of signals, see section 1.2.3 and Appendix A.

[6] In section 1.2.2 we will present the exact procedures of the experiment in more detail.

update their beliefs regarding the outcome of the lottery.

In **treatment 2**, we changed the payoff structure of the lottery. In Kőszegi and Rabin (2009), people are loss averse with respect to anticipated belief changes. We suspected that anticipation effects might be more pronounced the more meaningful the outcome is to subjects. While in treatment 1 stakes and the payoff difference between winning and losing were already high, we decided to use a lottery in treatment 2 which has an almost 10-times higher payoff difference. Subjects could either gain 500 Euro or zero.[7] The lottery worked as follows. In three rounds three dice were thrown simultaneously. Subjects won if in at least one round, all three dice showed a six. The lottery has an expected value of about 7 Euro, and a standard deviation of roughly 58.7. As in treatment 1 each of the three rounds of dice rolls represented a piece of information.

Subjects' choices between clumped or piecewise information in treatments 1 and 2 allow us to qualitatively examine on an individual level which information structure is preferred. **Treatments 3** and **4** allow us to specify a willingness to pay, i.e., a quantitative measure.[8] In these treatments, subjects could not choose between the two information

Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
(BonnEconLab)	(Phonecall)	(Phonecall)	(Phonecall)	(Experimenter's Office)
- **Main Decision** *Treatment 1:* clumped vs. piecewise *Treatment 2:* clumped vs. piecewise *Treatment 3:* willingness to pay for lottery *Treatment 4:* willingness to pay for lottery **- Measure for Loss Aversion** **- Measure for Risk Aversion** (only treatments 3 & 4)	**Clumped Condition:** - Information about outcome of lottery **Piecewise Condition:** - First piece of information about outcome of lottery (result of first dice roll)	**Clumped Condition:** - No further information about lottery **Piecewise Condition:** - Second piece of information about outcome of lottery (result of second dice roll)	**Clumped Condition:** - No further information about lottery **Piecewise Condition:** - Third and final piece of information about outcome of lottery (result of third dice roll)	Payment

Figure 1.1: Illustration of experimental design.

conditions. Instead they found themselves in one of the two conditions and were asked to

[7]In addition, in treatment 2 subjects received a show-up fee of 15 Euro.

[8]Note that treatments 3 and 4 were conducted before treatments 1 and 2. While we do not think that this changes the interpretation or validity of our results in any way, we report this here for the sake of completeness and to avoid any misunderstandings.

state their willingness to pay to participate in the lottery. The information conditions were identical to treatments 1 and 2. Subjects in treatment 3 received information clumped, subjects in treatment 4 received information piecewise. The lottery was the same as in treatment 1.

The only decision subjects had to make was to choose their willingness to pay for the lottery. We used a multiple price list format to elicit certainty equivalents.[9] In particular, subjects had to make 25 choices between the lottery and a certain amount which was increased from 13 Euro to 37 Euro in increments of 1 Euro. One of the 25 choices was afterwards randomly selected and implemented. If subjects behaved consistently, they (at maximum) switched once between the lottery and the fixed payment. This switching point was used as subjects' willingness to pay for the lottery. Comparison of the average willingness to pay for the lottery between the treatments 3 and 4 allows us to analyze if and to what degree subjects preferred clumped over piecewise information.

1.2.2 Procedural Details

In all four treatments the experiment went over 5 days, starting on a Monday and ending on Friday of the same week. On Monday subjects met in the experimental lab. They were welcomed and assigned into cabins. Then instructions were passed and read aloud.[10] Subjects were instructed in detail about the lottery and the information conditions. In treatments 1 and 2, subjects were informed about both information conditions, in treatments 3 and 4 they were only informed about the information condition of the respective treatment. Then subjects had to make their choice. In treatments 1 and 2 they decided which information condition they preferred. In treatments 3 and 4 they stated their willingness to pay for the lottery.

In all treatments we also elicited a measure of loss aversion, following the procedure of Fehr and Goette (2007). Subjects faced two lottery choices. In choice 1 they had to decide whether they want to participate in a lottery where they could win 3 Euro with probability $\frac{1}{2}$ or loose 2 Euro with probability $\frac{1}{2}$. In choice 2 they had to decide if they want to participate in a lottery that consisted of four independent repetitions of the lottery in choice 1. Subjects were told that in the end of the experiment one of the two

[9] See Holt and Laury (2002) for the multiple price list format.

[10] Instructions are provided in Appendix A.

choices was randomly selected and implemented. In treatments 3 and 4 we also elicited a risk measure. Subjects faced 25 choices between a lottery and a fixed payment. The lottery was the same across choices and paid zero or 3 Euro, each with probability 0.5. The fixed amount was increased in 10 Cent increments, starting from 30 Cent and going up to 270 Cent. Again, one choice was randomly picked and implemented (see Dohmen et al. (2011)).

Note that our central measures of interest (choice between clumped and piecewise information in treatments 1 and 2 and willingness to pay for the lottery in treatments 3 and 4) were all elicited on the first day of the experiment, i.e., on Monday. When subjects left the laboratory, they received a letter which reminded them of their duties for the next days, i.e., when to call the experimenter and when to pick up the money. After all subjects had left the lab on Monday, the experimenter conducted the dice rolls. From Tuesday to Thursday subjects had to call the experimenter once a day.[11] Subjects were told that failing to call would lead to the loss of all their earnings from the experiment.[12] During the phone calls, subjects received information about the outcome of the lottery. In the clumped information condition, subjects were informed on Tuesday whether they won in the lottery or not and which numbers were thrown for them. In the piecewise condition, subjects received one piece of information each day. Thus they usually did not know before Thursday whether they won in the lottery or not. Note that in both conditions subjects had to call once a day from Tuesday to Thursday and that the duration of the phone calls always was approximately one minute. This was made clear to subjects in the instructions. On Friday subjects had to come to the experimenter's office to receive their earnings from the experiment.

Note that information in this setting is non-instrumental in the sense that the lottery is an exogenous event which cannot be influenced by the subjects. One might however argue that information has at least some instrumental value as it may allow subjects to improve their decision on whether to stop participating in the experiment or not, i.e., to stop calling or not to pick up their money, depending on their chances of winning the lottery. If this were the case, subjects should have preferred the clumped condition over

[11] Subjects could call from 9am to 12pm and from 2pm to 5:30pm. Alternatively they could show up personally in the experimenter's office which only one subject chose to do.

[12] In treatments 3 and 4, some subjects did not participate in the lottery but received a fixed payment, depending on the outcome of the price list format. Nevertheless these subjects still had to call from Tuesday to Thursday and this was made clear in the instructions.

the piecewise condition, because it provided them with all the information on Tuesday, allowing them to decide on Tuesday whether the revenues from the experiment outweigh the cost of calling and picking up the money. We argue that the minimum payoff from the experiment (15 Euro) is big enough for subjects to continue with the experiment, even if they know they lost in the lottery. This is supported by the low number of subjects who failed to call or to collect their revenues from the experiment and by the fact that these numbers do not differ between treatments.[13] Furthermore, in case this argument were valid, it would only bias our results in favor of Kőszegi and Rabin's model.[14]

All experiments were conducted using paper and pencil. A total of 104 subjects participated in the experiment, 24 in treatments 1 and 2 respectively, 30 in treatment 3 and 26 in treatment 4. We ran 2 sessions per treatment. Subjects were students from different fields.

1.2.3 Hypotheses

Here we intuitively derive the predictions of Kőszegi and Rabin (2009). In Appendix A we formally derive the proposition by Kőszegi and Rabin (2009) that individuals prefer information in one piece, and derive predictions for our treatments.

A central assumption in Kőszegi and Rabin (2009) is that utility depends on anticipated changes in beliefs about current and future consumption. Beliefs are derived from rational expectations and people are loss averse with regard to changes in their beliefs. Loss aversion in belief changes implies an aversion towards gradual resolution of uncertainty. Piecewise information exposes people to fluctuations in their beliefs. Since bad news decrease utility more than good news increase it, these expected fluctuations in beliefs do not cancel in utility terms. Consequently people seek to avoid piecewise information.

In addition the model assumes that people care (weakly) less about changes in beliefs, the further away the time of belief change lies from the actual point of consumption. In

[13]In treatment 1, one subject failed to collect its revenues. In treatment 3 one subject failed to call, in treatment 4, 2 subjects failed to call.

[14]One might also argue that information in our setting might be instrumental in the sense that early information allows subjects to improve their inter-temporal consumption smoothing. We believe that this effect is negligible in our setting, given that consumption smoothing occurs over a whole life-span. Again, if this effect were present, it would only bias our results in favor of Kőszegi and Rabin's model.

other words, a person is assumed to be less sensitive to changes in beliefs, the more time lies in between news and the time of consumption. This implies that people (weakly) prefer to receive information sooner rather than later. Note that this assumption also has consequences for the preference for clumped information. When comparing conditions where information is clumped to piecewise information conditions, the time information arrives is necessarily affected. It is impossible to collapse different pieces of information into one piece, without changing the time the pieces of information are received. Therefore, the precise prediction of Kőszegi and Rabin (2009) is that people prefer to receive information clumped rather than piecewise, as long as no information is delayed through clumping.

Therefore, subjects in treatments 1 and 2 should strictly prefer the clumped condition over the piecewise condition and consequently select the clumped condition.

HYPOTHESIS 1: *In treatments 1 and 2 subjects choose to receive information in one piece.*

Likewise, the model predicts that the average willingness to pay for the lottery should be higher in treatment 3 (where subjects receive clumped information) compared to treatment 4 (where subjects receive information piecewise).

HYPOTHESIS 2: *Average willingness to pay for the lottery should be higher in treatment 3 compared to treatment 4.*

1.3 Results

First, consider treatments 1 and 2, where subjects could directly choose between the two information conditions. *Figure 1.2* summarizes results from the two treatments. In treatment 1, only 11 out of 24 subjects preferred to receive information clumped rather than piecewise. In treatment 2, 14 out of 24 preferred the clumped information condition. Comparing choices between treatments 1 and 2, we find no significant difference. Using a Fisher Exact Test we cannot reject the null-hypothesis choices do not differ between the treatments (p-value is 0.56). Using a simple Probit regression, regressing the choice between the information conditions on a constant and a treatment dummy delivers similar results. The coefficient of the treatment dummy is not significantly different from zero (p-value =0.39).

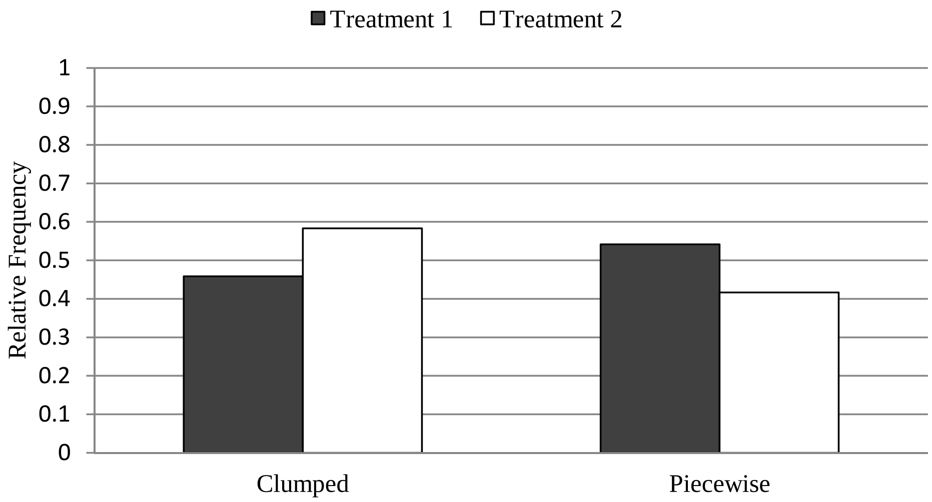


Figure 1.2: Relative frequency of choices (clumped or piecewise information) for treatments 1 and 2.

Given that we find no treatment difference, we henceforth analyze pooled data for treatments 1 and 2. 25 out 48 subjects preferred to receive information in one piece. This is clearly inconsistent with Kőszegi and Rabin (2009), which predicts that all subjects should prefer the clumped information condition. However, when evaluating the predictive power of the model with our data, we need to incorporate an error structure that captures possible inconsistencies and mistakes of subjects. Thus, the statistical model we evaluate is one where subjects make a mistake with probability p_e. Since the model predicts that all subjects prefer the clumped condition, p_e denotes the likelihood that the piecewise condition is chosen. With probability $(1 - p_e)$ they make no mistake and choose the clumped condition. As a first step we simply assume $p_e = 0.2$, i.e., we evaluate the model allowing for error rates of up to 20 percent. Given that Kőszegi and Rabin (2009) predict a strict preference for clumped information, we believe that an error rate of 0.2 is fairly high. We use a simple Binomial Test to test the null hypothesis that our data are generated by a preference for clumped information and an error rate of 20 percent or lower, i.e., that $p_e \leq 0.2$. We reject the null hypothesis at any conventional level (p-value < 0.00001).

In the next step we ask which error rate we would have to assume such that the data is compatible with the model's prediction, i.e., such that we cannot reject the null hypothesis that people prefer clumped information. More precisely, we ask for which value of p_e we cannot reject the nullyhypothesis (at the 5 percent level) that people prefer clumped information, using a one-sided Binomial Test. We find that this threshold value of p_e is 0.354. Thus, we cannot reject the null hypothesis that $p_e \leq 0.354$ (p-value = 0.0502). We conclude that our data is only compatible with the prediction of Kőszegi and Rabin (2009), if we are willing to assume that subjects make mistakes with a probability of more than 35 percent.

It might be that people are heterogenous in their attitudes towards different resolutions of uncertainty. Thus one could ask which individual characteristics determine preferences towards the resolution of uncertainty. Obvious candidate is our measure of loss aversion. The aversion to piecewise information in Kőszegi and Rabin (2009) is driven by loss aversion in belief changes. Thus, it could be that more loss averse subjects have a preference for clumped information. We split our sample according to a high or low degree of loss aversion.[15] For subjects with a low degree of loss aversion, 57.14 percent preferred the clumped condition over the piecewise condition. For subjects with a high degree of loss aversion, exactly 50 percent preferred to receive information clumped. Thus we do not find that subjects with a high degree of loss aversion prefer the clumped condition more frequently.

We summarize our findings from treatments 1 and 2 as follows:

RESULT 1: *Putting treatments 1 and 2 together, only 25 out of 48 subjects preferred the clumped information condition. We can reject the hypothesis that people prefer clumped over piecewise information, even if we allow for error rates of 20 percent. Our results are only compatible with a preference for clumped information if we are willing to assume error rates of more than 35 percent.*

Next, consider behavior in treatments 3 and 4. Average willingness to pay for the lottery is 25.93 Euro and is below the expected value of the lottery of about 32 Euro. *Figure 1.3* shows a histogram of subjects' willingness to pay for the lottery for both treat-

[15] Recall that we used two lottery choices to elicit a measure of loss aversion. We classify subjects who reject both gambles as having a high degree of loss aversion. Subjects who accept both gambles or reject the gamble in choice 1 and accept the gamble from choice 2 are classified as having a low degree of loss aversion. Note that a total of 5 subjects did not behave consistently in the two loss aversion choices. Inconsistency means that they reject the gamble from choice 2 but accept the gamble from choice 1.

ments. Kőszegi and Rabin (2009) predict that subjects should have a higher willingness to pay in treatment 3, where information was clumped. However, *Figure 1.3* suggests the opposite. While about 37 percent of subjects in the clumped condition (treatment 3) report a willingness to pay of 23 Euro or lower, only about 27 percent do so in the piecewise condition (treatment 4). On the other hand, while about 27 percent in the piecewise condition report a willingness to pay of 32 Euro or higher, the fraction is only 3 percent in the clumped condition. The average willingness to pay is 24.83 in treatment 3 compared to 27.19 in treatment 4. Using an OLS regression, regressing willingness to pay for the lottery on a constant and a treatment dummy, we can reject the null hypothesis that willingness to pay is higher in the clumped information condition (p-value < 0.05).[16] This result is robust when controlling for our measure of risk aversion or gender.[17]

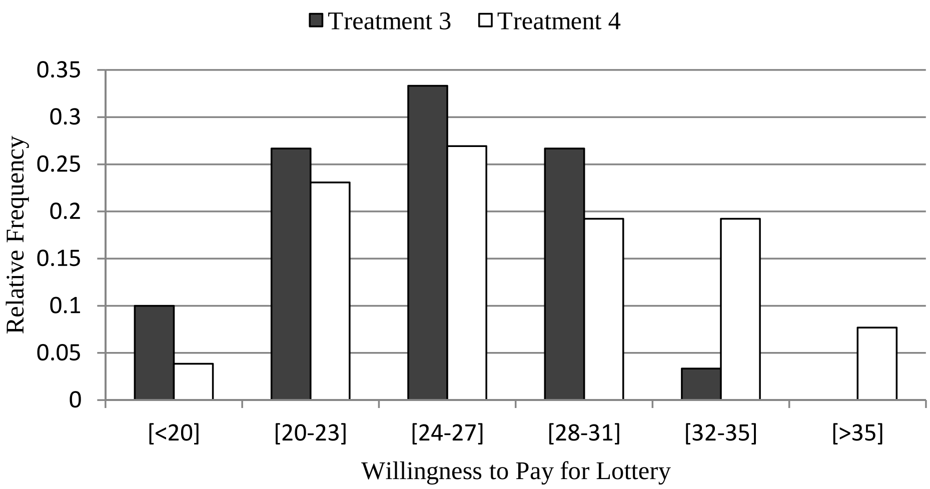


Figure 1.3: Relative frequency of willingness to pay for lottery for treatment 3 (clumped information) and treatment 4 (piecewise information).

Now consider our measure of loss aversion. Again, we split our sample according to high and low degree of loss aversion.[18] For subjects with a low degree of loss aversion,

[16]Kőszegi and Rabin (2009) provide a directed null hypothesis to test. Consequently we use one-sided test statistics to test their predictions.

[17]Note that out of the 56 subjects, 3 failed to make consistent choices in the multiple price list format. In the analysis above we used their average switching point in the price list format. Our results are robust when using the first switching point instead, or excluding them from the sample. When using the first switching point for these three subjects, average willingness to pay for the lottery is 24.17 in treatment 3 and 26.69 in treatment 4. Regressing willingness to pay for the lottery on a constant and a treatment dummy, we can still reject the null hypothesis that willingness to pay is higher in the clumped information condition (p-value < 0.05). When we exclude the three inconsistent subjects from the sample, average willingness to pay for the lottery is 24.85 in treatment 3 and 27.16 in treatment 4. Regressing willingness to pay for the lottery on a constant and a treatment dummy, we again reject the null hypothesis that willingness to pay is higher in the clumped information condition (p-value $= 0.06$).

[18]Note that a total of 9 subjects did not behave consistently in the two loss aversion choices.

average willingness to pay is 25.64 in the clumped information treatment and 28.38 in piecewise information treatment. For subjects with a high degree of loss aversion, this difference is smaller. Average willingness to pay is 24.42 in the clumped condition and 24.44 in the piecewise condition. The smaller treatment difference is somewhat in line with Kőszegi and Rabin (2009). Note however that also for subjects with a high degree of loss aversion, average willingness to pay is not higher in the clumped information condition.

RESULT 2: *The average willingness to pay for the lottery is higher in treatment 4 compared to treatment 3. We reject the null hypothesis that subjects have a higher willingness for the lottery when information is clumped.*

1.4 Conclusion

We examined individuals' attitudes towards information regarding exogenous events. While "standard" theory predicts indifference between different types of resolutions of uncertainty, other theories propose that people care about the timing of information. In this chapter we used a controlled lab experiment to test a prediction developed in Kőszegi and Rabin (2009) that people prefer to receive information in one piece rather than piecewise. Our experimental data does not support the hypothesis that piecewise information is utility-decreasing. In the following we discuss several possible explanations.

First, there is a general problem of testing dynamic models of decision making as these models usually do not specify the length of a time period. In principle time periods could be seconds, minutes, days or months. From a theoretical perspective this makes perfect sense. In fact it seems impossible to determine exact specifications as these are likely to depend on various factors, e.g, the context of the decision-problem. From an empirical perspective this is challenging. It could be that failure to support Kőszegi and Rabin (2009) is due to failure to create timing structures that affect different time periods. Note, however, that in our experiment we made a high effort to make variations in the timing structure meaningful by running the experiment over days. Note also that while leaving the length of a time period open, Kőszegi and Rabin (2009) also do not exclude any specifications.

Second, while we did not find support for the hypothesis on an aggregate level, it might be that some subjects do have preferences for receiving information in one piece. People might be heterogenous in their preferences for different information structures and it would be interesting to analyze which individual characteristics determine these preferences. Note, however, that our data on individual degree of loss aversion does not deliver a subgroup that shows a clear preference for information in one piece.

Third, the prediction that piecewise information is utility decreasing might only hold in certain decision environments. The model by Kőszegi and Rabin (2009) requires people to anticipate utility consequences of future belief changes and incorporate them in their current choices. These anticipation effects might only be present in contexts of particular significance, e.g., news about the own health condition or that of close relatives, news about the future career or maybe news about important political events. Note, however, that expected payoffs and payoff differences between winning and losing of the lotteries in our experiment are rather large. In one treatment, the payoff difference between winning and losing was 500 Euro, which is probably more than half of the monthly income of an average student in our sample.

Chapter 2

Image and Misreporting

2.1 Introduction

Individuals hold private beliefs about their performance, skills, abilities and achievements. Transmission of this private information is crucial for the efficiency of economic interactions. For instance, efficient allocation of tasks within a firm relies on information about employees' skills and abilities. The same is true for decisions about job promotions or efficient specialization. In insurance contexts, the design of efficient insurance plans is difficult when individuals hold private information about their underlying risk. In this chapter, we analyze whether individuals' image concerns can lead them to misreport private information in situations, where from a traditional, purely pecuniary perspective, truthful revelation would be optimal. Individuals that care about how they are perceived by their environment, will take this perception into account when making choices or assessing own performance and abilities in front of others. We illustrate with a simple model how image concerns make people misreport their own performance, skill or ability. In equilibrium, some individuals with low performance will choose to report high performance. Consequently, reports become less informative. Then we provide evidence from a lab experiment documenting the consequences of a desire for a favorable image on statements about own performance.

In our model, decision makers' choice is to publicly report private information about their own type. Correctly stating their private information is optimal in direct monetary terms. However, we assume that decision makers' utility consists of two components, a "standard" part, reflecting direct monetary concerns and an image part, reflecting reputational concerns. The way we model image concerns is a shortcut that captures all potential benefits from signaling a high type. The nature of reputational concerns could be strategic. In labor market contexts, signaling of abilities and skills may improve hiring prospects and lead to higher wages or promotion. Benefits could also be in the form of social approval. Alternatively, decision makers could value reputation for hedonic reasons. People simply enjoy being regarded as a high type. We show the existence of a unique Perfect Bayesian Nash equilibrium, where a decision maker misreports private information. Low skilled types choose to signal a high type, if image concerns are relevant. We also show that misreporting increases in the relative importance of image utility. Apart from sending positive signals about skills and abilities, our model also captures situations where decision makers might want to appear humble or modest in front of others. If modesty is the dominant signaling motive, misreporting might go in the opposite direction, i.e., decision makers downplay own skills and display underconfidence. While we focus on social image concerns, our model is also compatible with a self-signaling interpretation where decision makers learn about their own type by inference from own choices (e.g., as in Bénabou and Tirole (2004) and (2006)).

We test the main prediction of our model, that image concerns lead to misreporting of private information, in a laboratory experiment. The experiment has two stages. In stage 1, subjects go through a series of general knowledge quiz questions. In stage 2, subjects are asked to give a binary and incentivized self-assessment concerning their quiz performance. We study two main treatments: In the audience treatment, we exogenously increase subjects' image concerns in stage 2 by making them report their self-assessment to the other subjects present in the lab. After all subjects have given their binary assessment, one after the other has to stand up and report his or her self-assessment to the group. This procedure has been used by Ariely et al. (2009) to increase image concerns in the context of prosocial decision making. In the private treatment, subjects do not report their assessment to the group. Our data reveals significant evidence for image effects. In the audience treatment, stated self-assessments are significantly higher than in the private treatment. We also document a gender difference in stated self-assessments. This difference seems to be driven by a stronger response of men to the presence of an audience.

To further assess if subjects' reports are also affected by a desire to appear modest in front of others, we conduct a feedback treatment. The treatment is identical to our audience treatment. The only difference is that after subjects report their self-assessments to the group, the experimenter will also report the true performances to the group. If misreporting in the audience treatment was solely driven by the signaling of skills or ability, we should observe that reported self-assessments do not differ between the private and the feedback treatment, because in the latter, true performances will be revealed. If concerns to appear modest are relevant, we should observe stated self-assessments in the feedback treatment below the level we found in the private treatment. When comparing stated self-assessments between the feedback and private treatment, we find no evidence for modesty concerns on the aggregate level. However, we do find some evidence that subjects with rather low quiz performance want to appear modest in the feedback treatment.

Our findings show that image concerns play an important role in the transmission of private information about skill, ability or performance. Even if truthful reporting is optimal in monetary terms, decision makers misreport. This contributes to a large literature that has documented significant biases in stated self-assessments. If individuals are asked to assess their own type in terms of performance or ability, their self-assessments are frequently overly optimistic. One of the most prominent examples of highly optimistic beliefs is a study by Svenson (1981) on relative self-assessments in the context of car driving skills. He finds, for instance, that 40% of subjects place themselves in the top 20% of car drivers with regard to driving skills.[1] Our theoretical and experimental results suggest, that documented biases in self-assessments might be produced by a desire to gain a favorable image. By trying to signal a high type, decision makers appear overconfident. This can occur even with perfect knowledge about their own performance, skill or ability. Decision makers can appear overconfident without any inherent biases in self-assessments. Thus, in our approach, overconfident behavior is rather the outcome of a preference, e.g. a desire to signal skills or ability, than a mistaken self-perception. This might explain why people do not "learn" about their overconfidence over time.

Our findings are also relevant from a mechanism design perspective. They show that mechanisms designed with a purely monetary focus do not necessarily lead to truthful revelation of private information. If people have strong image concerns, these ought to be taken into account when designing optimal mechanisms, e.g., insurance plans or employ-

[1]Other empirical studies on overconfidence include for example Camerer ad Lovallo (1999) and Hoelzl and Rustichini (2005). For a recent overview, see Benoit and Dubra (forthcoming). Several studies examine the consequences of overconfidence for behavior in different contexts. Examples are Dohmen and Falk (2011) in the context of tournament entry, Malmendier and Tate (2008) for CEO behavior or DellaVigna and Malmendier (2006) for overestimation of future gym attendance.

ment contracts. Likewise, our findings are informative from a methodological perspective. They suggest that appropriate monetary incentives alone might not be sufficient to ensure truthful revelation of self-assessments in experiments or surveys. The presence of image concerns creates a trade-off between image concerns and monetary outcomes which leads to biases in stated self-assessments. Minimization of image concerns via, for instance double-blind procedures, might provide a solution to this problem.

While our focus is on direct transmission of private information, our results apply more generally. In many decision contexts that require prior self-assessment, decision makers' choices allow them to signal skill, ability or performance to others. Consider the choice to enter a tournament. The decision to enter or not clearly depends on individuals' private self-assessment. The money-maximizing choice for individuals with low skills and abilities is usually not to enter the tournament. In the presence of image concerns, however, individuals with low skills might yet decide to enter, as this allows them to signal skill and abilities to others. In the context of participation in welfare programs, image concerns and social approval seeking might lead to low participation rates due to fear of reputation losses. Moffitt (1983) presents data from different welfare programs in the U.S. in the 1970's. He reports that as much as 30 - 60 % of the citizens who are eligible for welfare do not apply and argues that this is a consequence of the fear of stigmatization of welfare recipients.

This chapter relates to a few recent papers that considered the social signaling component of biases in self-assessments. Burks et al. (2010) compare different explanations for overconfidence in a large survey study with truck drivers. Their results suggest a strong connection between image concerns and overconfidence. Truckers reporting that they care about how others perceive them, significantly overplace their performance in an IQ test and a numeracy task. Charness et al. (2011) provide experimental evidence that men exploit the possibility to send an exaggerated productivity signal in a strategic interaction of a tournament entry to deter entry of other individuals while women do not. In their paper, they also find evidence for a consumption value from overconfidence.[2] or Moebius et al. (2011)). In a related experiment, Reuben and Rey-Biel (2010) find that

[2]Eil and Rao (2011) and Moebius et al. (2011) also provide evidence for a consumption value from overconfidence. Bénabou and Tirole (2002) provide a theoretical argument for a value of overconfidence as it can increase motivation of individuals with imperfect willpower. Other models have assumed a value of self-confidence and show how overconfident self-assessments can be produced by selectively choosing information or by asymmetrically processing information (see Brunnermeier and Parker (2005), Kőszegi (2006)

subjects exaggerate past performance in order to become a group leader.

More broadly, this chapter relates to several papers that study consequences of image concerns on economic decision making in different contexts. So far the literature has mainly analyzed effects of social approval for prosocial decision making. Non-anonymity or the presence of an audience has been shown to increase prosociality (see Gächter and Fehr (1999), Rege and Telle (2004), Andreoni and Petrie (2004) and Ariely et al. (2009)). Theoretical papers analyzing image concerns in a prosocial context include Bénabou and Tirole (2006), Ellingsen and Johannesson (2008) and Andreoni and Bernheim (2009). Closest to our modeling approach is the paper by Bénabou and Tirole (2006). They show how extrinsic incentives can crowd out prosocial behavior, because they destroy the image rewards from prosocial activity.

The remainder of this chapter is organized as follows. The next section introduces our model. Section 2.3 presents the experimental design, section 2.4 the results from our experiment and section 2.5 concludes.

2.2 The Model

We provide a simple framework that allows illustrating how image concerns can influence reports of private information. The next two sections introduce the model, assuming that decision makers have perfect knowledge about their performance. In section 2.2.3 we relax the assumption of perfect knowledge. In section 2.2.4 we show how a desire to appear modest could be captured with our model framework.

2.2.1 Set-Up

Consider decision makers D that differ in a parameter p which is an element of $P = \{0, 1, ..., \bar{p}\}$. Depending on the context, p captures D's ability, skill, performance or achievement. p is D's private information but is commonly known to be distributed according to a probability function f defined over P. At first, we assume that decision makers have perfect knowledge about p. In section 2.2.3 we provide a version of the model where decision makers have imperfect knowledge about their type and show that this produces qualitatively the same results. Decision makers choice x is to report some measure related to p in public. We assume a binary report: is p larger than some value

$\overline{r}$? This report could be absolute (is p higher than a certain number?), or relative to others (is p higher than the average or the median performance of other decision makers?). Thus, we have that $x \in \{Yes, No\}$. Decision makers win a monetary prize y if their stated report is correct, otherwise they earn 0. Thus, choice x and prize y reflect contexts where truthful reporting of private information is optimal in direct monetary terms. In experimental settings, choice x and prize y simply capture an incentivized self-assessment. More generally, choice x could be a decision that depends on p, e.g., the choice to enter a tournament, and the prize y reflects direct monetary consequences from that choice. Note that the prize y might also capture direct non-monetary utility consequences from misreporting, e.g, costs of lying.[3]

We assume that utility has two sources, direct (monetary) payoffs and image utility. Money enters linearly in the utility function and the two components are additively separable. Thus utility is given by

$$U(x) = y\mathbf{1}(x) + \alpha\beta E(p \mid x).$$

The first part captures direct utility over money. $\mathbf{1}(x)$ is an indicator function taking the value 1 if the stated report is correct and 0 otherwise. The second part incorporates image utility. $E(p \mid x)$ is the public's expectation about D's performance, skill or ability p, conditional on D's choice x. Thus, the public infers decision makers' p from their reports, and social approval depends on that. α and β specify the strength of image concerns. α is an individual parameter, i.e., decision makers differ in α. Some D care more about their image or respond more strongly to social approval than others. α is assumed to be constant across contexts and environments. While α is D's private knowledge, it is commonly known to be drawn from a distribution described by a density function g over $[0, \overline{\alpha}]$ with $g(\alpha) > 0, \ \forall \alpha \in [0, \overline{\alpha}]$. We assume that performance or ability p and the desire for social approval α are drawn independently. β instead, is identical for all decision makers and we assume $\beta > 0$. β might depend on the context of the decision problem, e.g., the size of the public, the social distance between D and the public or the strategic value of a favorable image. Thus, β is the parameter that is exogenously manipulated

[3]Gneezy (2005) and Fischbacher and Heusi (2008) examine lying behavior in different contexts. They find evidence that subjects lie, but also that there is some cost of lying that prevents subjects from lying 100%. Note that throughout the chapter we focus on direct monetary utility, but always mean to include non-monetary interpretations such as costs of lying.

in our experiment. An alternative interpretation of decision makers' image concerns is a desire for a positive self-image (similar as in Bénabou and Tirole (2004) and (2006)).[4] In this case, decision makers receive a private signal about their performance or ability prior to their decision. Thus, when deciding, they hold information about their p. However, for their later self-evaluation, this knowledge is not available for example due to reasons of imperfect recall. Since actions are easier to recall than signals, decision makers base their self-evaluation on past stated reports.

2.2.2 Equilibrium

We now show under which conditions there exists a unique Perfect Bayesian Equilibrium where decision-makers misreport their private information. In the absence of image concerns, D's behavior would be straightforward. Decision makers choose $x = Yes$, if their performance, skill or ability p is higher than $\overline{r}$ and $x = No$ otherwise. In the presence of image concerns however, there exists a trade-off between stating a truthful report and gaining social approval. In equilibrium, all decision makers with $p > \overline{r}$ will choose $x = Yes$. For decision makers with $p < \overline{r}$ there exists a threshold type α^* such that all D with $\alpha > \alpha^*$ will choose $x = Yes$ and all D with $\alpha < \alpha^*$ will choose $x = No$. This is stated formally in the following Proposition:

PROPOSITION 1: If $\overline{\alpha}$ is sufficiently large, i.e.,
$\overline{\alpha}\beta \left[\sum_{p>\overline{r}} \frac{f(p)p}{\sum_{p>\overline{r}} f(p)} - \sum_{p\leq\overline{r}} \frac{f(p)p}{\sum_{p\leq\overline{r}} f(p)}\right] > y$, there exists a unique Perfect Bayesian Equilibrium where decision makers with $p < \overline{r}$ and $\alpha > \alpha^*$ choose $x = Yes$. Decision makers with $p > \overline{r}$ choose $x = Yes$ and those with $p < \overline{r}$ and $\alpha < \alpha^*$ choose $x = No$.

Next, we verify that behavior described above is indeed an equilibrium and show that if $\overline{\alpha}$ (the highest possible realization of α) is sufficiently large, such an equilibrium always exists. In Appendix B we show that this equilibrium is unique.

First, we state precisely what we mean by $\overline{\alpha}$ being sufficiently large. We assume that there exist decision makers with image concerns large enough, such that they choose $x = Yes$ if $p < \overline{r}$ and all other decision makers simply maximize monetary outcomes.

[4] In the psychology literature, the idea that people construct their self-image from past actions can be found in Bem (1972).

More precisely, $\overline{\alpha}$ is large enough, such that the image gains from choosing $x = Yes$, $\overline{\alpha}\beta\left[\sum_{p>\overline{r}} \frac{f(p)p}{\sum_{p>\overline{r}} f(p)} - \sum_{p\leq\overline{r}} \frac{f(p)p}{\sum_{p\leq\overline{r}} f(p)}\right]$, outweigh the monetary costs y.

In equilibrium, all D with $p > \overline{r}$ choose $x = Yes$. It is straightforward to show that this is optimal, given that it maximizes both monetary outcomes and image utility. For decision makers with $p < \overline{r}$, behavior depends on the strength of image concerns. There exists a threshold type α^*, such that all D with $p < \overline{r}$ and $\alpha > \alpha^*$ will choose $x = Yes$ and those with $p < \overline{r}$ and $\alpha < \alpha^*$ choose $x = No$. The threshold type α^* with $p < \overline{r}$ must be indifferent between potential image gains from choosing $x = Yes$ and monetary losses from reporting incorrectly. We have the following indifference condition:

$$\alpha^*\beta\left[\sum_{p>\overline{r}} f(p)p + \int_{\alpha^*}^{\overline{\alpha}} g(z)dz \sum_{p\leq\overline{r}} f(p)p\right] \frac{1}{\sum_{p>\overline{r}} f(p) + \int_{\alpha^*}^{\overline{\alpha}} g(z)dz \sum_{p\leq\overline{r}} f(p)} = y + \alpha^*\beta \sum_{p\leq\overline{r}} \frac{f(p)p}{\sum_{p\leq\overline{r}} f(p)}. (2.1)$$

The left hand side captures image utility in case D chooses $x = Yes$, which is simply a weighted average of the average performance, skill or ability of decision makers with $p > \overline{r}$ and those with $p < \overline{r}$, with weights depending on how many Ds misreport. The right hand side captures image utility when choosing $x = No$, which is simply the average performance or ability of Ds with $p < \overline{r}$ plus the prize y for reporting correctly. Rearranging equation 1 leads the following:

$$\alpha^*\beta\left[\frac{1}{\sum_{p>\overline{r}} f(p) + \int_{\alpha^*}^{\overline{\alpha}} g(z)dz \sum_{p\leq\overline{r}} f(p)} \left(\sum_{p>\overline{r}} f(p)p + \int_{\alpha^*}^{\overline{\alpha}} g(z)dz \sum_{p\leq\overline{r}} f(p)p\right) - \sum_{p\leq\overline{r}} \frac{f(p)p}{\sum_{p\leq\overline{r}} f(p)}\right] = y. (2.2)$$

One can see from equation (2) that decision makers with $\alpha < \alpha^*$ and $p < \overline{r}$ optimally choose $x = No$. As the expression in square brackets (gains in reputation) remains unchanged, but the strength of image concerns is smaller ($\alpha\beta < \alpha^*\beta$), image gains in total weigh less in utility terms than monetary losses, i.e., they will state a truthful report $x = No$. Ds with $\alpha > \alpha^*$ instead optimally choose $x = Yes$ as their image gains loom larger than their monetary losses. Note also, that if $\overline{\alpha}$ is sufficiently large, the threshold type α^* and thus the equilibrium, always exists. To see this, take the left hand side of equation (2) and vary α^*. If α^* approaches zero, the left hand side approaches zero as well. As α^* approaches $\overline{\alpha}$, the left hand exceeds y by assumption. Furthermore, the left hand side is continuous and strictly increasing in α^*. Consequently, there necessarily exists an α^* for which equation (2) holds.

PROPOSITION 2: An increase in β reduces the threshold type α^*. Consequently, more decision makers with $p < \bar{r}$ misreport by choosing $x = Yes$.

Proposition 2 shows how reports change in β, for example, when the size of the public, the social distance between D and the public, or the strategic value of reputation changes. Considering equation (2), one can see that a change in β affects the threshold type α^*. An increase in β reduces the threshold type, in other words, more decision makers with $p < \bar{r}$ will choose $x = Yes$. Thus, our model predicts that an exogenous increase in image concerns increases the number of decision makers that misreport information. Consequently, reports become less informative.

2.2.3 Model with Imperfect Knowledge

So far, we assumed that decision makers perfectly know their p. However, one could argue that in most real-life situations, individuals only have imperfect knowledge about their skills or abilities. Also, in our experiment subjects are likely to be uncertain about their performance. In this section, we analyze what happens if decision makers have imperfect knowledge about their type but know more than the public. The crucial difference to the case with perfect knowledge is that type-uncertainty weakens the informativeness of decision makers choices. Intuitively, it is more difficult for the public to infer ability from choices, if decision makers themselves are uncertain about their ability.

The set-up is identical to above. The only difference is that decision makers do not perfectly know their p. Instead, they hold a point belief $\hat{p} \in \{0, 1, ..., \bar{p}\}$ and $\hat{p}$ is (potentially) different from p.[5] D's choice x is again to report whether p is larger than some value $\bar{r}$, i.e., $x \in \{Yes, No\}$. Given their imperfect knowledge about p, it is possible that decision makers wrongly assess whether their p is larger or smaller than $\bar{r}$. We specify the imperfect knowledge about p as follows. Let $\phi(p)$ denote the likelihood that decision makers point belief $\hat{p}$ is larger (smaller) than $\bar{r}$ although the true p is smaller (larger). Thus $\phi(p)$ is the probability that $\hat{p} > \bar{r}$ although $p < \bar{r}$ or $\hat{p} < \bar{r}$ although $p > \bar{r}$. We make the following assumptions about $\phi(p)$. First of all, we naturally assume that $\phi(p) < \frac{1}{2}$ for all p. Second, we assume that $\phi(p)$ is strictly increasing in p for $p < \bar{r}$, and strictly decreasing in p for $p > \bar{r}$. In other words, the likelihood that Ds think that their p is

[5]To focus on the effect of type uncertainty on the informativeness of choices, we abstract from risk by assuming point beliefs about ability.

larger (smaller) than $\bar{r}$, although it is smaller (larger) increases the smaller the difference between p and $\bar{r}$. Intuitively, the binary self-assessment should be easier, the further away actual performance is from $\bar{r}$ and consequently, the frequency of mistakes should be lower.

We now show that decision makers still have incentives to misreport their private information $\hat{p}$. The key difference between a set-up with imperfect knowledge and one with perfect knowledge is, that the public's inference about performance from choices x changes. Since the public is aware that decision-makers only have imperfect knowledge about their performance, the informativeness of reports x about performance p is reduced. However, the informativeness does not vanish. One can show that if all decision makers report truthfully, i.e. they maximize monetary utility in the absence of image concerns, the public infers higher ability from reports $x = Yes$ compared to reports $x = No$, that is $E(p \mid x = Yes) > E(p \mid x = No)$. We have that $E(p \mid x = Yes) = \frac{\sum_{p>\bar{r}}(1-\phi(p))f(p)p+\sum_{p<\bar{r}}\phi(p)f(p)p}{\sum_{p>\bar{r}}(1-\phi(p))f(p)+\sum_{p<\bar{r}}\phi(p)f(p)}$ is greater than $E(p \mid x = No) = \frac{\sum_{p<\bar{r}}(1-\phi(p))f(p)p+\sum_{p>\bar{r}}\phi(p)f(p)p}{\sum_{p<\bar{r}}(1-\phi(p))f(p)+\sum_{p>\bar{r}}\phi(p)f(p)}$.

Thus, we can state the following proposition:

PROPOSITION 3: If $\overline{\alpha}$ is sufficiently large, there exists a unique Perfect Bayesian Equilibrium where decision makers with $\hat{p} < \bar{r}$ and $\alpha > \alpha^*$ choose $x = Yes$. Decision makers with $\hat{p} > \bar{r}$ choose $x = Yes$ and those with $\hat{p} < \bar{r}$ and $\alpha < \alpha^*$ choose $x = No$.

Proposition 3 corresponds to Proposition 1 in the set-up with perfect knowledge.[6] In Appendix B we state formally the requirement that $\overline{\alpha}$ is sufficiently large. Proposition 3 shows that also with imperfect knowledge, decision makers have incentives to misreport private information. The intuition is simple. Although decision makers are not perfectly informed about their own skills, performance or ability, they know more than the public. Consequently reports x have some informative value for the public and thus the signaling motive for decision makers still exists.

For variations in common image utility β, the same comparative statics hold as in section 2.2.2.

PROPOSITION 4: An increase in β reduces the threshold type α^*. Consequently, more decision makers with $\hat{p} < \bar{r}$ misreport by choosing $x = Yes$.

[6]The logic of the proof is the same as for Proposition 1.

2.2.4 Modesty

In addition to signaling skill or ability, there might exist other signaling motives. Decision makers may want to appear humble or modest in front of others, i.e. they downplay own skills or achievements when reporting them to others. In our model, there is a simple way to capture parts of this additional signaling motive. We model modesty in the context of self-assessments as a reluctant view about oneself. This can be signaled towards others by reporting low self-assessments.

We assume that decision makers hold a point belief $\hat{p} \in \{0, 1, ..., \bar{p}\}$ and $\hat{p}$ is potentially different from p. $\hat{p}$ is D's private knowledge, but is commonly known to be distributed according to a probability function h defined over P. We now consider mistaken point beliefs in more detail. First, take decision makers with $\hat{p} < p$. Instead of being interpreted as a simple consequence of receiving imperfect signals on p, downward biases in self-evaluation are now being interpreted as a character trait, namely having a modest and reluctant view about oneself. By the same argument, biases upwards ($\hat{p} > p$) could capture traits such as overconfidence. We now want to allow for the signaling of modesty or (not) signaling overconfidence. Utility is given by

$$U(x) = y\mathbf{1}(x) + \alpha\beta E(\hat{p} \mid x).$$

The fact that D's image or approval stems from the public's belief about $\hat{p}$ captures the additional signaling motive. Depending on the strength of the correlation between $\hat{p}$ and actual performance p, image concerns still reflect the desire to appear skilled and able to the public. In addition, now decision-makers might also want to signal modesty or not being overconfident. The sign of common image utility β determines which signaling-motives dominate. Positive values of β capture dominance of the desire to signal skill and ability. Negative values of β might reflect situations where social approval stems from modesty.

The structure of this game is identical to that in the model with perfect knowledge. Consequently, we can state the following proposition.

PROPOSITION 5: If $\overline{\alpha}$ is sufficiently large, i.e. $\mid \overline{\alpha}\beta \left[\sum_{\hat{p}>\overline{r}} \frac{h(\hat{p})\hat{p}}{\sum_{\hat{p}>\overline{r}} h(\hat{p})} - \sum_{\hat{p}\leq\overline{r}} \frac{h(\hat{p})\hat{p}}{\sum_{\hat{p}\leq\overline{r}} h(\hat{p})}\right] \mid > y$, there exists a unique Perfect Bayesian Equilibrium. If $\beta > 0$, decision makers with $\hat{p} < \overline{r}$ and $\alpha > \alpha^*$ choose $x = Yes$. Decision makers with $\hat{p} > \overline{r}$ choose $x = Yes$ and those with $\hat{p} < \overline{r}$ and $\alpha < \alpha^*$choose $x = No$. If $\beta < 0$, decision makers with $\hat{p} > \overline{r}$ and $\alpha > \alpha^*$ choose $x = No$. Decision makers with $\hat{p} < \overline{r}$ choose $x = No$ and those with $\hat{p} > \overline{r}$ and $\alpha < \alpha^*$choose $x = Yes$..

Proposition 5 shows how image concerns for skill, ability or modesty affect stated self-assessments. The sign of common image utility β determines which signaling motive dominates. If $\beta < 0$, decision-makers in equilibrium underreport their private information in order to appear humble or modest. If $\beta > 0$, decision makers overreport due to a desire to signal ability or skill. Which of the motives dominates in reality is ultimately an empirical question and our experiment can be viewed as an attempt to answer it.

2.3 Experimental Design

Our model suggests that the desire for social approval will tempt decision makers to misreport their private information in public. To test this hypothesis, we introduced a simple choice environment where subjects held private information about their skill or performance. Then, we manipulated image concerns exogenously by varying whether private information is reported to a public or not.

Table 2.1 summarizes our experimental between-subjects design. We study two main treatments, an *audience treatment* and a *private treatment*. In both treatments, the experiment started with a short introductory game. Subjects one after the other were asked to stand up and provide the group with some personal information such as name, age, and field of study.[7] The main part of the experiment consisted of two stages. In stage 1, subjects were asked to answer 20 multiple-choice quiz questions. The questions covered various general knowledge topics such as history, economics, math, or art. Subjects were given four possible answers and had to select one. We incentivized the quiz, such that subjects earned 40 cents for every correct answer. The number of correctly answered questions serves as our measure of performance. Subjects received no feedback regarding the number of correctly answered quiz questions. Therefore, they held private but not necessarily perfect information about their performance. In stage 2, subjects faced a

[7]The purpose of the introductory game was to reduce the social distance between participants. Gächter and Fehr (1999) show in the context of a public goods game that social approval incentives are only effective in combination with a procedure to increase familiarity among group members.

simple incentivized self-assessment task.[8] We asked them to assess whether their own performance was better or worse than the average quiz-performance of a group of other participants.[9] The group of other participants consisted of 95 different subjects who had also performed the quiz. Subjects received 5 euros for a correct self-assessment. The only difference between our two treatments was the following: In the *audience treatment*, all subjects entered their self-assessment into the computer, and then reported their self-assessment to the other subjects present in the lab. Subjects knew in advance that they had to report their assessment to the other subjects. Thus, after all subjects privately assessed their relative quiz-productivity and entered it in the computer, one after the other had to stand up, say their name and report their self-assessment to the group.[10] This procedure of introducing an audience to increase image concerns has been used for example in Ariely et al. (2009) in the context of pro-social behavior. The *private treatment* was identical to the audience treatment, however subjects did not state their self-assessment towards the other subjects.

The experiment ended with a questionnaire. We elicited subjects' risk preferences, image concerns and several socio-demographic characteristics like gender and age. We measured subjects' risk preference by using a question from the German Socio-Economic Panel Study (GSOEP): "How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: 'completely unwilling to take risks' and the value 10 means: 'completely willing to take risks'." Dohmen et al. (2011) show that this survey question is very well suited for analyzing risk preferences, because it is highly correlated with incentivized lottery choice measures. In addition, we used the question "How important is the opinion of others to you?" as a survey-based measure of image concerns. Subjects could choose between five answers from 'not at all' to 'very important'. To gain further insights on whether a high or low stated self-assessment is associated with social (dis)approval, we asked two questions. First, we were interested in whether subjects enjoyed the quiz

[8]Subjects were only informed about the self-assessment task after they finished stage 1.

[9]Studies that want to document relative overconfidence usually use comparisons to percentiles such as the median. For our question, identifying overconfidence is not the main goal, because we are particularly interested in the treatment effect on reported self-assessments. Therefore, we decided to use the simpler and more comprehensive average as measure of comparison.

[10]While subjects reported their private information (self-assessments) in front of the audience, their previously entered self-assessment was also shown on their computer screen to make sure subjects could not lie about their entered self-assessment.

("How much did you enjoy the quiz?"). Second, we asked whom subjects would hire if they were the boss of a firm on the basis of reported self-assessments. The three possible answers were: Somebody who reports 'better than average', 'worse than average', and 'I do not care'. In addition, we wanted to learn more about subjects' perception of others' self-assessments. We asked: "When stating their self-assessments, do you think the other participants overestimated, underestimated or correctly estimated their performance?"

The version of our model we presented in section 2.2.4 suggests that subjects might signal more than just ability through their choice. The stated self-assessment can also be informative about traits like modesty. While we suspected that the desire to signal skills or ability would dominate, we still wanted to analyze if concerns to appear modest affected stated self-assessments as well. Therefore, we ran a *feedback treatment* where we controlled for the signaling of ability. The feedback treatment was identical to the audience treatment. The only difference was that after subjects reported their self-assessment, the experimenter informed the public whether the assessment was correct or not. In this situation, the public learns the true relative performance and therefore subjects can no longer signal ability. However, they could signal modesty by reporting to be worse than average. Thus, if misreporting in the audience treatment is solely driven by the signaling of skills or ability, we should observe that reported self-assessments should not differ between the private and the feedback treatment. However, if concerns to appear modest are relevant, stated self-assessments in the feedback treatment should be below the level of the private treatment.

Stage 1	Stage 2	Treatments	Questionnaire
Multiple-choice quiz - Number of correct answers is our measure of performance - 40 cents / correct answer	Self-assessment - Are you better or worse than the average? - 5euros / correct self-assessment	1. Private: no further action 2. Audience: reporting self-assessment to an audience 3. Feedback: public feedback after reported self-assessment	- Risk - Survey questions - Demographics

Table 2.1: Design of the experiment

2.3.1 Experimental Procedures

A total of 143 subjects participated in the experiment, 47 in the private treatment, 48 in the audience treatment, and 48 in the feedback treatment. We were interested in potential gender differences and therefore invited an equal amount of women and men to each session. All six sessions of the experiment were conducted in the *B*onn*E*con*L*ab, subjects were recruited via ORSEE (Greiner (2004)) and the experiment was run using the experimental software z-Tree (Fischbacher (2007)). A session took on average 50 minutes and subjects earned 9.50 euros[11] on average. We distributed the instructions for stage one and two immediately before the stage started and they were read aloud.[12]

2.3.2 Hypotheses

In the experiment, we systematically increase image concerns of subjects by introducing an audience. When comparing the private and the audience treatment, we hypothesize that signaling ability is the dominant signaling motive. Thus, by Propositions 2 and 4 of our model, reported self-assessments should be higher in the audience treatment compared to the private treatment.

HYPOTHESIS 1: *Subjects choose "better than average" more frequently in the audience treatment than in the private treatment.*

Our feedback treatment allows us to control for the signaling of ability. Thus, if a desire for appearing modest is present, stated self-assessments should be lower in the feedback treatment compared to the private treatment.

HYPOTHESIS 2: *Subjects choose "better than average" less frequently in the feedback treatment than in the private treatment.*

[11] 1 euro was worth about 1.4 Dollars at the time.

[12] Instructions can be found in Appendix B.

2.4 Results

We start with the analysis of our two main treatments. In section 2.4.1, we compare reports of the audience and the private treatment. In addition, we show the influence of gender on our treatment effect and analyze individuals' perceptions of others' stated self-assessments. In section 2.4.2, we present results from the feedback treatment.

2.4.1 Main Results

RESULT 1 *There is a treatment difference in stated self-assessments: Subjects in the audience treatment report "better than average" significantly more often compared to subjects in the private treatment.*

Dependent variable: Relative self-assessment= $\begin{cases} 1 & \text{if better} \\ 0 & \text{if worse} \end{cases}$

	(1) All	(2) All	(3) All	(4) Private	(5) Audience
Dummy treatment	0.20**	0.25**	0.27**		
	(0.09)	(0.12)	(0.12)		
Quiz performance		0.07***		0.07*	0.08**
		(0.03)		(0.04)	(0.04)
Dummy quiz performance			0.16		
			(0.11)		
Dummy gender		-0.31***	-0.37***	-0.21	-0.28**
		(0.11)	(0.10)	(0.18)	(0.12)
Controls		included	included	included	included
N	95	95	95	47	48
-LL	62	50	53	24	18

Notes: Probit estimates. Marginal effects (evaluated at the mean of independent variables) reported; robust standard errors are in parentheses. Significance at the 1, 5, and 10 percent level is denoted by ***, **, and *, respectively. *Dummy treatment* =1 if audience treatment and 0 if private treatment. *Dummy gender* =1 if female. *Dummy quiz performance* =1 if better than average. *Controls* include the survey based risk measure, image concerns, age, and relationship status.

Table 2.2: Determinants of stated self-assessment

We find that 68% of subjects in the audience treatment report to be "better than average", compared to 48% of subjects in the private treatment. This sizable effect is also statistically significant in probit regressions. Table 2.2 reports the marginal effects of three probit regressions (columns 1-3), regressing a treatment dummy and several controls on reported self-assessment, where 1 indicates a report "better than average".[13] Column 1 of Table 2.2 shows that the raw treatment effect is significant at the 5% level. Our finding is robust when controlling for different measures of quiz performance, gender and several additional controls. In column 2, we take the number of correctly solved quiz questions as a control for quiz performance. In column 3, we use a different measure: we create a performance dummy, taking the value one if performance was actually better than average and zero otherwise. In both regressions, the treatment effect remains significant at the 5% level.[14] In addition, nonparametric testing with a Fisher-exact test also confirms result 1 ($p - value = 0.06$, two-sided).

A different way to look at our data is to analyze the treatment effect for different intervals of actual quiz performance. According to our model, the treatment effect should be driven by subjects who place themselves below the average, when privately evaluating own performance, but want to signal high quiz performance towards others. In line with section 2.2.3 of our model, we assume that most subjects with low quiz performance privately place themselves below average, while those with high performance, mostly place themselves above average. Consequently, our model predicts that stated self-assessments for subjects with rather high quiz performance should be similar between treatments, while reports for subjects with low quiz performance should differ between treatments. This is indeed what we find. Figure 2.1 depicts the percentages of subjects in the audience and the private treatment who report to be better than average for different intervals of actual quiz performance, centered around the average of the comparison group (14.3 questions). Among subjects that clearly solved more questions than average (more than 15 correctly solved questions), 72 % report to be better than average in the audience treatment, compared to 69 % in the private treatment. For subjects with low quiz performance (less than 13 correctly solved questions), however, we have a very pronounced treatment difference. While 57 % report to be better than average in the audience treatment, only 27 % do so in the private treatment. This suggests, in line with our model, that our treatment effect is mainly driven by subjects who privately place themselves below average, but want to signal high performance towards others.

[13]Columns 4 and 5 of Table 2.2 are discussed later.

[14]Note that the average quiz performance over all treatments is 14.4 correctly solved quiz questions. The distributions of quiz performance do not significantly differ across treatments ($p - values > 0.34$ of Kolmogorov-Smirnov tests). The comparison group of 95 participants had an average quiz performance of 14.3 which is also not significantly different from performances of subjects in our treatments.

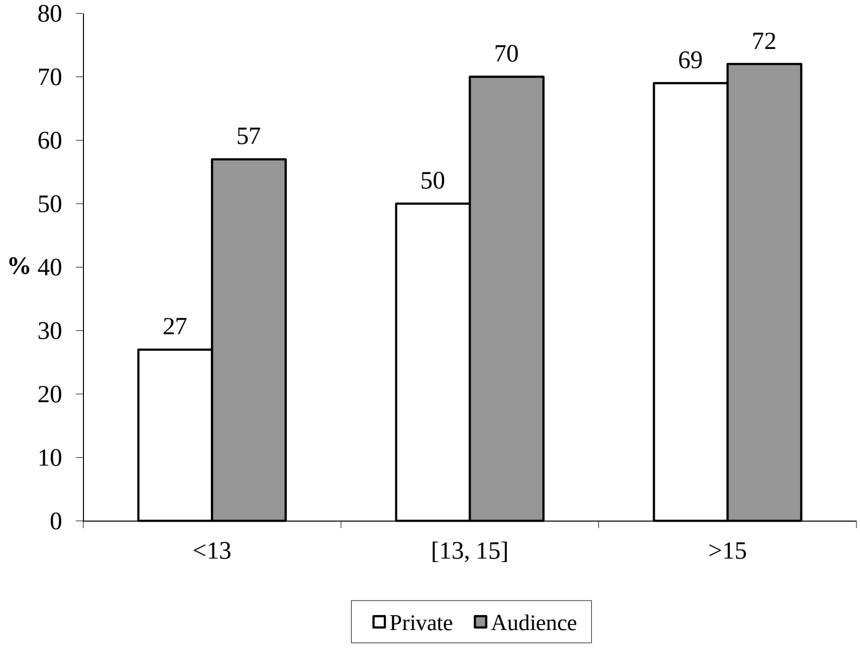


Figure 2.1: Percentage of "better than average" reports for high, low and close to average quiz performance, for subjects in the audience treatment and the private treatment.

Additional, more indirect evidence that high reported self-assessments are associated with social approval comes from two survey questions. First, 64% of our subjects stated that they enjoyed the quiz or enjoyed it very much. Only 10% indicated they did not like the quiz. Second, when subjects were asked to imagine they owned a firm and had the opportunity to hire new workers, none of the subjects was willing to hire a worker that reports "worse than average" in the audience treatment and only 13% would do so in the private treatment.

RESULT 2 *There is a gender difference in reported self-assessments: Men report "better than average" significantly more often. This difference seems to be driven by a stronger response of men to the presence of an audience.*

We find a gender difference in reported self-assessments. By inspection of Table 2.2 we find in regressions (2) and (3) that the probability to choose "better" is higher for men than women. The marginal effect of the gender dummy is significantly different from zero. A gender difference in self-assessments has been found in many studies and provides a possible explanation for documented gender differences in selection into competitive environments (see for example Gneezy and Rustichini (2004), Gneezy et al. (2003), Niederle and Vesterlund (2007) and Dohmen and Falk (2011). Columns 4 and 5 of Table 2.2 show separate Probit regressions for the private and the audience treatment. The data indicates that the gender effect is mostly driven by more men overreporting in the audience treatment. While men report to be "better than average" significantly more often than women in the audience treatment, the effect is not significant in the private treatment. This finding might provide a possible explanation for gender differences in overconfidence. It suggests, that men feel a stronger desire to signal skills or abilities towards others which results in overconfident appearance.[15]

RESULT 3 *The public is aware of misreporting due to image concerns when evaluating subjects' reports.*

Does the audience anticipate that the report "better than average" might be driven by image concerns? To answer that question, we asked our subjects in the questionnaire about their perception and beliefs regarding the reported self-assessments of the other participants in the experiment. Table 2.3 summarizes the answers. We find that a majority of subjects in the audience treatment (56%) thinks that others misreport and state too optimistic assessments. Only 26% hold a similar view in the private treatment. A Fisher-exact test confirms a significant difference ($p = 0.01$), where we categorize subjects' perceptions in "overreport" or not. Thus, we find evidence that the audience anticipates misreporting and adjusts beliefs accordingly. This finding supports the mechanism of our model. The decision maker chooses to signal a high self-assessment, the public anticipates this and adjusts beliefs about the decision maker's type downwards.[16]

2.4.2 Feedback Treatment

RESULT 4 *Comparing the feedback treatment to the private treatment, we do not find evidence for strong modesty concerns on the aggregate level.*

[15] Note, however, that this interpretation should be taken with caution. In Appendix B we report the marginal effects of a probit regression with interactions of a gender dummy and treatment dummy (I_Treatment*Women). The marginal effect of this interaction describes the difference of the gender effect in the audience treatment compared to the private treatment. The difference is negative. In line with our interpretation, men report especially in the audience treatment that they are better than average, however not significantly more often than in the private treatment.

[16] Ludwig and Nafziger (2011) explore subjects' beliefs about other subjects' confidence bias and find that the majority believes that others are unbiased, and only few think that others are overconfident.

	Private treatment	Audience treatment
Overreport	26 %	56 %
Correct	40 %	42 %
Underreport	34 %	2 %

Table 2.3: Subjects' beliefs about the other participants' self-assessments

We now analyze results from our feedback treatment. The purpose of the treatment was to identify whether subjects' reports are also affected by concerns for appearing modest in front of others. We find that 56% of subjects choose to report "better than average". Compared to the private treatment with a frequency of 48%, there is no significant difference ($p - value = 0.54$ using a Fisher-exact test). Table 2.4 reports the marginal effects of probit regressions with and without controls for the private and feedback treatment. The treatment effect is insignificant in all regressions. When asking our subjects about their perception of other subjects' reports, 81% indicated that they think others chose a correct self-assessment. Thus, on an aggregate level, we do not find evidence for strong modesty concerns.

Note, however, that we do find evidence for modesty concerns for subjects with low quiz performance. Figure 2.2 depicts the percentages of subjects in the feedback and the private treatment who report to be better than average for the same intervals of actual quiz performance as in section 2.4.1 (Figure 2.1). Among subjects with rather low quiz performance (less than 13 correctly solved questions) 27 % report to be better than average in the private treatment, while only 10 % do so in the feedback treatment. However, this desire to appear modest seems to counteract with a desire to appear confident in own skills or performance for high performance subjects. Similar to a desire to appear modest in front of others, some subjects might want to display confidence in their own performance. Figure 2.2 shows that for subjects with high quiz performance (more than 15 correctly solved questions), the treatment effect goes in the opposite direction. 69 % report to be better than average in the private treatment, compared to 93 % in the feedback treatment. Thus, it seems that the feedback treatment has a differential impact on stated self-assessments, depending on actual quiz performance. Knowing that the experimenter will subsequently provide feedback to the audience, low performance subjects want to

Dependent variable: Relative self-assessment= $\begin{cases} 1 & \text{if better} \\ 0 & \text{if worse} \end{cases}$

	(1)	(2)	(3)
Dummy treatment	-0.07	-0.03	-0.05
	(0.10)	(0.12)	(0.11)
Dummy gender		-0.09	-0.22*
		(0.13)	(0.12)
Quiz performance		0.11***	
		(0.03)	
Dummy quiz performance			0.21*
			(0.11)
Controls		included	included
N	95	95	95
-LL	65	48	54

Notes: Probit estimates. Marginal effects (evaluated at the mean of independent variables) reported; robust standard errors are in parentheses. Significance at the 1, 5, and 10 percent level is denoted by ***, **, and *, respectively. *Dummy treatment* =1 if feedback treatment and 0 if private treatment. *Dummy gender* =1 if female. *Dummy quiz performance* =1 if better than average. *Controls* include the survey based risk measure, image concerns, age, and relationship status.

Table 2.4: Determinants of stated self-assessment in the private and feedback treatment

appear modest, while high performance subjects want to signal confidence in their own performance.

2.5 Concluding Remarks

In this chapter we studied the consequences of image concerns on reports of private information. We illustrated with a simple model how a desire for social approval can give rise to overconfident behavior. In addition, we conducted a controlled lab experiment that supports predictions of our model. In the experiment, subjects stated a higher self-assessment when an audience is present than in private. We also find that men choose

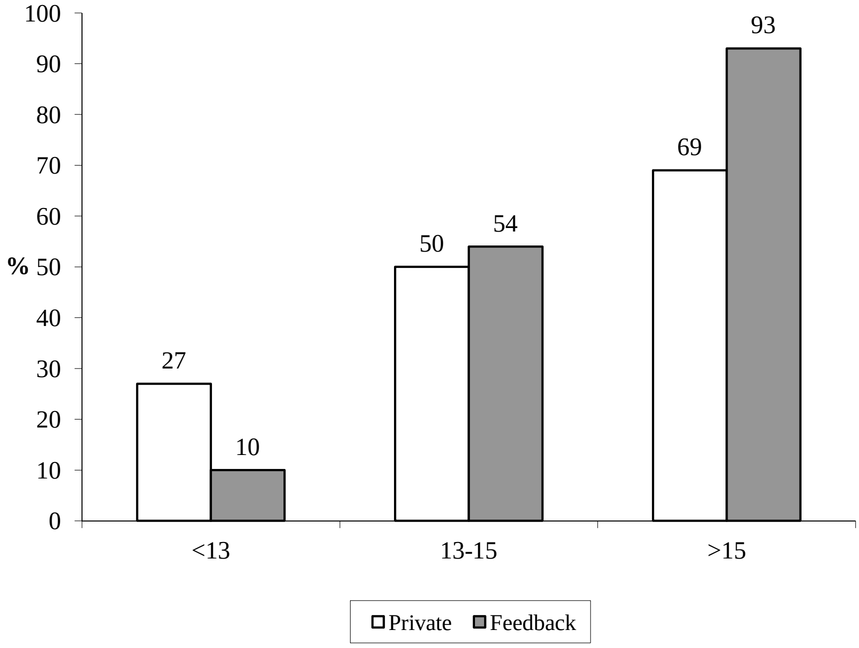


Figure 2.2: Percentage of "better than average" reports for high, low and close to average quiz performance subjects in feedback treatment and private treatment.

more often than women to signal ability and confidence especially when an audience is present.

Our findings show that biases in self-assessments might be produced by image concerns. As a consequence, decision makers can appear overconfident even with perfect knowledge about their own performance, skill or ability, in other words, without inherent biases in self-assessments. This is also an explanation why overconfidence is persistent. Receiving feedback and learning one's type over time might not prevent decision makers from appearing overconfident. Other explanations for overconfident behavior have been suggested: Bénabou and Tirole (2002) provide a theoretical argument for a value of self-serving beliefs as these can increase motivation of individuals with imperfect willpower. Other models assume a value of self-confidence and show how overconfident self-assessments can be produced by selectively choosing information or by asymmetrically processing information, putting more weight on positive than on negative information (see for example Brunnermeier and Parker (2005), Kőszegi (2006) or Moebius et al. (2011)). Recently, several experimental papers have provided support for biases in information-processing and information demand (see Eil and Rao (2011) and Moebius et al. (2011) and Charness et al. (2011)). Benoit and Dubra (forthcoming) provide a different explanation. They argue that most of the evidence for relative overconfidence can in fact be reconciled by correct

Bayesian updating from common priors. In other words, evidence in the form "40% of subjects place themselves in the top 20% of good car drivers" should not be interpreted as evidence for overconfident self-assessments as it can be the outcome of correct updating from unbiased information. While all approaches are important and in concert provide a good explanation for documented behavior, our experimental results highlight the crucial role of image concerns for stated self-assessments.

In our experiment, we manipulated image concerns by letting subjects report their self-assessment to an audience. The audience was mainly composed of students that did not know each other and thus social distance between decision makers and the public was rather high. We expect that in more intense social contexts, e.g. talking to one's supervisor, boss, parents or friends, the magnitude of our finding might be even larger. Furthermore, we did not provide direct strategic reasons for image or reputational concerns. We could have implemented an instrumental value of appearing skilled or able as follows: subjects would randomly be assigned to the roles of principals and agents. In each session there would be twice as many agents as principals. Agents would go through our quiz questions and then (anonymously) state a self-assessment towards the principal assigned to them. The principal has to select one of the two agents for an additional quiz and has incentives to select the agent he thinks is most able. Agents would be given incentives for being selected. We suspect that agents would overstate self-assessments to increase the likelihood of being selected by the principal. Therefore, stated self-assessments in such a treatment should be higher compared to our control treatment.

Finally, while the main focus of the chapter is on social image concerns, our model is also compatible with a self-signaling interpretation. Instead of signaling skill or confidence to others, decision makers care about how they perceive their own self. In this interpretation, self-image is built from past actions. While beliefs about performance are available when making choices, later self-evaluation is built on past actions because actions are easier to recall than beliefs (see Bénabou and Tirole (2004) and (2006)). Although this is not explicitly modeled in our framework, the self-signaling interpretation might give rise to inherent biases in self-assessment. Interestingly, these biases would not stem from selective choice of information or asymmetric information processing (like for example in Brunnermeier and Parker (2005), Kőszegi (2006) or Moebius et al. (2011)) but from self-evaluation based on biased choices.

Chapter 3

Consistency as a Signal of Skills

3.1 Introduction

In this chapter, we examine the role of consistency in actions or statements as a device to signal skills such as ability or personality. We show with the help of a simple model how a rational signaling motive can induce people to behave consistently. Testing the model with a principal-agent experiment we show that: (i) consistency is indeed a signal of ability, (ii) principals understand the informativeness of consistency and (iii) agents anticipate this and, as a consequence, act more consistently. An important implication of the model is that publicly stating a point of view creates a strong pressure to stick to it even in the presence of new and challenging information. In the second part of the chapter we highlight this implication, showing the crucial role of early commitment for subsequent behavior. Furthermore, we underscore the potential of consistency as a device of social influence.

Our model is built on the notion that (in)consistent behavior is associated with (low) skills. To illustrate, consider the following examples: take an engineer who's job it is to assess the energy efficiency of gasoline motors. He is presented a new motor and, after some tests and examinations, states that the motor is very energy efficient. After some weeks he is presented exactly the same motor again. This time, however, after analyzing the motor, he criticizes its poor energy efficiency. Likewise, consider a new colleague who told you that he is a vegetarian and how important this is for him. The next day, you see him in a steakhouse ordering a huge rib-eye steak. What would you infer from this about the engineer's ability or the personality of your new colleague?

In our model there are two types of decision-makers who repeatedly face the same choice problem. High types perfectly know their preferred outcome while low types are uncertain about it. Before making a decision, low types receive noisy signals about their preferred outcome. Regarding behavior over time, high types make consistent choices. However, for low types, beliefs and consequently choices may be inconsistent. We assume that decision-makers' utility consists of two parts, a "standard" part reflecting direct material concerns and a reputational part capturing decision-makers' image concerns for being a high type. This creates a trade-off between choosing according to updated beliefs and reputational concerns.

We show the existence of a Perfect Bayesian Equilibrium where decision-makers display consistent behavior because it allows them to signal skills. Note that the informativeness of consistent behavior is not limited to ability in task-related contexts but applies more generally. Summarizing evidence from social psychology, Cialdini (1984) stresses that consistency is a signal of "personal and intellectual strength". Depending on the context, consistency is associated with different personal characteristics. In repeated social interactions, being a high type, reflected by consistent behavior, signals predictability and reliability which are important prerequisites for relationship formation and trust. Consistency can also be viewed as a sign of personal identity. Identity is shaped by past actions and thus, without continuity in actions, the formation of a sense of self-identity is not possible. The way we use image concerns in our model is a shortcut that captures all potential benefits from signaling a strong type. These benefits can be of monetary nature, e.g., career benefits, but could also reflect social approval or appreciation. In other contexts these benefits could be stable relations or friendships. Benefits could also be hedonic in nature - people simply like it if others think well of them. Note also that the recipient of the signal can be others but also the decision-maker himself. In this case the decision-maker learns about her own type by observing and inferring from her own behavior (e.g., as in Bénabou and Tirole, 2004 and 2006).

Note that we focus on situations where decision-makers face the same choice problems repeatedly and where the preferred outcome remains stable over time. In such situations, skills are reflected by consistent behavior. By contrast, in changing, unstable environments where it is commonly known that the preferred outcome changes, quick adjustment to new information or environments could be seen as a positive trait. In fact sticking to previous points of view could be considered rigid or even stupid. Put differently, if the circumstances that lead to a first decision have clearly changed before a second decision is taken, it is not a sign of strength to stick to the first period's choice. In this case decision-makers will, due to reasons of consistency, make sure that the public knows that circumstances have in fact changed. This explains why decision-makers (a good example are politicians) exert much rhetoric effort to explain the reasons for why they have changed their mind and why acting differently should not be interpreted as inconsistency.

We conduct an experiment that tests the basic logic of our model in a simple principal-agent framework. In particular, we address three questions: First, are consistent agents more able? Second, do principals understand the informativeness of consistency as a signal of ability? Third, do agents anticipate behavior of the principals and use consistency as a signaling device? In the experiment, the decision context is a simple estimation task. Principals receive information about the estimation behavior of two agents and need to select one for an additional estimation task. We find that consistent agents are indeed more able. Principals understand this and select agents who provide consistent estimates significantly more often. Anticipating the signaling value of consistency, agents' estimates in this treatment are more consistent compared to a control treatment where we eliminated the role of the principals, i.e., the need to signal ability.

In the second part of the chapter we highlight two central implication of our model. First, we study the role of early commitment. Intuitively, actively committing to an opinion, belief, intention or action is a precondition for observing consistent or inconsistent behavior. Without commitment, i.e., without taking a stand or an action, possible inconsistencies are impossible to detect. Thus consistency as a signal of skills requires prior commitment. We experimentally test this intuition. Subjects have to perform an estimation task and receive valuable information regarding the solution of the task. In the main treatment, subjects commit to a first estimate prior to receiving the helpful information and without knowing that they will later receive it. After they have received this information they are free to revise their first estimate. In the control treatment, no prior commitment is made. We find that the deviation of the final estimate from the valuable information is significantly higher in the main treatment than in the control treatment. Prior commitment makes subjects neglect valuable information leading to lower payoffs.

Second, we demonstrate how the desire for consistency can be used to influence behavior. We do so in the context of surveys. The way a taste for consistency can affect answers to survey questions is simple. The basic idea is to "tempt" a person to make a biased statement in a first question. In a second step, she faces a question related to that statement and the pressure to respond consistently. Thereby, the simple addition of one (or several) question(s) can affect answering behavior in subsequent questions. We test this intuition in three controlled survey studies. Each survey consisted of one main question of interest and one (or several) additional question(s). We randomly varied whether the additional question(s) was (were) included in the survey or not. In all three surveys, we are able to manipulate responses in a systematic way.

Consistency has further implications, for example for the design of committee or jury procedures. Institutionally requested commitments on a certain opinion at an early stage of negotiations can decrease the quality of final choices as these potentially do not reflect the full level of available information. In bargaining contexts, early requests can increase the danger of bargaining failure as negotiation outcomes below these requests cannot be reached without revealing inconsistency. The desire to be consistent with stated intentions can also be a powerful means to circumvent problems of self-control. Statements like "I want to exercise more" in front of relatives or friends create pressure to live up to that commitment. From a methodological point of view our results suggest interdependencies in behavior that can be relevant for the design of experiments, surveys and empirical research in general. We discuss these implications in more detail in section 3.5.

Our model belongs to a class of models where decision-makers try to signal positive traits or skills through their actions (as for example in Bernheim, 1994, Prendergast and Stole, 1996, Bénabou and Tirole, 2006, Ellingsen and Johannesson, 2008 or Andreoni and Bernheim, 2009). Eyster (2002) and Yariv (2005) have suggested models of consistent behavior. In Eyster (2002) people have a taste for rationalizing past mistakes by taking current actions that justify these mistakes. His model offers an explanation for the well-known "sunk-cost effect". It also predicts procrastination in search contexts and overbidding in wars of attrition. Yariv (2005) proposes a model where people have a taste for consistency of beliefs held over time, thereby reducing cognitive dissonance. Her model can explain a variety of phenomena such as underpinnings of overconfidence and underconfidence, persistence of actions over time or why people sometimes might prefer to receive less accurate information. Different to our approach, Eyster (2002) and Yariv (2005) directly assume a taste for consistency while in our model consistency results from a desire to signal skills. Also, their models focus on internal consistency while our model stresses the role of consistency as a signaling device. Ellingsen and Johannesson (2004) and Vanberg (2008) refer to the taste for consistency as a possible reason for why people incur costs of lying and thus keep their promises. In fact, breaking promises can be viewed as a particularly obvious form of inconsistency.

Our work also relates to a literature in social psychology. Cialdini (1984) summarizes much of the evidence and discusses several explanations for consistent behavior. In particular he highlights the role of consistency as a signal of positive attitudes. An alternative interpretation of consistent choices relies on the notion of cognitive dissonance. This approach basically assumes that consistent behavior reflects a desire to avoid cognitive dissonances. Early work in this direction was developed in Heider (1946), Newcomb (1953) and Festinger (1957).[1] Another potential driver of consistent behavior is based on the idea that thinking is not costless. If thinking does involve cognitive costs, it may in fact be optimal to stick to a particular behavioral strategy and not to change behavior in response to new information or new signals.[2]

Note that our notion of consistency is conceptually different from so-called anchoring effects. In a classical anchoring experiment, Tversky and Kahneman (1974) for example generated a random number between 0 and 100 and asked subjects if the number of African nations in the United Nations was greater than that number. Then they asked subjects to estimate the number of African nations in the United Nations. They find that the randomly generated number (the anchor) significantly affected estimates. Studies on anchoring manipulations differ from the notion of consistency in that they document effects of random unrelated numbers on subsequent choices. In our work, subsequent choices are affected by prior actions or statements. In addition the role of reputation stressed in our model is not present in work on anchoring effects, which also implies that anchoring effects cannot explain our experimental findings.

The remainder of the chapter is organized as follows: In the next section we introduce our model. In section 3.3 we present our first experiment that tests the basic logic of our model. Section 3.4 contains the second experiment that highlights the role of early commitment. Section 3.5 contains the three survey studies and section 3.6 concludes.

[1] Akerlof and Dickens (1982) incorporate cognitive dissonance theory into an economic model. In an empirical study, Mullainathan and Washington (2009) examine the consequences of cognitive dissonance in the context of voting behavior.

[2] In the final section we discuss why neither cognitive dissonance nor cost of thinking can explain the behavioral patterns observed in our experiments.

3.2 The Model

We model the intuition that consistency is a signal of skills. We choose a very simple set-up with two types of decision-makers who choose between two alternatives. The model formalizes a rational signaling motive for behaving consistently and delivers behavioral predictions for our experiments.[3]

3.2.1 Set-up

There are two periods, $t = 1, 2$. In both periods, a decision-maker (D) chooses x_t from a choice set $X = \{Red, Blue\}$ in public. D has a preferred outcome $\mu \in X$. There are two different types of decision-makers. High types D_H (highly skilled types) perfectly know their preferred outcome μ. Low types D_L are uncertain about their preferred outcome. Her type is D's private knowledge but is commonly known to be drawn from a distribution with probability α that D is of type D_H and $(1 - \alpha)$ that she is of type D_L. Low types and the public are holding an uninformative prior on μ, i.e., $Pr(\mu = Red) = Pr(\mu = Blue) = \frac{1}{2}$. In both periods, before making a choice, D_L privately receives a signal m_t about μ. Signals are of strength p_t, i.e., $p_t = Pr(m_t = Red|\mu = Red) = Pr(m_t = Blue|\mu = Blue)$. We assume that signals are informative with $\frac{1}{2} < p_t < 1$. We allow that the strength of signals differs between periods 1 and 2, only requiring that $p_2 \geq p_1$. Thus our setup captures situations where the quality of information D_L receives may change over periods. The assumption that signals in period 2 are at least as strong as signals in period 1 is only made to focus on scenarios where contradicting signals lead to changes in beliefs about the preferred outcome.[4]

Upon receiving a signal m_t, D_L updates beliefs about μ following Bayes' rule. Throughout the chapter $l(m_t, m_{t-1})$ denotes D_L's beliefs about μ in terms of the probability she assigns on Red being her preferred outcome conditional on signals m_t and m_{t-1} (if m_{t-1} exists), i.e., $l(m_t, m_{t-1}) = Pr(\mu = Red|m_t, m_{t-1})$.

Since the prior on μ is uninformative, D_L's updated period 1 belief is determined by

[3] We have also developed an alternative and more general version of the model with a continuous type and choice set. Insights and predictions of this model are qualitatively similar to the more parsimonious set-up. We therefore decided to use the latter. The continuous version is available on request.

[4] If the period 2 signal is less informative than the period 1 signal, updated period 2 beliefs always remain in line with the signal from period 1 even in case of contradicting signals. Consequently types D_L would always choose consistently anyway.

the signal she received in period 1, m_1. Accordingly, the updated belief on μ in period 1 is $l(m_1 = Red) = p_1 > \frac{1}{2}$ and $l(m_1 = Blue) = (1 - p_1) < \frac{1}{2}$, respectively. In period 2, D_L receives an additional signal m_2 and updates again. Thus:

$$
\begin{aligned}
&(1) \quad l(m_2 = Red, m_1 = Red) = \frac{p_1 * p_2}{p_1 * p_2 + (1-p_1)(1-p_2)} > \frac{1}{2} \\
&(2) \quad l(m_2 = Blue, m_1 = Red) = \frac{p_1 * (1-p_2)}{p_1 * (1-p_2) + (1-p_1) * p_2} \leq \frac{1}{2} \\
&(3) \quad l(m_2 = Blue, m_1 = Blue) = 1 - \frac{p_1 * p_2}{p_1 * p_2 + (1-p_1)(1-p_2)} < \frac{1}{2} \\
&(4) \quad l(m_2 = Red, m_1 = Blue) = 1 - \frac{p_1 * (1-p_2)}{p_1 * (1-p_2) + (1-p_1) * p_2} \geq \frac{1}{2}
\end{aligned}
$$

In both periods, the decision-maker chooses x_t in order to maximize utility. D's utility function consists of two components. The first is "standard" outcome-based utility. Standard utility is 1 if D chooses $x_t = \mu$ and 0 otherwise. The decision-maker also cares about her reputation. She receives utility if the public perceives her as being a high type. An alternative interpretation of the reputational concern is a desire for a positive self-image (similar as in Bénabou and Tirole, 2004 and 2006). In this case decision-makers receive a perfect signal about their type prior to their decision. Thus, when deciding, they know their type. However, for their self-assessment, this knowledge is not readily available, e.g., due to reasons of imperfect recall. Since actions are easier to recall than signals, decision-makers use past actions for their self-assessment. Thus the model is compatible with the intuition that people care about their self-image and construct self-image from past actions. In the following we describe the model mostly in terms of public reputation but always mean to include a self-signaling interpretation. Reputational concerns are expressed by

$$-\beta * Pr(type = D_L | x_t, x_{t-1}).$$

$Pr(type = D_L | x_t, x_{t-1})$ denotes the public's (or D's) belief about D's type, conditional on D's decisions in the current period x_t and the previous period x_{t-1} (if it exists). Parameter β is assumed to be positive and specifies how much D cares about her reputation. Differences in β might reflect, e.g., the size and importance of the public or the social distance between D and the public. Note that the way we model reputational concerns is a shortcut that captures all benefits from signaling a strong type. The value of reputation could be instrumental in the sense that people expect benefits in future interactions. In labor relations, e.g., signaling higher abilities may improve hiring prospects or lead to higher wages and promotion.[5] Benefits could also be in the form of social approval or appreciation. Alternatively, β could reflect a hedonic value of reputation. People simply enjoy being regarded as a high type. Depending on the context, skills are associated with different personal characteristics. In task-related choices, similar solutions to similar problems signal high ability. In repeated social interactions, consistent behavior signals predictability and reliability. These are important prerequisites for relationship formation and trust (see Brown, Falk and Fehr, 2004). They also help solving coordination problems. Consistency is also a sign of personal identity. Identity is shaped by past actions or statements. Without continuity in actions or statements, the formation of a sense of self-identity is not possible. In that sense, a high type (via consistent behavior over time) is a prerequisite for strong personal identity.

Putting these two components together, in periods 1 and 2 D_L's expected utility is given by

$$E(U_t(x_t)) = \begin{cases} l(m_t, m_{t-1}) * 1 - \beta * Pr(type = D_L | x_t, x_{t-1}) & \text{if } x_t = Red \\ (1 - l(m_t, m_{t-1})) * 1 - \beta * Pr(type = D_L | x_t, x_{t-1}) & \text{if } x_t = Blue. \end{cases}$$

In periods 1 and 2, D_L maximizes $E(U_t)$ facing a trade-off between maximizing outcome-based utility and gaining reputation.[6]

Note that we assume myopic, non-forward looking behavior. Decision-makers are not anticipating period 2 decisions when deciding in period 1. However, predicted behavior is actually identical if decision-makers are forward looking. We show this formally in Appendix C. Intuitively, choosing according to their beliefs in period 1 is optimal not only from a standard utility perspective. It also maximizes the likelihood of consistent behavior. It is therefore not possible to improve in terms of consistency even in situations

[5]Likewise, in our principal-agent experiment, the signaling of ability increases agents' chances of being selected by the principal.

[6]Utility of the high types is straightforward. Suppose a type D_H with $\mu = Red$. Her utility is described by $U_t(x_t) = \begin{cases} 1 - \beta * Pr(type = D_L | x_t, x_{t-1}) & \text{if } x_t = Red \\ 0 - \beta * Pr(type = D_L | x_t, x_{t-1}) & \text{if } x_t = Blue. \end{cases}$

where the decision-maker anticipates future decisions.

3.2.2 Equilibrium

We now turn to equilibrium behavior. In period 1 the equilibrium we consider is straightforward. D always maximizes standard utility. Types D_L choose $x_1^* = Red$ if $m_1 = Red$ and $x_1^* = Blue$ if $m_1 = Blue$. Types D_H choose their preferred outcome μ. Due to the uninformative prior of the public, D cannot affect her image through x_1, i.e., $Pr(type = D_L|x_1 = Red) = Pr(type = D_L|x_1 = Blue)$. Thus simply maximizing standard utility is optimal and the behavior described above constitutes a Perfect Bayesian Equilibrium.[7]

In period 2 there now exists a choice history x_1^* and D faces a possible trade-off between standard utility and reputational concerns. We characterize the equilibrium conditional on x_1^*. To simplify notation we consider the case where $x_1^* = Red$ without loss of generality.

Types D_H who chose *Red* in period 1 know with certainty that their preferred outcome is *Red*. For types D_L who chose *Red* in period 1 we need to distinguish two possibilities. In period 2 they either receive a signal $m_2 = Red$ or $m_2 = Blue$. In the first case we have that $l(m_2 = Red, m_1 = Red) > \frac{1}{2}$, in the second case $l(m_2 = Blue, m_1 = Red) \leq \frac{1}{2}$. If D would simply maximize standard utility, types D_H and types D_L with $m_2 = Red$ would choose $x_2 = Red$ whereas types D_L with $m_2 = Blue$ would choose $x_2 = Blue$.[8] In the presence of image concerns, however, D_L types with contradicting signals face a trade-off between maximizing standard utility and signaling skills. We show under which conditions there exists a Perfect Bayesian Equilibrium in period 2 where low types with contradicting signals nevertheless choose consistently, thereby sacrificing expected standard utility.

In the equilibrium, types D_H, types D_L with signals $m_2 = Red, m_1 = Red$ and types D_L with signals $m_2 = Blue, m_1 = Red$ all behave consistently choosing *Red* in period 2. For this to be an equilibrium it suffices to check incentive compatibility for low types with contradicting signals. If incentive compatibility is fulfilled for these types it

[7] As is common for this type of signaling models, there exist other equilibria. For example, if reputational concerns are large enough, there exists a Perfect Bayesian Equilibrium where D always chooses *Red* (or *Blue*) regardless of her beliefs about μ and her type and the public holds off-equilibrium beliefs that D is a low type, i.e., $Pr(type = D_L|Blue) = 1$ (or $Pr(type = D_L|Red) = 1$).

[8] If signals in periods 1 and 2 are of equal strength, D_L's with contradicting signals are indifferent in terms of standard utility between *Red* and *Blue*.

is straightforward that for types D_H and types D_L with signals $m_2 = Red, m_1 = Red$ it is fulfilled as well. Decision-makers D_L with signals $m_2 = Blue, m_1 = Red$ hold beliefs $l(m_2 = Blue, m_1 = Red) \leq \frac{1}{2}$. We need to check that they are better off choosing *Red* than choosing *Blue*. Standard utility is $l(m_2 = Blue, m_1 = Red) * 1$ when they choose *Red* and $(1 - l(m_2 = Blue, m_1 = Red)) * 1$ when they choose *Blue*. Reputational utility is $-\beta * Pr(type = D_L|Red, Red) = -\beta * (1 - \alpha)$ if they choose consistently and $-\beta * Pr(type = D_L|Blue, Red) = -\beta$ if they choose inconsistently.[9] We end up with the following condition: D_L's with contradicting signals prefer *Red* over *Blue* if

$$\alpha * \beta \geq 1 - 2 * l(m_2 = Blue, m_1 = Red).$$

The above condition captures the trade-off between standard utility and reputational concerns. The left hand side describes reputational gains from choosing consistently while the right hand side represents the costs in terms of standard utility. If image concerns are sufficiently high, there exists a rational signaling-motive for behaving consistently, because consistent behavior allows the signaling of skills. Note that if reputational concerns are small, there exists an equilibrium in period 2 where D always chooses to maximize standard utility. We summarize this result in the following proposition.

PROPOSITION 1: Assume w.l.g. that $x_1^* = Red$. If reputational concerns β are sufficiently large, i.e., $\alpha * \beta \geq 1 - 2 * l(m_2 = Blue, m_1 = Red)$, there exists a Perfect Bayesian Equilibrium in period 2, where D always chooses $x_2^* = x_1^* = Red$.

Proposition 1 states the main result of our model. In period 2, decision-makers do not simply maximize standard outcome-based utility. Instead, they act consistently with their period 1 choice in order to signal a high type. Thus, in equilibrium, they are willing to sacrifice outcome-based utility to increase their reputational utility.

[9] Note that we assume off-equilibrium beliefs such that the public infers a low type from inconsistent behavior. We believe that this assumption is plausible. High types know their preferred outcome with certainty. Consequently they have the highest expected costs in terms of standard utility from choosing inconsistently. Also, if only a small fraction of decision-makers would not have any reputational concerns and only maximize standard utility, only low types would choose inconsistently. In addition, our equilibrium satisfies the Intuitive Criterion (Cho and Kreps, 1987).

3.3 Consistency as a Signal of Ability

In this section we present evidence from a principal-agent experiment designed to test the basic logic of the model by implementing a strategic motive to signal ability. In particular, we ask three questions: Is consistency a signal of ability? Do principals understand the informativeness of consistent behavior? Do agents anticipate behavior of the principals and use consistency as a signaling device? In total we study three main treatments as summarized in *Figure 3.1*.[10]

Design: The decision context in the experiment was a simple estimation task. In the **agent treatment**, subjects had to perform two estimation tasks. Both tasks consisted of estimating how many times the letter "e" appeared in a text with 1,966 letters. We had selected two texts where the number of e's was identical and subjects were informed about this. The correct number was 233. Subjects saw the first text for 60 seconds on their computer screen. Then they had 60 seconds to state their estimate about the number of e's (first estimate). Without getting any feedback on the first task, agents then saw the second text for 60 seconds. Again they had 60 seconds to provide their estimate for the second text (second estimate). Subjects were paid for accuracy. For both estimates the following rule applied: The maximum profit for each estimate was five Euro, which the agent received if her estimate was less than 1 percent above or below the correct number. For every percentage point the estimate deviated from the correct number, 10 Cents were deducted. If the estimate was more than 50 percent above or below the true value, profits were zero. Negative earnings were not possible.

	Agent Treatment	**Principal Treatment**	**Principal-Agent Treatment**
Questions	- Correlation of consistency and ability - Level of consistency in the absence of a strategic signaling motive	- Do principals understand information content of consistent choices	- Level of consistency in the presence of a strategic signaling motive - Comparison with agent treatment

Figure 3.1: Summary of Treatments for Experiment 1.

The agent treatment allows us to test if consistency of estimates is indeed correlated

[10]Instructions for all experiments we conducted are displayed in Appendix C.

with estimation ability. To analyze if others infer higher ability from consistency, we ran a second treatment, the **principal treatment**. In the principal treatment, subjects were confronted with estimates of subjects from the agent treatment. Each principal was randomly assigned to two subjects from the agent treatment. A principal's choice was to select one of the two subjects. Principals were paid according to the estimation precision of the selected subject. Thus, principals had incentives to select the subject who they thought is most able in solving the estimation task.

For their decision, principals were informed about the absolute difference between the first and the second estimate for both subjects assigned to them. This information was provided on an answer sheet.[11] Note that principals did not know the correct result of the two estimation tasks. However they knew that the correct result was the same in both tasks. On their answer sheet principals had to select "their" subject. Principals were paid according to the accuracy of the selected subject's first estimate. The maximum payoff was ten Euro, which was paid if the subject's estimate was less than 1 percent above or below the correct number. For every percentage point the estimate deviated from the correct number, 20 Cents were deducted. If the estimate was more than 50 percent above or below the correct value, a principal's payoff was zero.

Note that we did not display the four single estimates of the two subjects. The reason is that we wanted to allow principals only the kind of inference the public can make in our model. If we would have given principals the four single estimates, other inferences would have been possible. First, if principals themselves held a belief about the correct result of the estimation task, they could have inferred estimation ability from how close a subject's estimates are to that belief.[12] Second, the two estimates of subject 1 could serve as a signal about the true preferences of subject 2, given that both have the same preference μ (they want to estimate the identical number). If principals observe this signal, subjects' two estimates reveal different information about D's type than in the absence of that knowledge. Therefore we provided only absolute differences between estimates to the principals, thereby only allowing the kind of inference on D's type that is assumed in our model. In footnote 17 we present results from an additional treatment we conducted

[11] Information was provided anonymously, i.e., principals could not link information to actual subjects in the experiment.

[12] As mentioned above, principals did not know the correct result of the tasks. However, they knew that subjects had to estimate how many times a certain letter appeared in a text. Thus it is likely that they held some belief about the correct solution to the estimation tasks.

where principals were given the four single estimates.

To answer our third question (do agents use consistency as a signaling device?) and complete the test of our model we need an additional treatment. In the **principal-agent treatment**, we increase agents' incentives to signal estimation ability compared to the agent treatment. In our model, this increase corresponds to an increase in reputational concerns β. Upon arrival, subjects were randomly assigned to the roles of principals and agents.[13] In each session there were twice as many agents as principals. Subjects were seated in separate rooms according to their roles. The treatment involved two stages. All subjects were informed about both stages at the beginning. In the first stage, agents had to perform the two estimation tasks used in the agent treatment, i.e., agents had to estimate how many times the letter "e" appeared in the two texts. The payoff scheme for the two tasks was identical to that of the agent treatment.

After all agents had completed their two estimates, the second stage began. Two agents were randomly assigned to one principal. The decision of the principal was to select one of the two agents for a third estimation task, which was known to be similar to the first two estimation tasks. The principal was paid according to the precision of the selected agent's estimate in this third task. The payoff scheme for principals was identical to that of the principal treatment. Thus principals had an incentive to select the agent who they thought is most able in solving the estimation task. For their decision, principals were informed about the absolute difference between the first and the second estimate for both agents assigned to them. As in the principal treatment, this information was provided on an answer sheet.

On their answer sheet principals had to select "their" agent. Agents had an incentive to be selected and to estimate as precise as possible in the third task. They received a prize of 10 Euro for being selected. In addition, they were paid according to accuracy identically to the payment scheme in the first two estimates.

Principals' selection decisions from the principal-agent treatment provide additional information on whether estimation ability is inferred from consistent estimates. To examine whether agents anticipate this and actually use consistency as a signaling device, we simply compare estimation behavior between the principal-agent and the agent treatment.

[13]In the instructions we used a neutral framing. Subjects were called "participants A" and "participants B", respectively.

Given that principals selected the most able agent, we implemented a strong motive to signal ability in the principal-agent treatment (compared to the agent treatment).

Procedural Details: A total of 209 subjects participated in ten sessions. In the principal-agent treatment, 64 subjects participated as agents and 32 as principals. 72 subjects participated in the agent treatment. 41 subjects participated in the principal treatment. Subjects were mostly students from various fields at the University of Bonn and were recruited using the online recruitment system by Greiner (2003). No subject participated in more than one session. The experiment was run using the experimental software z-Tree (Fischbacher, 2007). Principals made their choice on an answer sheet. Sessions lasted on average about 60 minutes in the principal-agent treatment and 45 minutes in the agent treatment and the principal treatment. Average earnings were 13.05 Euro for principals and 12.06 Euro for agents, including a show-up fee of eight Euro for principals and four Euro for agents.[14]

Hypotheses: Consistent estimates signal estimation ability. While highly skilled types (in this case high estimation ability) should solve the estimation tasks very well and give consistent estimates, low types only receive noisy signals and therefore are likely to give inconsistent estimates. Thus, on average, more consistent agents (low absolute difference between estimates) should give better estimates. Note that while this should hold true for the agent treatment, the correlation between consistency of estimates and estimation ability might be less tight in the principal-agent treatment. The reason is that given the strategic value of consistency in the principal-agent treatment, low types should try to imitate high types and estimate more consistently. In this pooling equilibrium consistency will be less informative about estimation ability compared to the agent treatment.

Principals value consistent behavior. They infer high ability from consistent estimates. Thus agents who estimate more consistently should have a higher probability of being selected by principals. In our model the value of consistency is expressed as an increase in the reputational concern β. While reputational concerns in the agent treatment (β_c) are not necessarily zero (e.g., due to self-signaling motives), reputational concerns in the principal-agent treatment (β_m) are higher, i.e., $\beta_m > \beta_c$. This follows simply from the strategic value of reputation for high ability. In the principal-agent treatment we imple-

[14]One Euro was worth about 1.35 U.S. dollar at the time.

mented a tournament incentive structure where agents win a prize if they outperform the other agent. These tournament incentives lead to a strategic value of acting consistently that is not present in the agent treatment. A higher β leads to greater importance of the desire to be and appear consistent relative to the goal of maximizing standard utility. Consequently, the likelihood of an equilibrium where all decision-makers behave consistently is higher in the principal-agent treatment compared to the agent treatment. Proposition 1 states the condition under which an equilibrium where all decision-makers act consistently exists: $\alpha * \beta \geq 1 - 2 * l(m_2 = Blue, m_1 = Red)$. It can be easily seen that an increase in β makes this condition more likely to hold.[15] Our main hypotheses can be summarized as follows:

HYPOTHESIS EXPERIMENT 1: (i) *More consistent subjects are more precise in their estimates.* (ii) *The likelihood that an agent is selected by a principal decreases in the absolute difference between first and second estimate.* (iii) *The absolute difference between first and second estimate is smaller in the principal-agent treatment than in the agent treatment.*

Results: The first result concerns the connection between consistency of estimates and estimation precision. As predicted by our model, smaller differences between estimates do reflect differences in estimation ability. More consistent agents are more able: The correlation between absolute difference in estimates and precision of estimates (measured as the sum of absolute distance between first estimate and true value and second estimate and true value) in the agent treatment is strong and significantly positive (corr. coefficient is 0.429, with p-value <0.001)). OLS regression, regressing precision of estimates on a constant and the absolute difference in estimates in the agent treatment yields the same result (coefficient of absolute difference in estimates is 0.569, with p-value < 0.001).

Thus using the difference between the estimates as a signal of ability is justified. Note, however, that the informativeness of differences in estimates as a signal of estimation ability basically vanishes in the principal-agent treatment. The correlation between absolute difference in estimates and precision of estimates in the principal-agent treatment is not significantly different from zero (p-value $=$ 0.47). This suggests, in line with our model,

[15]The simplistic character of this prediction follows from the simple structure of our model. A more general version of the model makes the prediction that an increase in reputational concerns should make all decision-makers behave more consistently, leading to a shift in the distribution of distances between the two estimates. This version with a continuous type and choice set is available upon request.

that in the principal-agent treatment we have a pooling equilibrium where low types imitate high types by estimating consistently. In this pooling equilibrium, consistency of estimates loses its informativeness on estimation ability.

Next we consider the selection decisions of principals. In line with our hypothesis, a higher absolute difference between the two estimates decreases the likelihood of being selected in the principal treatment. *Figure 3.2* shows that the likelihood of being selected is about 86 percent for differences between zero and 20 and declines for larger differences. For differences larger than 100, e.g., the likelihood drops to about 11 percent. The decrease in likelihood is significant as shown by a Probit regression. When we regress the probability of being selected on the absolute differences between the estimates we get a negative and significant coefficient (p-value <0.001). The marginal effect is -0.007, indicating that an increase in the absolute difference by one point decreases the likelihood of being selected by about 0.7 percent. Further evidence comes from the observation that among all principals 78 percent select the agent with the smaller absolute difference. A binomial test rejects the null hypothesis that principals randomized with equal probability (p-value <0.001). At the end of the experiment we asked principals how they made their selection choice.[16] Among those principals who actually selected the more consistent subject, 84 percent indicated that they took a small deviation between the two estimates as a signal of estimation precision. This suggests that principals were fully aware of the informativeness of consistency.[17] [18]

[16]The question we asked was the following: "In the following, please briefly provide us with the reasons for you choice."

[17]Examples of statements from principals were: "I decided on the basis of the smaller difference. Since the correct result was the same in both tasks, I thought that the participant with the smaller difference was also better"; "Smaller difference between estimates seemed more convincing to me"; "The difference between the two estimates of the other participant was too extreme".

[18]We were interested in studying consequences of providing principals with the four single estimates. Therefore we conducted an additional control treatment identical to the principal treatment but principals now received the four single estimates of the two subjects assigned to them. Note that this treatment is not an appropriate test of our model, as principals can now use additional information for their inference that they do not have in our model. 41 subjects participated in that treatment. We find that our main result is robust when providing principals with the four single estimates. When we regress the probability of being selected on the absolute differences between the estimates in the additional treatment, we again get a negative and significant coefficient (p-value = 0.07, two-sided). The marginal effect is -0.003, indicating that an increase in the absolute difference by one point decreases the likelihood of being selected by about 0.3 percent. Among all principals in the treatment 66 percent selected the more consistent subject. A binomial test rejects the nullhypothesis that principals randomized with equal probability (p-value = 0.06, two-sided). Note, however, that more consistent subjects are chosen less frequently compared to the principal treatment. The reason is that some principals indeed used different strategies when inferring estimation ability from the single estimates. When asked at the end of the experiment how they made their selection choice, 24 out of 41 principals indicated that they took consistency of estimates as a

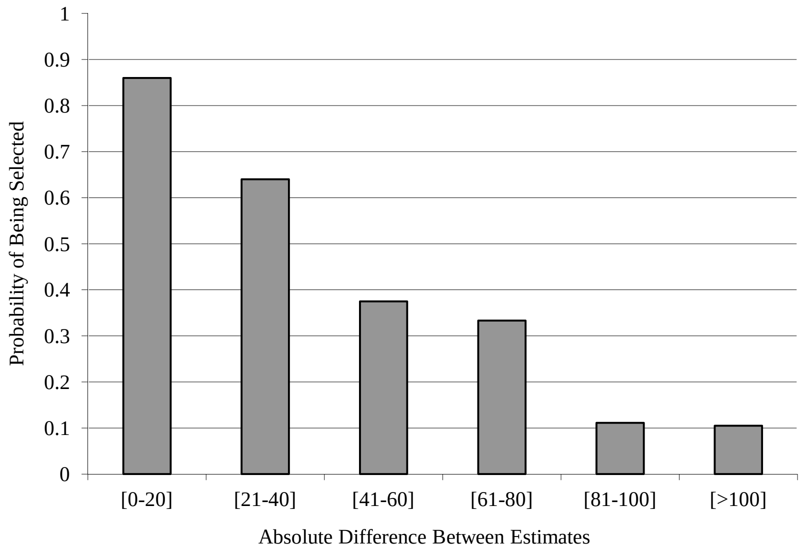


Figure 3.2: Probability of being selected by principal (in the principal treatment) dependent on the absolute difference between two estimates.

Principals' behavior in the principal-agent treatment is very similar to that in the principal treatment. In a Probit regression, regressing the probability of being selected on the absolute differences between the estimates in the principal-agent treatment, we get a negative and significant coefficient (p-value <0.01). The marginal effect is -0.012, indicating that an increase in the absolute difference by one point decreases the likelihood of being selected by about 1.2 percent. Among all principals in the principal-agent treatment, 75 percent selected the agent with the smaller absolute difference. A binomial test rejects the null hypothesis that principals randomized with equal probability (p-value <0.01).

We now turn to agents' behavior. The correct answer for both estimations was 233. Using all estimates (first and second estimates from principal-agent and agent treatment), the average estimate was about 218.[19] The variance in estimates was rather high. The standard deviation of all estimates in the first task was 83.29, in the second it was 74.39. At the end of the experiment, we asked agents in the agent treatment to briefly describe

signal of estimation ability. However, a substantial fraction of principals indicated different strategies. 12 percent indicated that they held a belief about the correct estimation result and chose the subject whose estimates came closest to that belief. 10 percent described that they took the average of all four estimates and selected the subject whose first estimate was closest to that average. In our model these kinds of inferences are not possible. Thus, results of the additional treatment show that principals also value consistent behavior in situations where four estimates are provided, but that alternative or additional inferences are used as well.

[19]The result that the average estimate from a large population is close to the true value is also referred to as the "wisdom of the crowds" (see for example Surowiecki, 2004).

their estimation strategy for the two estimation tasks.[20] Almost all decision-makers who answered the question described a similar procedure. First they counted the number of e's for a couple of rows. Then they counted the total amount of rows in the text and projected the total number of e's in the text.

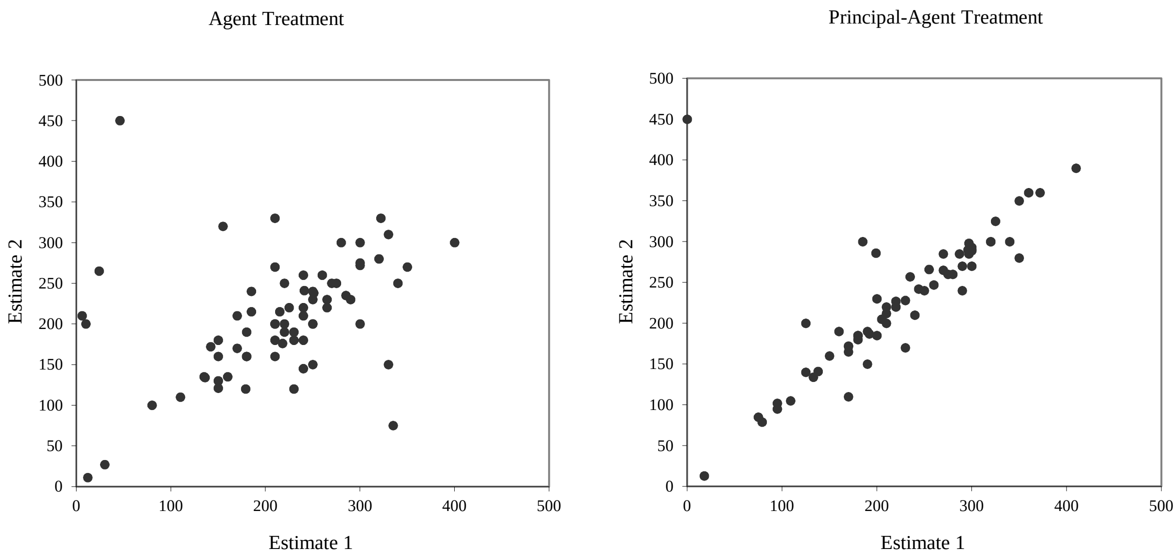


Figure 3.3: Scatterplots of first and second estimates for principal-agent and agent treatment.

Given that principals choose consistent agents, the model predicts that agents choose more consistently in the principal-agent than in the agent treatment. This is in fact what we find. *Figure 3.3* shows scatterplots of first and second estimates. The left panel displays observations from the agent treatment, the right panel from the principal-agent treatment. While first and second estimates are correlated in both treatments, the correlation is much tighter in the principal-agent treatment, i.e., decisions are more consistent. The correlation coefficients are 0.37 in the agent treatment and 0.94 in the principal-agent treatment, respectively. *Figure 3.4* shows a histogram of absolute differences between the two estimates for both treatments. While about 56 percent of agents report an estimation difference below 10 in the principal-agent treatment, the respective number is only 23 percent in the agent treatment. More than 40 percent of agents in the agent treatment indicated differences larger than 30 while less than 15 percent did so in the principal-agent treatment. The average absolute difference between estimates in the agent treatment is 53.8 (std. dev. 70.5) while it is only 17.3 (std. dev. 23.2) in the principal-agent treat-

[20]The question we asked was the following: "Please briefly describe how you proceeded in the two estimation tasks. How did you get to your estimation results?"

ment. This difference is significant at any conventional level (p-value < 0.0001, using either Wilcoxon rank-sum test or OLS, regressing the absolute difference on a constant and a treatment dummy).

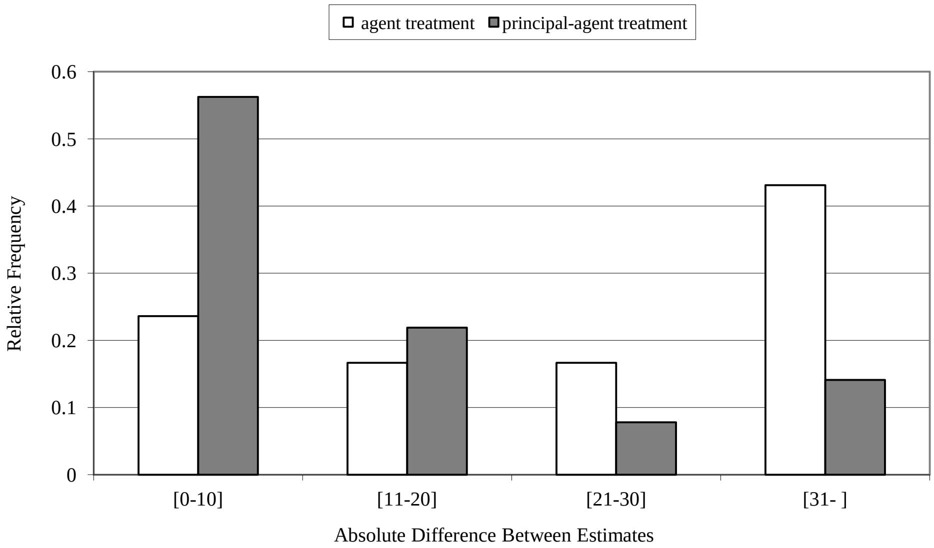


Figure 3.4: Relative frequency of deviations between estimates for principal-agent and agent treatment.

We summarize our main results as follows:

RESULTS EXPERIMENT 1: *(i) Consistency of estimates is a signal of estimation ability. (ii) Principals understand this. The likelihood of being selected decreases in the absolute difference between the two estimates. (iii) Agents anticipate this and, consequently, the absolute difference between first and second estimate is significantly lower in the principal-agent treatment, compared to the agent treatment.*

3.4 The Role of Commitment

Our second experiment studies the role of early commitment. Once an individual has committed to an opinion or belief, she cannot easily change her mind without revealing some inconsistency. In contrast, without commitment observers will not be able to detect possible inconsistencies and therefore decision-makers can maximize utility without taking reputational costs into account. We test this intuition and prediction of our model in the context of an estimation task and show how commitment to an opinion can make people disregard valuable information.

Design: We study two main treatments, one with commitment (main treatment) and one without (control treatment). The different steps of the experiment are illustrated in *Figure 3.5*. The **main treatment** is shown in the upper panel. First, subjects were explained the task: Subjects had to estimate the number of peas in a bowl.[21] Subjects were paid according to the precision of their estimate. If their estimate was less than 5 percentage points above or below the true value of 3000, subjects earned 10 Euro. For every 5 percentage points the estimate deviated from the true value, we deducted 50 Cents. For example, a subject whose estimate deviated 17 percent from the true value earned 8.50 Euro. Negative earnings from the estimation task were not possible.

Subjects were seated around a table which was placed in the middle of the lab.[22] After subjects had been informed about the task the bowl was shown. The bowl with peas was placed in the middle of the table. Subjects were asked to raise their hand once they had written down their estimate on an answer sheet that had been distributed at the beginning of the experiment. As soon as a subject indicated that he or she had written down an estimate, the experimenter went to the subject and recorded the subject's estimate. This means that subjects had written down their first estimate and knew that the experimenter knew that estimate. At this point, subjects had committed to their first estimate. After all subjects had stated their estimates, the experimenter announced that he would now provide subjects with additional and "helpful" information regarding the estimation task. Each subject received an information sheet containing the following sentence. "In the past it has often been the case in various estimation tasks, that the average estimate of all participants is often relatively close to the true value. The estimation task you are facing has also been conducted with a different group of participants. They have also been paid according to precision of their estimates. The average estimate of the number of peas in the bowl of this group was 2615. If you want to, you can now revise your estimate." After they received the information sheet, subjects had time to revise their estimate on their answer sheet. Of course, only the final estimate was relevant for earnings. After all subjects had indicated that they had specified their final estimate, the experimenter collected their answer sheets and the estimation task ended.

The additional information we provided to subjects was based on a separate experi-

[21] A picture of the bowl is shown in Appendix C.

[22] Subjects were seated sufficiently far away from each other, such that they could not see what other subjects were writing down.

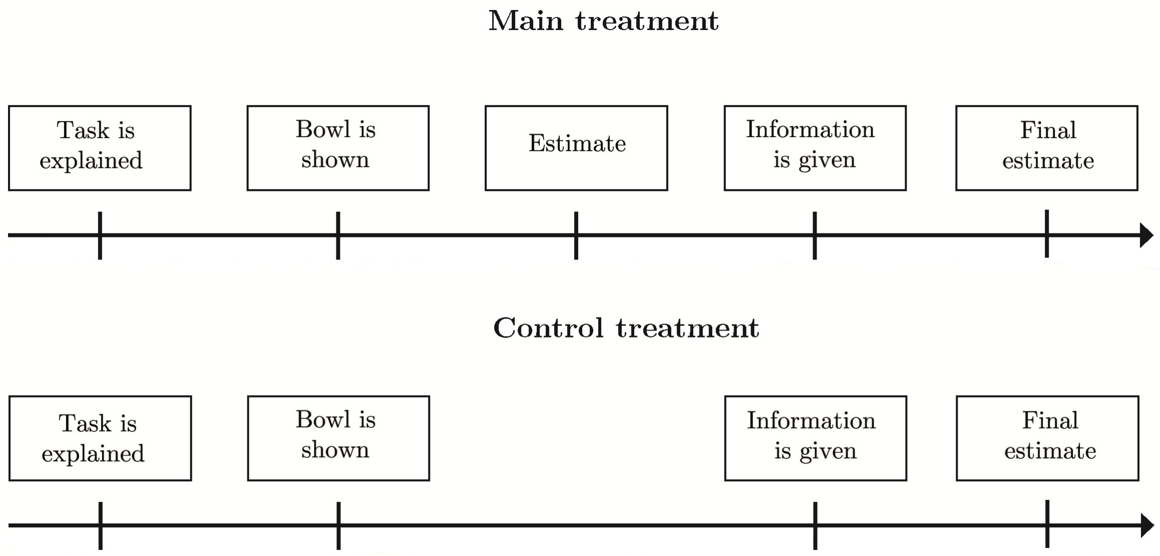


Figure 3.5: Timing of the experiment

ment we had conducted with 61 different subjects. They faced the same estimation task and were also paid according to the precision of their estimates. The average estimate of that group was 2615. In the results section we show that the additional information was in fact valuable to subjects.

Subjects in the **control treatment** also learned about the task and were asked to provide an estimate. The only difference between main and control treatment was that in the latter subjects did not state an estimate prior to receiving the additional information (see lower panel of *Figure 3.5*). In the control treatment, subjects saw the bowl with peas for some time prior to receiving the information sheets.[23] The time was approximated to be the same as what subjects in the main treatment needed. During this time subjects could form a belief about the correct number of peas, but did not state this to the experimenter, i.e., no commitment to a first estimate was made. After subjects received the information sheets, they stated their estimate on their answer sheet. Answer sheets were collected by the experimenter and the estimation task ended.

Procedural Details: A total of 105 subjects participated in eight sessions, 54 in the main and 51 in the control treatment. Subjects were mostly undergraduate students from various fields at the University of Bonn and were recruited using the online recruitment

[23]Note that in both treatments subjects did not know that they would receive helpful information before they actually received it.

system by Greiner (2003). No subject participated in more than one session. The experiment was conducted with paper and pencil. Sessions lasted on average about 40 minutes. Subjects earned on average 12.14 Euro, including a show-up fee of 5 Euro.

Hypotheses: Here we present the intuition of our model's prediction. A formal prediction is derived in Appendix C. In both treatments subjects see the bowl and form a belief about the correct number of peas. In the main treatment subjects commit to their belief by stating it towards the experimenter. Then, in both treatments subjects receive a public signal, i.e., valuable information. In the main treatment subjects necessarily reveal inconsistency if they respond to the public signal by changing their final estimate accordingly. On the contrary, subjects in the control treatment can respond to the public signal without revealing inconsistency due to the lack of commitment to their prior belief. Therefore the desire to behave consistently will make subjects partially neglect the valuable information in the main treatment. Consequently final estimates in the main treatment will be further away from the public signal than in the control treatment. Since subjects in the main treatment disregard valuable information, it follows directly that the quality of estimates and therefore earnings are lower in the main than in the control treatment.[24] We summarize our hypothesis as follows:

HYPOTHESIS EXPERIMENT 2: *The absolute difference between the final estimate and the information value of 2615 should be higher in the main treatment, compared to the control treatment. Final estimates in the main treatment will be further away from the correct value of 3000 compared to the control treatment.*

Results: Pooling data from both treatments the average (final) estimate was 2411.44. The estimation problem is very difficult and final answers ranged from 400 to 6000. Accordingly, the variance was large as indicated by a standard deviation of 1026.81. We chose a difficult task on purpose as it offers an ideal context to provide subjects with helpful information. To show that the public signal was in fact valuable we simply count the number of subjects in the main treatment whose estimate in the first estimation was further from 3000 than 2615. It turns out that this holds for 49 out of 54 subjects. This

[24] We abstract here from self-signaling motives. Subjects could signal skills towards themselves by being consistent with their private belief. Since this motive is present in both treatments, it does not change our predictions. One might even argue that self-signaling should be stronger in the main treatment. There private beliefs might be more salient for self-evaluation because they were actually stated.

means that about 91 percent of subjects could improve their (first) estimate by simply choosing 2615 or by moving in this direction.

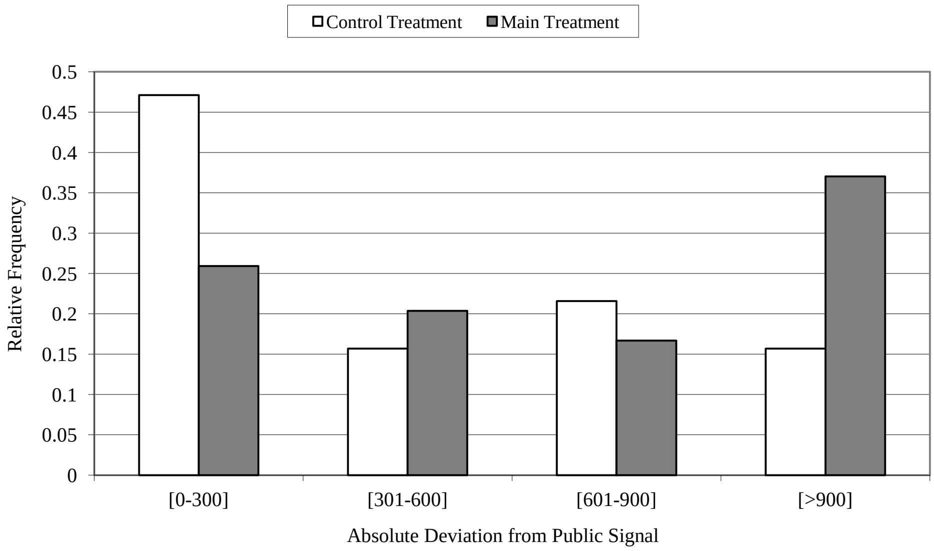


Figure 3.6: Relative frequency of absolute deviations from the public signal (2615) for both treatments.

We now turn to our main variable of interest, the absolute deviation between final estimate and the public signal of 2615. *Figure 3.6* shows a histogram of absolute deviations from the public signal. In the control treatment, about 47 percent of all estimates are in the interval +/-300 around the public signal. In contrast only about 26 percent of all estimates in the main treatment lie within this interval. The figure also shows that extreme deviations from the public signal are more frequent in the main treatment than in the control treatment. On average, the deviation in the main treatment is 341.68 points higher than in the control treatment. The difference in deviations from the public signal is statistically significant (p-value = 0.03, using Wilcoxon rank-sum test (two-sided) or OLS, regressing the absolute difference on a constant and a treatment dummy (p-value = 0.02)).

Figure 3.7 suggests that early commitment in the main treatment affects subjects' final estimate. The figure depicts a scatterplot with subjects' first and final estimates together with a line indicating the public signal 2615. This reveals that many subjects are either at or close to the 45-degree line indicating a strong resistance to take into account new and valuable information. It also shows that if subjects change, they change in the direction of 2615, as predicted by the model. The correlation between first and final estimate is 0.61 (p-value < 0.001).

The disregard of the valuable public signal is associated with a decrease in the quality of estimates and earnings. On average, estimates in the main treatment are 352.31 points further away from the correct value than estimates in the control treatment. The effect is statistically significant (p-value = 0.04, using Wilcoxon rank-sum test (two-sided) or OLS, regressing the deviation from the true value on a constant and a treatment dummy (p-value = 0.02)). We summarize our results as follows:

RESULT EXPERIMENT 2: *Commitment is key: it induces subjects to act consistently at the cost of neglecting valuable information and receiving lower payoffs.*

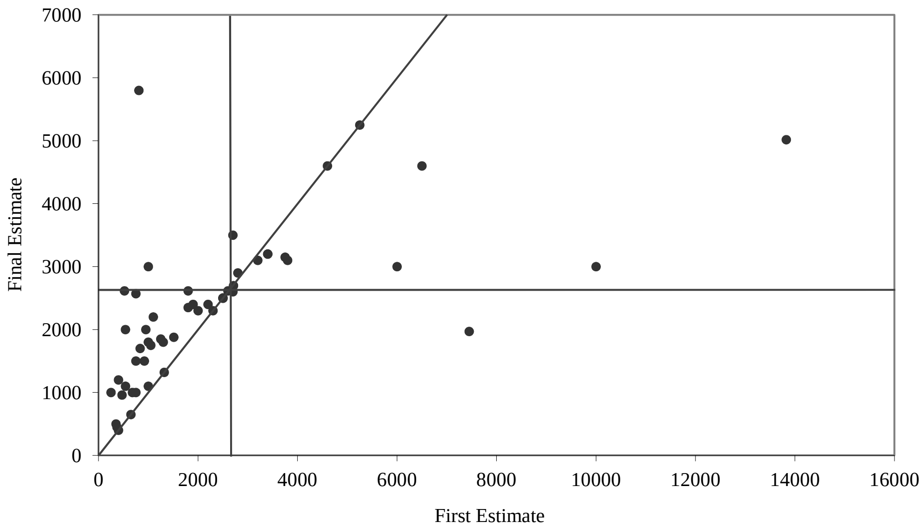


Figure 3.7: Scatterplot of first and final estimates in main treatment.

Discussion: Our findings are distinct from and cannot be explained by a simple confirmatory bias. Confirmatory bias describes a tendency to interpret new information as being in favor of or confirming the existing hypothesis.[25] In our study, however, subjects in both treatments had time and incentives to form an own hypothesis prior to receiving the public signal. Thus, if a confirmatory bias exists, it should be present in both treatments. To make this point even clearer, we ran an additional control treatment. In this control treatment, we asked subjects to indicate towards the experimenter once they had found an estimate by raising their hand. After that, subjects received the public signal and had time to revise their estimate. Thus, we can be sure that all subjects in this treatment have built an own hypothesis prior to receiving the public signal. However, subjects did

[25]Rabin and Schrag (1999) model confirmatory bias and show how it can lead to overconfidence. They also provide a review on the psychological literature on confirmatory bias.

not inform the experimenter about their first estimate, i.e., no commitment was made. A total of 47 subjects participated in this treatment. We find that our main result continues to hold. On average, the deviation between the final estimate and the public signal in the main treatment is 324.01 points higher than in the additional control treatment. The difference in deviations from the public signal is statistically significant (p-value = 0.06, using a Wilcoxon rank-sum test (two-sided) or simple OLS regression, regressing the absolute difference on a constant and a treatment dummy (p-value = 0.03)).

Note that our findings can also not be explained by subjects putting higher effort into forming a private belief in the main treatment. While given our experimental setup we believe it is plausible to assume that the quality of private signals is the same across treatments, one might argue that for some reason subjects in the main treatment try harder and thus receive better private signals. Consequently, subjects in the main treatment should also put higher weight on their private signals relative to the public signal in the Bayesian updating process. However, this does not necessarily explain that the absolute difference between final estimate and public signal is higher in the main treatment. The reason is that higher effort should make first estimates "better" and thus they should be closer to the valuable public signal to begin with. This would reduce the deviation between final estimates and the public signal. Also, higher estimation effort would predict a higher quality of final estimates in the main treatment compared to the control treatment, which is exactly the opposite of what we find.

3.5 Survey Manipulation

In this part, we examine how a desire for consistency can influence response behavior in surveys. The way consistency can affect survey responses is simple. The basic mechanism is to "tempt" a somebody to give a biased response in a first question. In a second step, she faces a question related to that response and the pressure to answer the survey consistently. Thereby, the addition of one (or several) question(s) can influence response behavior in subsequent questions. We test this intuition in three controlled survey studies. All surveys consisted of one main question of interest and one (or several) additional question(s). We randomly varied whether the additional question(s) was (were) included in the survey or not. The three surveys were designed to cover important topics. The first was about subject's attitudes towards a political/moral question (the punishment of a murderer).

The second survey asked for the degree to which participants followed social norms (in the context of lying behavior). The third survey was about general well-being.[26]

In the literature on survey methodology, effects of asking several related question and question ordering are known as context effects (see Tourangeau (1999)). Two main effects are distinguished, contrast effects and assimilation effects. Contrast effects describe negative correlations between previously asked and current questions while assimilation effects describe positive correlations (see for example Schwarz et al. (1991)). The main method to analyze these effects is known as split-ballot experiments (see for example Groves et al. (2009)). Our work contributes to this literature by providing a simple formal argument as well as experimental evidence on how a desire for consistency affects survey responses when several related questions are asked in a systematic way. In particular our findings underscore that the direction of consistency effects is predictable and show how they can be used to manipulate response behavior in surveys. In addition, our findings show that effects of inclusion of additional questions will be particularly severe, if image concerns of respondents are high.

3.5.1 Design

We wanted to examine how the inclusion of related additional questions can affect response behavior in a survey. We designed three survey studies where we selected three important topics: political/moral attitude, norm compliance and general well-being.

The basic design of all three survey studies was very simple. We always study two conditions which were randomly varied between subjects. In the control condition, subjects only had to answer our main question of interest. In the manipulation condition, subjects had to answer one or several related questions before they had to answer our main question of interest. The questions were as follows:

Survey 1 - Punishment of a murderer: In survey 1, our main question of interest was if subjects would agree that a murderer should be imprisoned for the rest of his life. Subjects were asked to read a short text that described a horrible deed of a murderer. After reading the text, subjects were asked to answer the following question: "Do you agree with the following statement? I would approve if the offender would be sent to

[26]For overviews on research about life satisfaction, see for example Frey and Stutzer (2001) or Layard (2005).

prison for the rest of his life, never to be released." Subjects could either agree or disagree with the statement by checking the appropriate box. In the manipulation treatment, subjects had to answer a different but related question first, namely: "Do you agree with the following statement? Everybody deserves a second chance in life. Even dangerous criminals should be released after their imprisonment and be given a chance to start a new life." Subjects could either agree or disagree.

Survey 2 - Lying behavior: In survey 2, we were interested in lying behavior. We described to subjects a lying experiment similar to Fischbacher and Heusi (2011). Subjects were told to imagine that they participated in an experiment where they were paid according to the role of a die. Subjects would role a die in private and were then paid according to the number they reported to the experimenter. A reported "1" would earn them 2 Euro, a "2" would earn them 4 Euro and so on. Maximum earnings would be 12 if a "6" is reported. Thus it was clear to our subjects that in the experiment there were incentives to lie and report a high number. In the main question of this survey we asked subjects hypothetically which number they would report if they participated in such an experiment, conditional on the actual die role. Thus we asked subjects which number they would report if they rolled a "1", what they would report if they rolled a "2" etc. In the manipulation treatment we asked two additional questions before subjects were explained the hypothetical lying experiment. Subjects were asked to which degree they would agree with the following statements: (1) "I am an honest person"; (2) "Other people can rely on my words". Subjects could respond on a scale from 1 ("do not agree") to 5 ("completely agree").

Survey 3 - General well-being: In survey 3 we asked subjects about their general well-being. Our main question of interest was: " How satisfied are you at present with your life, all things considered?". Subjects could respond on a scale from 0 to 10. In the manipulation condition we asked several questions before the question on well-being. Questions were about health condition, satisfaction with field of study, number of friends and optimism about finding a good job after leaving university.[27]

Procedures: Subjects were mostly students at the University of Bonn and were recruited using the software by Greiner (2003). All surveys were conducted using paper and pencil in the BonnEconLab at the end of different and unrelated experiments.

[27]In the control condition we actually asked the additional questions as well, but after subjects had answered the question on general well-being.

3.5.2 Hypothesis

The inclusion of additional questions in the manipulation conditions should trigger subjects' consistency concerns when answering the respective main question of interest. If responses to the additional questions are "biased" in a systematic way, the bias should extend to the main question due to a desire to act consistently. Consequently, responses to the main question of interest should systematically differ between manipulation and control question in all our studies.

To illustrate this in more detail, consider survey 1. There we expected that most subjects in the control group would agree with the statement that the murderer should be imprisoned forever. We also expected that many subjects in the manipulation group would agree with the statement that everybody deserves a second chance. Therefore our hypothesis is that subjects in the manipulation group would feel a desire to be consistent with their first response and therefore agree less frequently with the statement that the murderer should be imprisoned forever, as compared to the control treatment.

HYPOTHESIS: *In all three surveys, responses to the main question of interest differ between the manipulation and the control condition.*

3.5.3 Results

Survey 1: A total of 95 subjects participated in survey 1. Figure 1 summarizes our results. We first analyze how many subjects in the control treatment agreed that the murderer should be sent to prison for the rest of his life. Given the horrible deed of the murderer, we expected that most subjects would agree. It turns out that 44 out of 48 subjects (91.7 percent) responded with "I agree". Now consider how many subjects in the manipulation treatment agreed with the statement that everybody deserves a second chance in life. Here we expected that many subjects would agree to give a second chance. This it what we find: 26 out of 47 subjects (55.3 percent) responded with "I agree".

Given these results, the model predicts that in order to be consistent with the statement that everybody deserves a second chance, fewer subjects in the manipulation treatment agree to imprison the murderer forever. This is confirmed by our data. Only 32 out of 47 subjects (68.0 percent) stated that they would approve if the offender would never live in freedom again. Thus the approval rate dropped from 91.7 percent in the

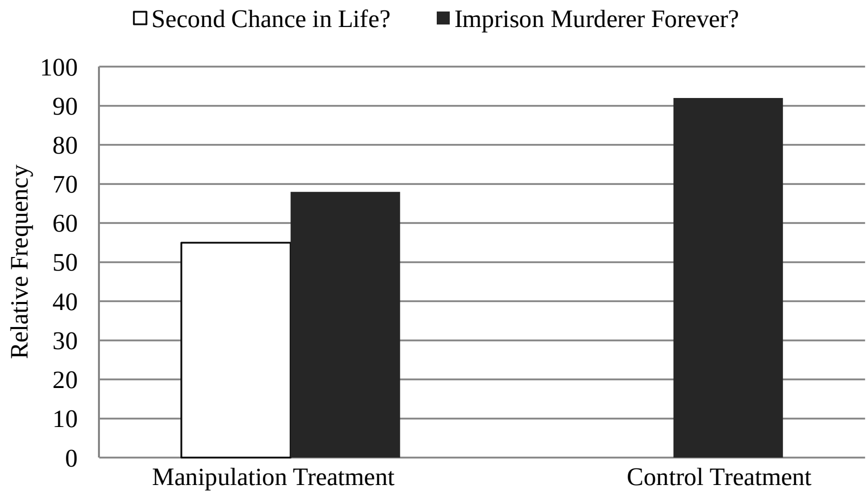


Figure 3.8: Relative frequencies of statement "Yes" for question on punishment of murderer for manipulation and control treatment plus relative frequency of statement "Yes" for question on second chance in life (only manipulation treatment).

control treatment to 68.0 percent in the manipulation treatment. This difference is statistically significant using either a Fisher exact test (p-value < 0.005) or simple Probit regression, regressing a dummy variable for "agree" or "disagree" on a constant and a treatment dummy (p-value < 0.005). Thus, we were able to significantly manipulate reported attitudes towards the punishment of a murderer, simply by including an additional question.

Survey 2: 68 subjects participated in survey 2. First, consider responses to the two questions on honesty we asked in the manipulation treatment. We expected that a very high number of subjects would state to be honest. This is indeed what we find. 86 percent of subjects either agreed or completely agreed with the statement "I am an honest person". For the statement "Other people can rely on my words" all subjects either agreed or completely agreed.

Given this very high degree of self-reported honesty, we expected that subjects would lie less in the hypothetical lying experiment compared to the control treatment, in order to be consistent with prior statements of being honest. This is confirmed by our data. When we use the sum of all reported dice rolls (six reports, one for each possible die roll) as our measure of hypothetical lying behavior, we find that subjects in the manipulation treatment report a sum of 24.43 on average. Reported dice rolls in the control treatment are on average 3.07 points higher. This difference is significant using a Ranksum test (p-value < 0.01) or OLS, regressing reported dice rolls on a constant and a treatment dummy (p-value for treatment dummy $= 0.02$).

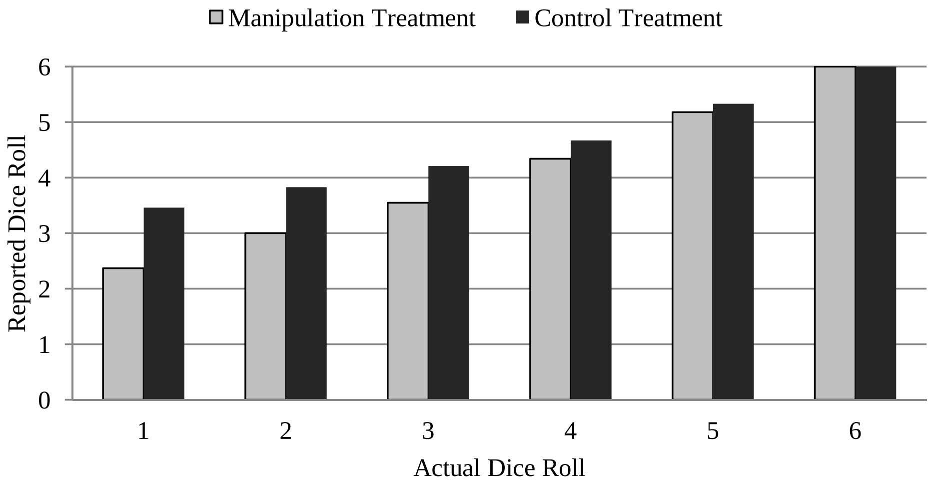


Figure 3.9: Reported and actual dice roll for manipulation and control treatment.

Figure 2 provides a more detailed description of self-reported lying behavior. It displays the average reported die roll conditional on the actual die roll for both treatments. We see that, at all levels, reports in the manipulation treatment are below those in the control treatment. For example, if the actual die roll was a one, subjects in the control treatment on average indicate a report of almost 3.5. The corresponding average report in the manipulation treatment is more than one point below.[28] The difference between the two treatments gets smaller for higher actual dice rolls, as the scope for lying gets smaller. (There is no point in lying and reporting a number below 6 if the actual die roll is 6.)

Survey 3: A total of 180 subjects participated in survey 3. If we compare stated subjective well-being between the treatments, we find that well-being is .48 points higher in the manipulation treatment (when questions about health, number of friends etc. were asked before). This difference is significant using a Ranksum test (p-value < 0.05).

Table 1 shows results from OLS regressions. In regression 1, we regress subjective well-being on a constant and a treatment dummy. The treatment effect is only marginally significant (p-value < 0.1). A possible concern could be that subjects in our two conditions actually differ for example in health condition or number of friends, i.e., randomization

[28]Note that Figure 2 also reveals that (hypothetical) lying behavior does not vanish in the manipulation treatment. Average reported die rolls in the manipulation treatment are always above the actual die roll.

between conditions was not fully successful. This could possibly explain the treatment difference we find. To address this issue, we exploit that we also have data on current health condition etc. for subjects in our control condition.[29] Our finding is robust when we control for these additional factors (see regressions 2 and 3). In fact, the treatment dummy in regressions 2 and 3 becomes highly significant.[30] Apart from number of friends, all the other controls are significant. Favorable current health condition, satisfaction with field of study and optimism to find a good job in the future significantly increase stated subjective well-being.

3.6 Concluding Remarks

We have presented a model that conceptualizes a signaling motive for consistency and allows the analysis of how it affects economic behavior. In the model, people behave consistently, because this allows the signaling of skills. Our first experiment then tested the basic logic of our model. In the context of an estimation task, we show that consistent behavior is a signal of ability and that principals understand and value this signal. Agents anticipate this and make more consistent estimates compared to a control treatment without principals, i.e., in a situation where signaling ability has no strategic value. We then presented results from an experiment that tested a central implications of our model. We show that explicit commitment has a crucial impact on subsequent behavior as commitment is a prerequisite for detecting inconsistencies. In the experiment such a commitment leads to a neglect of valuable information. Furthermore, we showed how a taste for consistency can systematically influence response behavior in surveys. In three different surveys, we demonstrated that the inclusion of one or several questions affects answers to subsequent related questions. Our results underscore the importance of a taste for consistency as a means of influence and highlight the fragility of response behavior.

In our model we highlight the role of signaling skills as a key driver of consistent behavior. Other motives that have been discussed in the literature are reduction of cognitive

[29] Remember that we also asked subjects in the control treatment about things like their current health condition, but only after the question on subjective well-being.

[30] We also conduct an ordered probit regression, regressing subjective well-being on a constant, a treatment dummy and several controls (health condition, satisfaction with field of study, optimism to find a good job in the future and number of friends). Our findings are robust to this change in specification. The p-value of the treatment dummy is 0.011.

Dependent Variable = Subjective Well-Being			
	(1)	(2)	(3)
Treatment dummy	0.479*	0.59***	0.54**
	(0.29)	(0.22)	(0.24)
Health Condition		0.25***	0.24***
		(0.05)	(0.05)
Studies		0.33***	0.33***
		(0.06)	(0.06)
Optimism		0.23***	0.24***
		(0.06)	(0.07)
Number of friends			0.003
			(0.009)
Constant	7.071***	1.25**	1.09*
	(0.20)	(0.57)	(0.61)
N	180	179	164
Adj R-squared	0.01	0.40	0.41

***p-value<0.01, **p-value<0.05,*p-value<0.10

Table 3.1: OLS, regressing subjective well-being on a constant, a treatment dummy (=1 if manipulation treatment) and several controls (health condition, satisfaction with field of study, optimism to find a good job in the future and number of friends).

dissonance and costs of thinking (see work cited in Cialdini, 1984). The most important difference between a signaling argument and fixed costs or cognitive dissonance concerns the role of third parties. We believe that the presence of observers creates a particularly strong desire to appear consistent. Moreover, our first experiment demonstrates that cognitive dissonance or cost of thinking alone are not sufficient to understand consistent behavior. While we would argue that these two motives play a role as well, they cannot explain observed treatment differences. This follows simply from the fact that both motives can be relevant in both treatments. It is also unclear how cognitive dissonance theory or costs of thinking can explain results from our second experiment. If, as we assume, subjects form a belief about the correct estimate in both treatments, the desire to avoid cognitive dissonance (by not following the public signal) should not differ between the treatments. Also, from a cost of thinking perspective there should be no treatment difference in final estimates.

We conclude by highlighting several important economic implications. The desire for behaving consistently makes people act against their immediate material interest. Early statements or choices have consequences for subsequent behavior. For instance we show that early statements on a matter can make people ignore valuable information. This is of interest for the design of committee or jury procedures. Institutionally requested commitments on a certain opinion or intention at an intermediate stage can decrease the quality of final choices as these potentially do not reflect the full level of available information. In the context of negotiations where full details of the negotiation are only sequentially revealed, statements of requests at an early stage of negotiations can increase the likelihood of negotiation failures. In the final stage of negotiations, when all information is revealed, negotiation outcomes below these requests cannot be reached without one party revealing inconsistency with early statements. This may be one reason why early requests are often formulated rather vaguely as this makes possible inconsistencies between requests and negotiation outcomes harder to detect.

On the other hand, in many bargaining situations early requests can be used strategically to increase bargaining power. Take a simple ultimatum game situation where the responder can state a minimum acceptable offer towards the proposer in the beginning. In principle this statement is cheap-talk and should not affect the outcome of the game. In the presence of a desire for consistency, however, the stated minimum acceptable offer serves as a credible commitment that the responder will reject any offer below that because otherwise he would reveal inconsistency. Anticipating this, the proposer may have to offer something close to the demanded minimum. Note that the credibility of the minimum acceptable offer depends on the reputational costs from inconsistency. Consequently, early requests made in public, e.g., in front of large audience have a higher credibility than requests that are only stated towards the other bargaining party.

The desire to be consistent with stated intentions can also be a powerful means to circumvent problems of self-control. Consider smoking: A public announcement "I will stop smoking!" creates a pressure to live up to that announcement. Continuing to smoke is only possible at the cost of revealing some inconsistency, thereby signaling a low type. Thus, explicit public commitments can be very effective in solving problems of self-control. In fact, one reason for the effectiveness of self-help groups could be that they "force" members to publicly announce their intentions.

Our results are also informative from a methodological point of view. When designing surveys and experiments, our results suggest interdependencies between behavior in related experiments conducted in one session, between behavior in repeated interactions or within-subjects designs, between experimental behavior and related questionnaire answers or between different survey answers. Depending on the context of the experiment or the survey study, the desire for consistency introduces a bias in behavior that potentially confounds results.

Chapter 4

A Consistency-Based Approach to Anchoring Effects

4.1 Introduction

Studies in psychology and economics have shown that subjective judgments and valuations can be influenced using simple anchoring manipulations. In a seminal study, Tversky and Kahneman (1974) generated a random number between 0 and 100 and asked subjects if the number of African nations in the UN was greater than that number or not. Then they asked subjects to estimate the number of African nations in the UN and show that the randomly generated number (the anchor) significantly affected estimates. Their results have been shown to be robust to incentivizing the estimation tasks and have been replicated many times in various contexts.[1] The susceptibility of judgments towards completely irrelevant randomly generated numbers is hard to reconcile with assumptions of rationality frequently made to describe economic behavior.

Even more troubling for economic theory are studies showing that anchoring manipulations also affect valuations for common consumption goods. Ariely, Loewenstein and Prelec (2003) elicit subjects' willingness to pay for different consumer products. Before that, however, they let subjects write down the last two digits of their social security number and ask if they are willing to buy the products for that price. They report that subjects' willingness to pay is affected by the last two digits of the social security number.

[1]See for example Jacowitz and Kahneman (1995), Strack and Mussweiler (1997), Chapman and Johnson (1999), and Epley and Gilovitch (2001).

Subjects with rather high last two digits of the social security number on average report a higher willingness to pay compared to subjects with rather low last two digits. [2] These results suggest that people do not have stable preference relations and that valuations of quite common consumption goods are to a large extend arbitrary. Arbitrariness of choices is particularly problematic as economic analysis, e.g., evaluation of consequences of tax changes, effects of wage cuts, behavior in social dilemmas, builds on the assumption that individuals act optimally in their choice context given their preference relation.

In this chapter we attempt to reconcile the evidence from anchoring manipulations in a model with rational decision-makers who have stable preferences or judgments. The model is built on two central assumptions. First, (some) decision-makers make random and unsystematic errors in their choices from time to time. Second, decision-makers have a desire to be consistent in their choices. In the model, there exists an outcome μ that is optimal in monetary terms and decision-makers face two choices. First, they make a binary choice whether μ is above or below a randomly determined number, the anchor. Second, they are asked to state their μ. We assume that decision-makers make unsystematic errors in their choices from time to time. These errors are modeled as random deviations from the utility maximizing choice. Considering the two choices, these random mistakes can lead to inconsistent choices for some decision-makers (first choices stating that μ is above (below) the anchor followed by second choices stating a μ below (above)). We also assume that decision-makers have a desire to be consistent over time. This creates a trade-off in the second choice between maximizing standard monetary utility and choosing consistently with the first choice. If the desire to be consistent is sufficiently large, decision-makers might be willing to sacrifice standard monetary utility in order not to behave inconsistently. Thereby the seemingly irrelevant random anchor influences decision-makers' second choices because it constitutes a "threshold value" that separates consistent from inconsistent choices. Low anchors drive choices down, high anchors drive choices up.

The assumption that people make mistakes is not new to economic theory. Trembling hand perfection is a refinement of Nash equilibrium that takes off-equilibrium play into account by assuming that people sometimes mistakenly choose wrong strategies (see Selten,

[2]Several other studies have shown anchoring effects in valuations, see for example Johnson and Schkade (1989), Kahneman and Knetsch (1993) and Green et al. (1998). Bergman et al. (2003) replicate the findings of (Ariely et al. (2003)) and show that subjects with low cognitive ability are more prone to anchoring effects. In a recent replication, Fudenberg et al. (forthcoming) find weaker anchoring effects.

1975). Likewise, Quantal Response Equilibrium is an equilibrium concept similar to Nash Equilibrium, assuming that players' behavior might be noisy, for example, due to bounded rationality. Players are assumed to be "... more likely to select better choices than worse choices, but do not necessarily succeed in selecting the very best choice" (Goeree, Holt and Palfrey, The New Palgrave Dictionary of Economics, Second Edition, 2008). In our set-up, the sources of errors could be manyfold. First, errors could be in the actual act of choosing. People might simply push the wrong button, mark the wrong alternative or enter a wrong number by mistake. Second, the source of mistakes could lie in the process of thinking about the best alternative. Cognitive resources are limited and consequently mistakes in determining the preferred choice are likely. Third, we want to stress that what we label mistakes in our set-up is compatible with an alternative assumption, that people have imperfect knowledge about the utility maximizing alternative. Instead of perfectly knowing their preferences, individuals might receive noisy (but unbiased!) signals. In that case actual observed behavior would be similar to that when people make unsystematic random mistakes.

The assumption that individuals have a taste for consistency builds on a large literature in social psychology. Both theoretical and empirical work stresses that the desire to behave consistently is an important determinant of human behavior. Cialdini (1984) notes that "once we make a choice or take a stand, we will encounter personal and interpersonal pressures to behave consistently with that commitment." He summarizes much of the evidence and discusses three main explanations for consistency preferences.[3] One is built on the notion of cognitive dissonance. According to this approach, consistent behavior reflects a desire to avoid cognitive dissonances. Early work in this direction was developed in Heider (1946), Newcomb (1953) and Festinger (1957). A different potential driver of consistent behavior is based on the idea that thinking is costly. If thinking does involve cognitive costs, it may in fact be optimal to stick to a particular behavioral strategy and not to change behavior in response to new information or new signals. A third explanation for a taste for consistency is that consistent behavior allows the signaling of positive traits. In chapter 3 we build a model where decision-makers have a taste for being consistent over time because this allows them to signal personal and intellectual strength. They test central implications of their model in a series of experiments and find empirical support.

[3]Emprical studies documenting a desire for consistency include Allgeier, Byrne, Brooks and Revnes (1979) and Asch (1956). Freedman and Fraser (1966), Cialdini et al. (1978) and Sherman (1980) analyze how the taste for consistency can be used to influence behavior.

In our model we directly assume a taste for consistency. In that sense our model can be viewed as a reduced form where the underlying cause of consistency preferences could be either of the three explanations mentioned.[4]

Note that our model captures judgments and valuations. The choices decision-makers face could be both, to estimate the height of the Mount Everest and to state their willingness to pay for an IPod. Note also that our model is compatible with recent empirical findings from Bergman et al. (2010). They show that anchoring is more pronounced for subjects with low cognitive ability. In our model, only decision makers that make mistakes are susceptible to anchoring manipulations while the choices of all other decision makers remain unaffected. As the frequency of making (or not making) mistakes is likely to be correlated with cognitive ability, our model is consistent with the observation that low cognitive ability implies higher effectiveness of anchoring manipulations.

We proceed as follows. In section 4.2 we first report in more detail the procedure used in most anchoring manipulations and then formally introduce our model. Section 4.3 concludes.

4.2 The Model

4.2.1 Anchoring Manipulations

We aim to explain anchoring manipulations in the context of judgments and valuations. These manipulations involve two choices (see for example Tversky and Kahneman (1974) or Ariely et al. (2003)). Choices could be regarding the valuation of a certain consumption good or judgments in the context of an estimation task. Subjects are first instructed about the choices they will face. In the first choice the anchor is presented and its random nature is made clear to subjects. For example a wheel of fortune is spun in front of subjects or subjects are asked to write down the last two digits of their social security number. Thus, the fact that the anchor is completely uninformative and unrelated to the choice problem is made very salient. Subjects are then asked to decide if their valuation or

[4] In the economics literature, Eyster (2002) and Yariv (2005) have suggested models of consistent behavior. Ellingsen and Johannesson (2004) and Vanberg (2008) refer to the taste for consistency as a possible reason for why people incur costs of lying and thus keep their promises. In chapter 3 we underscore the informativeness of consistency as signal of skills, and analyze implications for commitment and social influence.

judgment is higher or lower than the random anchor. For valuations, depending on their choice, subjects buy (or don't buy) the good at the price of the anchor. In the context of judgments, subjects win a prize if their first choice was correct. In the second choice, subjects are asked to state their valuation or judgment. For valuations this is usually done in an incentive-compatible way, using the Becker-DeGroot-Marschak procedure. In the context of judgments, subjects are simply paid according to the precision of their judgments. In some studies, choices were only hypothetical. In incentivized studies, at the end of the experiment one of the two choices was randomly chosen to be payoff-relevant to avoid hedging motives. The striking finding of these manipulations is that the anchors influence second choices in a systematic way. For low anchors, judgments or valuations are significantly below those when high anchors are presented to subjects.

4.2.2 The Model

In the model there are two periods $t = 1, 2$. In both periods, decision-makers D face a choice-problem. There exists an outcome μ which is an element of the real numbers, $\mu \in R$. In the context of valuations, μ represents the true valuation of the decision-maker, for example his willingness to pay for an IPod. In the context of judgments, μ is the correct judgment, for example the true height of the Mount Everest.

In period 1 decision-makers are asked whether their μ is above a random number r_a, the anchor, i.e., $x_1 \in \{Yes, No\}$. Decision-makers are paid a fixed amount p_f if they answer correctly. In the context of valuations, p_f reflects the gains from buying the good at the anchor price r_a if $\mu \geq r_a$ or the gains from not buying if $\mu < r_a$ respectively. In the context of judgments, p_f is the prize for answering correctly. In period 2, Ds face a second choice x_2. They are asked to state their μ. Thus, x_2 is an element of the real numbers ($x_2 \in R$) and decision-makers are paid according to the distance between x_2 and μ. For simplicity we assume a linear incentive scheme and say that in the second choice Ds are paid according to $p - |x_2 - \mu|$. This captures payment according to the quality of judgments in that context or incentive compatible preference elicitation in the valuation context.

D's utility depends on two parts, a monetary component and a taste for consistency. For simplicity we assume that money enters linearly in the utility function and that the two components are additively separable. The taste for consistency is captured as a function $\beta()$ that can take on two values, β_H and β_L with $\beta_H > \beta_L$. The function has three arguments, the two choices x_1, x_2 and the value of the anchor r_a. The function takes on the value β_H if x_1 and x_2 are consistent and β_L otherwise. Choices are consistent if D chooses $x_1 = Yes$ and $x_2 \geq r_a$ or $x_1 = No$ and $x_2 \leq r_a$. In sum, the taste for consistency is described as follows:

$$\beta(x_1, x_2, r_a) = \begin{cases} \beta_H & \text{if x}_1 = Yes \text{ and } x_2 \geq r_a \text{ or } x_1 = No \text{ and } x_2 \leq r_a \\ \beta_L & \text{otherwise.} \end{cases}$$

It immediately follows that consistency is not an issue in the first choice. Naturally, a taste for consistency can only affect choices in the presence of a choice history.

A final assumption completes the model. Decision-makers are assumed to make mistakes from time to time. These mistakes are modeled as random and unsystematic deviations from the utility maximizing choice. In period 1, decision-makers' choice is binary. p_e^1 indicates the probability that decision-makers make an error and do not pick the utility maximizing choice. Naturally we assume $p_e^1 < \frac{1}{2}$. In period 2, Ds choose from the real numbers. $p_e^2 < \frac{1}{2}$ denotes the probability of mistakes in period 2. In case D makes a mistake in period 2, this is captured by a normally distributed error term ϵ with zero mean and variance σ_r^2. Thus actual choices in case of mistakes in period 2 are the sum of the utility maximizing choice plus the error term. If no mistake is made x_2, is simply the utility maximizing choice. The nature of mistakes could be on the behavioral level, meaning that decision makers make mistakes in the actual act of choosing. It could also be cognitive. Decision makers might make mistakes when thinking about the choice problem and determining the best alternative. Note that our notion of mistakes is also compatible with incomplete knowledge about the best alternative. Decision-makers might only receive noisy (but unbiased) signals about their utility-maximizing alternative and consequently sometimes pick alternatives that are suboptimal ex-post.

Actual behavior in the first period is straightforward. In case D does not make an error, he will choose $x_1 = Yes$ if $\mu > r_a$ and $x_1 = No$ if $\mu < r_a$. In case he makes a mistake, i.e. with probability p_e^1, he chooses $x_1 = No$ if $\mu > r_a$ and $x_1 = Yes$ if $\mu < r_a$.

Choice 2 is more involved. We first abstract from errors in the second choice and focus on the utility maximizing decision. In period 2 decision makers face a trade-off between maximizing monetary payoffs and behaving consistently with their period 1 choice. They face the following problem:

$$\max_{x_2} U_2 = p - |x_2 - \mu| + \beta(x_1, x_2, r_a).$$

Obviously, the choice is easy for decision-makers that did not make a mistake in period 1, i.e., with $x_1 = No$ and $\mu \leq r_a$ or $x_1 = Yes$ and $\mu \geq r_a$. They simply choose $x_2 = \mu$, thereby maximizing monetary utility and assuring consistency. The interesting case is the one where decision-makers made errors by choosing $x_1 = No$ ($x_1 = Yes$) although $\mu > r_a$ ($\mu < r_a$). Now they face a trade off. Either they choose to maximize monetary payoffs (thereby being inconsistent) and receive $U_2 = p + \beta_L$. Or they choose consistently (thereby sacrificing monetary payoffs) and receive $U_2 = p - |r_a - \mu| + \beta_H$.[5] Thus, they will behave consistently if the gains from being consistent outweigh the monetary losses, i.e., if $\beta_H - \beta_L \geq |r_a - \mu|$. Now we can incorporate mistakes in period 2. Remember that with likelihood p_e^2 D's make mistakes in period 2, modeled as a normally distributed error term ϵ (with variance σ_r^2) that adds to the utility maximizing choice. Thus choices of decision-makers can be summarized as in the following Proposition.

PROPOSITION 1: If the taste for consistency is sufficiently large, i.e., $\beta_H - \beta_L \geq |r_a - \mu|$, decision-makers with $x_1 = No$ ($x_1 = Yes$) and $\mu > r_a$ ($\mu < r_a$) maximize utility by choosing x_2 such that it is consistent with their first choice, i.e., $x_2 = r_a$. Taking mistakes into account, actual choices of these decision-makers are described by $x_2 = r_a + \epsilon$ in case of mistakes and $x_2 = r_a$ otherwise. Choices of all other decision-makers are described by $x_2 = \mu + \epsilon$ in case of mistakes and $x_2 = \mu$ otherwise.

Proposition 1 shows how a random anchor (r_a) influences decision-makers choices. Intuitively, the anchor works because it constitutes a threshold value that separates consistent from inconsistent choices. In the next step we examine how anchoring-manipulations, i.e., variations in the value (high versus low anchor) systematically affect choices.

We compare two situations, one with a low anchor r_a^l and one with a high anchor r_a^h, i.e., $r_a^l < r_a^h$. Let $x_2(r_a)$ denote the period 2 choice given anchor r_a. First, consider the low anchor r_a^l. From Proposition 1 we know that (if the taste for consistency is sufficiently strong) decision makers that made a mistake in period 1 make choices according to

[5] The optimal way of behaving consistently in that case is to choose $x_2 = r_a$.

$x_2(r_a^l) = r_a^l + \epsilon$ or $x_2(r_a^l) = r_a^l$. Those who made no mistake in period 1 are not affected by the anchor and consequently their choices are described by $x_2(r_a^l) = \mu + \epsilon$ or $x_2(r_a^l) = \mu$.

Now consider the high anchor r_a^h. Choices of decision-makers who made a mistake in period 1 are described by $x_2(r_a^h) = r_a^h + \epsilon$ or $x_2(r_a^l) = r_a^h$, those of all others by $x_2(r_a^h) = \mu + \epsilon$ or $x_2(r_a^h) = \mu$. Thus, one can easily see that the random anchors affect choices in a systematic way. In expectation, period 2 choices with a low anchor are below period 2 choices with a high anchor, i.e., $E(x_2(r_a^l)) < E(x_2(r_a^h))$. This is summarized in the following Lemma.

LEMMA 1: Compare choices for two anchors r_a^l and r_a^h with $r_a^l < r_a^h$. If the taste for consistency is sufficiently large, we have that $E(x_2(r_a^l)) < E(x_2(r_a^h))$.

4.3 Concluding Remarks

Striking findings from anchoring manipulations have called into question utility-maximizing behavior of individuals and have suggested that human behavior are driven rather by arbitrariness than some stable preference relation. In this chapter we presented a model of utility-maximization that can explain evidence from anchoring experiments. The model makes two central assumptions. First, people make random, unsystematic mistakes from time to time. Second, people have a desire to be consistent in their choices. These two assumptions are enough to explain effectiveness of anchoring manipulations. Intuitively, random anchors work because they constitute a threshold between consistent and inconsistent choices. Thus, our model shows that results from anchoring manipulations are not necessarily evidence for arbitrariness in choices, but can be reconciled with a simple model of utility-maximization.

Our model is also compatible with findings that sensitivity to anchoring manipulations is correlated with cognitive ability. Bergman et al. (2003) show experimental evidence that subjects with low cognitive ability are more affected by anchoring manipulations compared to subjects with high cognitive ability. In our model only subjects that make mistakes are affected by anchors. As cognitive ability is likely to be predictive for the frequency of mistakes, the correlation between cognitive ability and effectiveness of random anchors is consistent with our model. Note also that our model provides some insights regarding the limits to the effectiveness of anchoring manipulations. Although this is not modeled explicitly in our set-up, the frequency of mistakes is likely to depend on the placement of the anchor. Intuitively, if the anchor is placed very far away from μ, the first choice of decision makers is very simple and consequently mistakes seem less likely. Consequently anchors are less likely to be effective. There exists, however, a simple trade-off as the size of anchoring effects (in cases when the anchor is effective) also depends on the placement of the anchor. The larger the distance between anchor and μ, the larger is the effect of the anchor on choices in period 2. Effectiveness of anchoring manipulations is also likely to depend on the difficulty of the choices people face. Intuitively, the frequency of mistakes is higher, the more difficult the choice problem is. In contexts where mistakes are unlikely, we predict that the effectiveness of anchoring manipulations is low.

Finally, we want to stress that the purpose of our model is to highlight an intuition why evidence on anchoring manipulations is compatible with a model of rational decision-making. This suggests, that some of the conclusions drawn from studies on anchoring manipulations, namely the arbitrariness of preferences and judgments might be too far-reaching. The purpose is not to fit experimental data from anchoring manipulations. Some of the assumptions made in the model are too strong and consequently our model will go wrong there. For example, we assume that high types perfectly know and act according to their preferred outcome. Thus we predict that a (non-trivial) fraction of decision makers will manage to pick exactly the correct choice. Especially in the context of judgments this seems unlikely.

References

Abeler, Johannes, Falk, Armin, Goette, Lorenz and David Huffman. 2011. "Reference Points and Effort Provision." *American Economic Review*, 101 (2): 470-492.

Ahlbrecht, Martin and Martin Weber. 1997. "Preference for Gradual Resolution of Uncertainty." *Theory and Decision*, 43: 167-185.

Akerlof, George and William Dickens. 1982. "The Economic Consequences of Cognitive Dissonance." *American Economic Review*, 72(3): 307-319.

Allgeier, A.R., Donn Byrne, Barbara Brooks and Diane Revnes. 1979. "The Waffle Phenomenon: Negative Evaluations of Those Who Shift Attitudinally." *Journal of Applied Social Psychology*, 9(2): 170-182.

Andreoni, James and Ragan Petrie. 2004, "Public Goods Experiments Without Confidentiality: A Glimpse Into Fundraising." *Journal of Public Economics*, 88(7-8): 1605-1623.

Andreoni, James and Douglas Bernheim. 2009. "Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects." *Econometrica*, 7(5): 1607-1636.

Ariely, Dan, Anet Bracha and Stephan Meier. 2009. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially," *American Economic Review*, 99(1): 544-555.

Ariely, Dan, George Loewenstein, and Drazen Prelec. 2003. "Coherent Arbitrariness: Stable Demand Curves without Stable Preferences." *The Quarterly Journal of Economics*, 118 (1): 73- 104.

Asch, Solomon. 1956. "Studies of Independence and Conformity: A Minority of One Against a Unanimous Majority." *Psychological Monographs*, 70 (Whole no. 416).

Bell, David E. 1985. "Disappointment in Decision Making under Uncertainty." *Operations Research*, 33 (1): 1-27.

Bellemare, Charles, Michaela Krause, Sabine Kroeger and Chendi Zhang.

2005. "Myopic Loss Aversion: Information Feedback vs. Investment Flexibility." *Economics Letters*, 87 (3): 319-324.

Bem, Daryl 1972. "Self-perception Theory," *Advances in Experimental Social Psychology*, 6: 1-62.

Bénabou, Roland and Jean Tirole. 2002. "Self-Confidence and Personal Motivation." *Quarterly Journal of Economics*, 117(3): 871-915.

Bénabou, Roland and Jean Tirole. 2004. "Willpower and Personal Rules." *Journal of Political Economy*, 112(4): 848-886.

Bénabou, Roland and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review*, 96(5): 1652-1678.

Benartzi, Shlomo and Richard Thaler. 1995. "Myopic Loss Aversion and the Equity Premium Puzzle." *Quarterly Journal of Economics*, 110 (1): 73-92.

Benoit, Jean-Pierre and Juan Dubra. (forthcoming). "Apparent Overconfidence?" *Econometrica.*

Bergman, Oscar, Tore Ellingsen, Magnus Johannesson and Cicek Svensson. 2010. "Anchoring and Cognitive Ability." *Economics Letters*, 107(1): 66-68.

Bernheim, Douglas. 1994. "A Theory of Conformity.", *Journal of Political Economy*, 102(5): 841-877.

Bowman, David, Deborah Minehart and Matthew Rabin. 1999. "Loss Aversion in a Consumption-Savings Model." *Journal of Economic Behavior and Organization*, 38(2): 155-178.

Brown, Martin, Armin Falk and Ernst Fehr. 2004. "Relational Contracts and the Nature of Market Interactions.", *Econometrica*, 72(3): 747-780.

Brunnermeier, Markus and Jonathan Parker. 2005. "Optimal Expectations." *American Economic Review*, 95(4): 1092-1118.

Burks, Steven, Jeffrey Carpenter, Lorenz Goette and Aldo Rustichini. 2010. "Overconfidence is a Social Signaling Bias." working paper.

Camerer, Colin, and Dan Lovallo. 1999. "Overconfidence and Excess Entry: An Experimental Approach." *American Economic Review*, 89(1): 306-318.

Caplin, Andrew and John Leahy. 2001. "Psychological Expected Utility Theory and Anticipatory Feelings." *Quarterly Journal of Economics*, 116(1): 55-79.

Caplin, Andrew and John Leahy. 2004. "The Supply of Information by a Concerned Expert." *The Economic Journal*, 114(497): 487-505.

Chapman, Gretchen and Eric Johnson. 1999. "Anchoring, Activation, and the Construction of Values." *Organizational Behavior and Human Decision Processes*, 79(2): 115-153.

Charness, Gary Aldo Rustichini and Jeroen van de Ven. 2011. "Self-confidence and Strategic Deter- rence." working paper.

Chew, Soo Hong and Joanna Ho. 1994. "Hope: An Empirical Study of Attitude Toward the Timing of Uncertainty Resolution." *Journal of Risk and Uncertainty*, 8(3): 267-288.

Cho, In-Koo and David Kreps. 1987. "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics*, 102(2): 179-221.

Cialdini, Robert. 1984. "Influence, the Psychology of Persuasion." New York: Harper Collins.

Crawford, Vincent and Juanjuan Meng. 2011. "New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income." *American Economic Review*, 101 (5): 1912-1932.

DellaVigna, Stefano and Ulrike Malmendier. 2006. "Paying Not to Go to the Gym." *American Economic Review*, 96(3): 694-719.

Dillenberger, David. 2010. "Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior." *Econometrica*, 78(6): 1973-2004.

Dohmen, Thomas and Armin Falk. 2011. "Performance Pay and Multi-dimensional Sorting: Productivity, Preferences and Gender." *American Economic Review*, 101(2): 556-590.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp and Gert Wagner. 2011. "Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences." *Journal of the European Economic Association*, 9(3): 522-550.

Eil, David and Justin Rao. 2011. "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself." *American Economic Journal: Microeconomics*, 3(2): 114-138.

Eliaz, Kfir and Andrew Schotter. 2007. "Experimental Testing of Intrinsic Preferences for Non- Instrumental Information." *American Economic Review (Papers and Proceedings)*, 97: 166-169.

Eliaz, Kfir and Andrew Schotter. 2010. "Paying for Confidence: An Experimental Study of the Demand for Non-Instrumental Information." *Games and Economic Behavior*, 70(2): 304-324.

Ellingsen, Tore and Magnus Johannesson. 2004. "Promises, Threats and Fairness.", *The Economic Journal*, 114(April): 397-420.

Ellingsen, Tore and Magnus Johannesson. 2008. "Pride and Prejudice: The Human Side of Incentive Theory." *American Economic Review*, 98(3): 990-1008.

Epley, Nicholas and Thomas Gilovich. 2001. "Putting Adjustment Back in the Anchoring and Adjustment Heuristic: Differential Processing of Self-Generated and Experimenter-Provided Anchors." *Psychological Science*, 12 (5): 391-396.

Ericson, Keith and Andreas Fuster. (forthcoming). "Expectations as Endowments: Evidence on Reference-Dependent Preferences from Exchange and Valuation Experiments." *Quarterly Journal of Economics.*

Eyster, Erik. 2002. "Rationalizing the Past: a Taste for Consistency." working paper.

Fehr, Ernst and Lorenz Goette. 2007. "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment." *American Economic Review*, 97(1): 298-317.

Fellner, Gerlinde and Matthias Sutter. 2009. "Causes, Consequences, and Cures of Myopic Loss Aversion: An Experimental Investigation." *The Economic Journal*, 119(537):

900-916.

Festinger, Leon. 1957. "A Theory of Cognitive Dissonance." Stanford, CA: Stanford University Press.

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics*, 10(2): 171-178.

Fischbacher, Urs and Franziska Heusi. 2008. "Lies in Disguise: An Experimental Study on Cheating." working paper.

Freedman, Jonathan and Scott Fraser. 1966. "Compliance Without Pressure: The Foot in the Door Technique." *Journal of Personality and Social Psychology*, 4(2): 195-202.

Frey, Bruno and Alois Stutzer. 2001. "Happiness and Economics: How the Economy and Institutions Affect Human Well-Being." Princeton, New Jersey: Princeton University Press.

Fudenberg, Drew, David K. Levine and Zacharias Maniadis. (forthcoming). "On the Robustness of Anchoring Effects in WTP and WTA Experiments." *American Economic Journal: Microeconomics.*

Gächter, Simon and Ernst Fehr. 1999. "Collective Action as a Social Exchange." *Journal of Economic Behavior and Organization*, 39(4): 341-369.

Gill, David and Victoria Prowse. (forthcoming). "A Structural Analysis of Disappointment Aversion in a Real Effort Competition." *American Economic Review.*

Gneezy, Uri and Jan Potters. 1997. "An Experiment on Risk Taking and Evaluation Periods." *Quarterly Journal of Economics*, 112 (2): 631-645.

Gneezy, Uri. 2005. "Deception: The role of consequences." *American Economic Review*, 95 (1): 384-394.

Gneezy, Uri, Muriel Niederle and Aldo Rustichini. 2003. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics*, 118 (3): 1049-1074.

Gneezy, Uri and Aldo Rustichini. 2004. "Gender and Competition at a Young Age." *American Economic Review (Papers and Proceedings)*, 94 (2): 377-381.

Goeree, Jacob, Charles A. Holt, and Thomas R. Palfrey. 2008. "Quantal Response Equilibrium." New Palgrave Dictionary of Economics, Second Edition, Editors Steven N. Durlauf and Lawrence E. Blume, Palgrave Macmillan, 2008.

Green, Donald, Karen E. Jacowitz, Daniel Kahneman, and Daniel McFadden. 1998. "Referendum Contingent Valuation, Anchoring, and Willingness to Pay for Public Goods."*Resources and Energy Economics*, 20: 85-116.

Greiner, Ben. 2003. "An Online Recruitment System for Economic Experiments." In *Forschung und wissenschaftliches Rechnen 2003.* GWDG Bericht 63, ed Kurt Kremer and Volker Macho, pp. 79-93. Goettingen: Ges. fuer Wiss. Datenverarbeitung.

Groves, Robert, Floyd Fowler, Mick Couper, James Lepkowski, Eleanor Singer, and Roger Tourangeau. 2009. "Survey Methodology." 2nd Edition. Hoboken, NJ: Wiley.

Gul, Faruk. (1991). "A Theory of Disappointment Aversion." *Econometrica*, 59: 667-686.

Haigh, Michael and John List. 2005. "Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis." *Journal of Finance*, 60 (1): 523-534.

Hastings Justine S. and Jeffrey M. Weinstein. 2008. "Information, School Choice, and Academic Achievement: Evidence from Two Experiments." *Quarterly Journal of Economics*, 123(4), 1373-1414.

Heider, Fritz. 1946. "Attitudes and Cognitive Organization." *Journal of Psychology*, 21: 107-112.

Hoelzl, Erich and Aldo Rustichini . 2005. "Overconfident: Do You Put Your Money on It?" *The Economic Journal*, 115(4): 305-318.

Holt, Charles A. and Susan K. Laury. 2002. "Risk Aversion and Incentive Effects." *American Economic Review*, 92 (5): 1644-1655.

Jacowitz, Karen E. and Daniel Kahneman. 1995. "Measures of Anchoring in Estimation Tasks." *Personality and Social Psychology Bulletin*, 21: 1161-1184.

Johnson, Eric and David Schkade. 1989. "Bias in Utility Assessments: Further Evidence and Explanations." *Management Science*, 35(4): 406-424.

Kahneman, Daniel and Jack Knetsch. 1993. "Anchoring or Shallow Inferences: The Effect of Format." Unpublished manuscript. University of California, Berkeley.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica*, 47 (2): 263-291.

Kocher, Martin, Michal Krawczyk and Frans van Winden. 2009. "Let Me Dream on! Anticipatory Emotions and Preference for Timing in Lotteries." working paper.

Kőszegi, Botond. 2006. "Ego Utility, Overconfidence, and Task Choice." *Journal of the European Economic Association*, 4(4): 673-707.

Kőszegi, Botond and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics*, 121(4): 1133-1165.

Kőszegi, Botond and Matthew Rabin. 2007. "Reference-Dependent Risk Attitudes." *American Economic Review*, 97(4): 1047-1073.

Kőszegi, Botond and Matthew Rabin. 2009. "Reference-Dependent Consumption Plans." *American Economic Review*, 99(3): 909-936.

Kreps, David and Evan Porteus.(1978. "Temporal Resolution of Uncertainty and Dynamic Choice Theory." *Econometrica*, 46(1): 185-200.

Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics*, 112(2): 443-478.

Langer, Thomas and Martin Weber. 2008. "Does commitment or feedback influence myopic loss aversion?: An experimental analysis." *Journal of Economic Behavior and Organization*, 67(3-4): 810-819.

Layard, Richard. 2005. "Happiness: Lessons From a New Science." London, Penguin Books.

Loomes, Graham and Robert Sugden. 1986. "Disappointment and Dynamic Consistency in Choice Under Uncertainty." *The Review of Economic Studies*, 53 (2): 271-282.

Sandra Ludwig and Julia Nafziger. 2011. "Beliefs about Overconfidence." *Theory and Decision*, 70 (4): 475-500.

Malmendier, Ulrike and Geoffrey Tate. 2008. "Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction." *Journal of Financial Economics*, 89(1): 20-43.

Moebius, M. M., M. Niederle, P. Niehaus, and T. S. Rosenblat. 2011. "Managing Self-Confidence: Theory and Experimental Evidence," working paper.

Moffitt, R. 1983. "An Economic Model of Welfare Stigma," *American Economic Review*, 73(5): 1023-1035.

Mullainathan, Sendhil and Ebonya Washington. 2009. "Sticking with Your Vote: Cognitive Dissonance and Political Attitudes." *American Economic Journal: Applied Economics*, 1 (1): 86-111.

Newcomb, Theodore. 1953. "An Approach to the Study of Communicative Acts." *Psychological Review*, 60: 393-404.

Niederle, Muriel and Lise Vesterlund. 2007. "Do Women Shy Away from Competition? Do Men Compete too Much?" *Quarterly Journal of Economics*, 122(3): 1067-1101.

Palacios-Huerta, Ignacio. 1999. "The Aversion to the Sequential Resolution of Uncertainty." *Journal of Risk and Uncertainty*, 18: 249-269.

Prendergast, Canice and Lars Stole. 1996. "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning." *Journal of Political Economy*, 104 (6): 1105-1134.

Rabin, Matthew and Joel Schrag. 1999. "First Impressions Matter: A Model of Confirmatory Bias." *Quarterly Journal of Economics*, 114 (1): 37-82.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review*, 83(5): 1281-1302.

Rege, Mari and Kjetil Telle. 2004. "The Impact of Social Approval and Framing on Cooperation in Public Good Situations." *Journal of Public Economics*, 88(7-8): 1625-1644.

Reuben, Ernesto, Pedro Rey-Biel, Paola Sapienza, and Luigi Zingales. (forthcoming). "The Emergence of Male Leadership in Competitive Environments." *Journal of Economic Behavior and Organization.*

Selten, Reinhard. 1975. "Reexamination of the perfectness concept for equilibrium points in extensive games." *International Journal of Game Theory*, 4(1): 25-55.

Schwarz, Norbert, Fritz Strack and Hans-Peter Mai. 1991. Assimilation and Contrast Effects on Part-Whole Questions Sequences:A Conversational Logic Analysis." *Public Opinion Quarterly*, 55, 3-23.

Sherman, Steven. 1980. "On the Self-Erasing Nature of Errors of Prediction." *Journal of Personality and Social Psychology*, 39(2): 211-221.

Strack, Fritz and Thomas Mussweiler. 1997. "Explaining the Enigmatic Anchoring Effect: Mechanisms of Selective Accessibility." *Journal of Personality and Social Psychology*, 73: 437-446.

Surowiecki, James. 2004. "The Wisdom of Crowds." New York: Random House.

Svenson, Ola. 1981. "Are We All Less Risky and More Skillful Than Our Fellow Drivers?" *Acta Psychologica*, 47: 143-148.

Thaler, Richard and Shlomo Benartzi. 2004. "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy*, 112(1): 164-187.

Tourangeau, Roger. 1999. "Context Effects on Answers to Attitude Questions." In *Cognition and Survey Research*, edited by M. G. Sirken, D. J. Herrmann, S. Schechter, N. Schwarz, J. M. Tanur, and R. Tourangeau. New York: John Wiley Sons Inc.

Tversky, Amos and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science*, 185: 1124-1131.

Vanberg, Christoph. 2008. “Why Do People Keep Their Promises? An Experimental Test of Two Explanations.” *Econometrica*, 76(6): 1467-1480.

Van Winden, Frans, Michal Krawczyk and Astrid Hopfensitz. 2011. “Investment, Resolution of Risk, and the Role of Affect.” *Journal of Economic Psychology*, 32: 918-939.

Yariv, Leeat. 2005. “I’ll See it When I Believe it - A Simple Model of Cognitive Consistency.” working paper.

Chapter 5

Appendices

Appendix A (Chapter 1)

Formal Prediction

We now formally derive the prediction of Kőszegi and Rabin (2009) we are testing in chapter 1. We closely follow the notation of the original model. The model is in discrete-time with $T + 1$ periods, 0 through T. Decision-makers consume K goods. In all periods $t \geq 1$, consumption $c_t = (c_t^1, ..., c_t^K)$ is realized. At the beginning of period t, the decision-maker holds beliefs $F_{t-1} = \{F_{t-1,\tau}\} \big|_{\tau=t}^{T}$, with $F_{t-1,\tau} = (F_{t-1,\tau}^1, ..., F_{t-1,\tau}^K)$ being the belief about the consumption vector in period τ. Then, some signals may arrive and the decision-maker accordingly forms new beliefs $\{F_{t,\tau}\} \big|_{\tau=t}^{T}$, where no uncertainty is left regarding consumption in period t.

Instantaneous period-t utility depends on consumption in t and on belief changes in t regarding contemporaneous and future consumption:

$$u_t = m(c_t) + \sum_{\tau=t}^{T} \gamma_{t,\tau} N(F_{t,\tau} \,|F_{t-1,\tau})$$

$m(c_t)$ denotes reference-independent consumption utility and the terms $N(F_{t,\tau} \,|F_{t-1,\tau})$ represent "gain-loss utility" from belief changes. $\gamma_{\tau,\tau} \geq \gamma_{\tau-1,\tau} \geq ... \geq \gamma_{0,\tau} \geq 0$ are the weights on gain-loss utilities. $\gamma_{t,t}$ is normalized to 1. The weights γ represent the importance of new information depending on how far in advance of actual consumption the news are received. Importance decreases, the earlier new information realized.

Gain-loss utilities are specified such that decision-makers make ordered comparisons between current and previous beliefs about consumption. It is assumed that decision-makers compare the worst percentile of outcomes under current beliefs to that under previous beliefs, the second-worst percentile under current and previous beliefs and so on.

Formally we define percentile p implicitly by stating that for any distribution F over $\mathbb{R}$ and any $p \in (0,1)$, the consumption level at p, $c_F(p)$ is defined by $F(c_F(p)) \geq p$ and by $F(c) < p$ for all $c < c_F(p)$. Then we can define gain-loss utility from the change in beliefs in consumption dimension k as:

$$N^k(F^k_{t,\tau} \Big| F^k_{t-1,\tau}) = \int_0^1 \mu(m^k(c_{F^k_{t,\tau}}(p)) - m^k(c_{F^k_{t-1,\tau}}(p)))dp$$

$\mu()$ is a "standard" gain-loss utility function with the following properties taken from Bowman et al. (1999):

1. $\mu(x)$ is continuous for all x, twice differentiable for $x \neq 0$, and $\mu(0) = 0$.
2. $\mu(x)$ is strictly increasing.
3. If $y > x \geq 0$, then $\mu(y) + \mu(-y) < \mu(x) + \mu(-x)$.
4. $\mu''(x) \leq 0$ for $x > 0$ and $\mu''(x) \geq 0$ for $x < 0$.
5. $[\mu'_-(0)]/[\mu'_+(0)] \equiv \lambda > 1$, where $\mu'_+(0) \equiv lim_{x\to 0}\mu'(|x|)$ and $\mu'_-(0) \equiv lim_{x\to 0}\mu'(-|x|)$.

Within these properties, loss aversion is captured in properties 3 and 5, diminishing sensitivity is captured by property 4.

Total gain-loss utility in period t is now assumed to be the sum of gain-loss utilities in each dimension, i.e. $N(F_{t,\tau} \Big| F_{t-1,\tau}) = \sum_{k=1}^{K} N^k(F^k_{t,\tau} \left| F^k_{t-1,\tau}\right.)$.

As a last step, it is assumed that the decision-maker wants to maximize the expected sum of instantaneous utilities,

$$U^t \equiv \sum_{\tau=t}^{T} u_\tau$$

We now have all the ingredients necessary to make predictions about informational preferences. Following Kőszegi and Rabin (2009), information here means information regarding "fixed but unknown future consumption". In other words, information has to be on exogenous events that cannot be influenced by the decision-maker.

For simplicity, we assume that consumption takes place only in period T. Decision-makers may receive information about consumption from period 1 to $T-1$. σ be a sequence of signals, $s_1, s_2, ..., s_J$ and $t(s_j \,|\sigma)$ denote the time of arrival of signal s_j under σ.

We want to make predictions about decision-makers preferences towards different information structures. For this purpose we introduce the following terminology. We call σ' to be (t_a, t_b, j)-equivalent to σ if both involve the same sequence of signals, if in both σ and σ' only s_j and s_{j+1} arrive between t_a and t_b (with $t_b > t_a$) and if for all $i \neq j, j+1$, we have that $t(s_i \,|\sigma') = t(s_i \,|\sigma)$. Thus, if two sequences of signals are (t_a, t_b, j)-equivalent, they only differ in the timing of the two signals s_j and s_{j+1}.

The model of Kőszegi and Rabin makes the following central prediction. Clumping information is utility-increasing as long as no information is delayed through clumping. This is captured in the following proposition.[1]

PROPOSITION from Kőszegi and Rabin (2009): Say that σ' is (t_a, t_b, j)-equivalent to σ and $t(s_{j+1} \,|\sigma') = t(s_j \,|\sigma') \leq t(s_j \,|\sigma) < t(s_{j+1} \,|\sigma)$. Then we have that $U(\sigma') > U(\sigma)$ for any $\gamma_{t,T} > 0$ nondecreasing in t.

Applying this proposition to our treatments, we get that subjects strictly prefer the clumped information condition over the piecewise condition and subsequently choose to be informed in one piece in treatments 1 and 2. This can be easily shown by iteratively applying the proposition. Consider a hypothetical information sequence σ^h where subjects learn the first piece of information on Tuesday and on Wednesday they learn the final outcome, i.e., whether they won or lost. Clearly, for the comparison of σ^h with the information sequence in the piecewise condition, one can see that proposition 1 applies, stating that subjects should strictly prefer σ^h. For the comparison of σ^h and the information sequence in the clumped condition, we can again apply proposition 1, giving us that subjects should strictly prefer the sequence of the clumped condition to σ^h. Therefore, subjects should strictly prefer the clumped condition over the piecewise condition and consequently select the clumped condition. Likewise, average willingness to pay for the lottery should be higher in treatment 3 compared to treatment 4.

[1]For a proof of this proposition we refer to Kőszegi and Rabin (2009).

Instructions

Instructions translated into English.

We present instructions for treatment 1. Instructions for treatment 2 are identical, only the description of the lottery was changed accordingly.

Instructions for treatments 3 and 4 only differed in that only one of the two information conditions was explained and that the price list format was explained. Explanations for the price list format are provided at the end of this Appendix.

General Explanations

You are participating in a decision making experiment. In this experiment you can earn money.
Please read these explanations carefully.

In this experiment you are participating in a lottery, which will be explained to you in detail below. You will get your earnings from the experiment in the experimenter's office at the following date.

Friday, XX.XX.XX, 9am-12pm and 2pm-5pm
(Note: exact dates varied between sessions)

In case you are unable to come at that date, there is a possibility to set another date with the experimenter. You will receive further information about this at the end of the experiment.

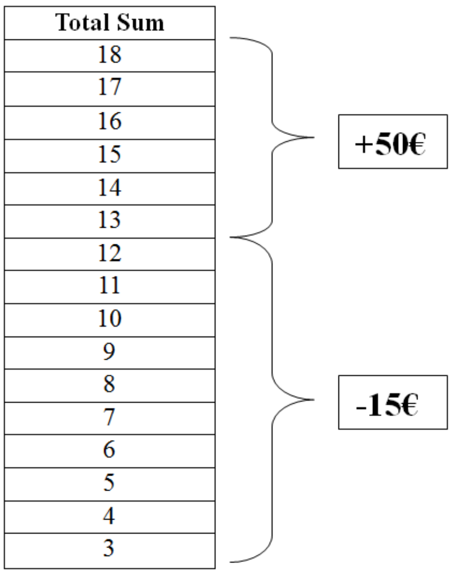
During the experiment you are not allowed to communicate with other subjects. If you have a question at any time, raise your hand and the experimenter will come to your place to answer it.

The Lottery

In this experiment you are participating in a lottery. Part of the lottery is a starting endowment of 30 Euro. In the lottery you can either win 50 Euro or lose 15 Euro. In case you lose 15 Euro they are deducted from your endowment of 30 Euro. In case you win 50 Euro they are added to your endowment of 30 Euro.

In the lottery a regular dice is thrown three times. After the three dice throws the numbers thrown are added up and determine whether you win or lose. If the total sum after three dice throws is larger or equal to 13, you win 50 Euro. If the total sum after three dice throws is smaller than 13, you lose 15 Euro.

The following graph illustrates this:



Your Obligations in this Experiment

From Tuesday to Thursday this week you have to call the experimenter once a day from 9am-12pm or 2pm-5pm. The phone calls will each not take longer than a minute. You will receive further information such as the telephone number and again the times you can call at the end of the experiment. **Note that you lose all your earnings from the experiment in case you fail to call every day.**

During the phone calls you will be asked a couple of questions. On Tuesday and Wednesday you will be reminded to call the next day. On Thursday you will be reminded of the day you will receive your earnings. In addition you will receive information regarding the outcome of the lottery during the phone-calls.

Your Decision

Your choice in this experiment is to decide how you want to be informed about the outcome of the lottery. There are two possibilities: either you learn the outcome in the first phone call on Tuesday or you will be informed step by step, from Tuesday until Thursday, about the outcome.

If you decide to learn the outcome on Tuesday, the experimenter will inform you during the phone-call on Tuesday whether you won or lost in the lottery. In addition he will tell you the outcome of each of the three dice throws so that you will know how your earnings from the lottery were determined. During the phone-calls on Wednesday and Thursday you will receive no further information about the lottery.

If you decide to be informed step by step from Tuesday until Thursday, the experimenter will inform you on Tuesday about the outcome of the first dice throw. During the phone-call on Wednesday he will tell you the result of the second dice throw. On Thursday you will be informed about the outcome of the third dice throw and thus whether you won or lost. Thus, you will not know the total sum of the three dice throws until Thursday.

The following graph illustrates the two possibilities how you are informed about the outcome of the lottery:

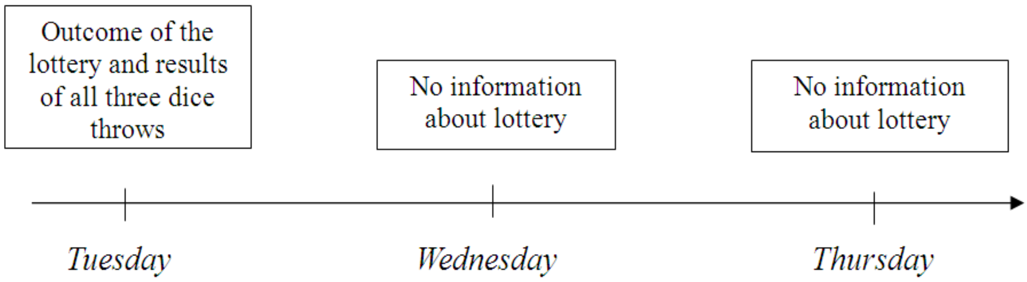


Figure 5.1: All information on Tuesday

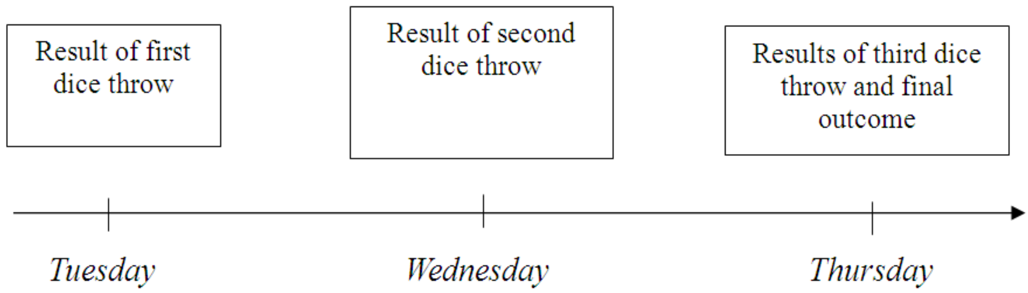


Figure 5.2: Information step by step

Your choice in this experiment is how you want to be informed about the outcome of the lottery. Choose carefully! Note that there are no "right" or "wrong" choices. Simply make your choice according to your own preferences.

Note that although you will be informed about the outcome of the lottery at a later stage, the experimenter will conduct the dice throws right after the end of the experiment. The experimenter will throw three dice for each participant and will enter the results in a table. At the day you receive your payments, you will be able to see the table containing the outcomes of the lottery for all participants in an anonymous way. Thus, you will be able to see how many participants won in the lottery and how many lost. But no other participant will be able to tell whether you won or lost in the lottery.

Summary

- In this experiment you participate in a lottery. Part of the lottery is a starting endowment of 30 Euro. Your earnings from the experiment will be paid to you on Friday the XX.XX.XX.
- In the lottery a regular dice is thrown three times. If the total sum after three dice throws is greater or equal to 13 you win 50 Euro in addition to your endowment of 30 Euro. If the total sum after three dice throws is smaller than 13 you lose 15 Euro which are deducted from your endowment of 30 Euro.
- You have to call the experimenter once a day from Tuesday until Thursday. If you fail to call you lose all your earnings from the experiment.
- Your choice in this experiment: you can decide how you want to be informed about the outcome of the lottery. There are two possibilities: either you learn the outcome during the first phone call on Tuesday or you will be informed step by step, from Tuesday until Thursday, about the outcome.

Please read the instructions again carefully. If you have any questions, raise your arm and the experimenter will be happy to help you. Shortly we will distribute control questions to make sure that you understand the instructions. After this you can make your choice.

Note: Explanations of the price list format in treatments 3 and 4.

Your Choice

In a series of decisions you can choose between a fixed payment and participation in the lottery. The following table should illustrate how these choices will look like. You will always have the choice between a fixed payment that you will receive on Friday the XX.XX.XX (Note: day was always identical to the day subjects would receive payments if they participated in lottery) and participation in the lottery. Note that you also need to call from Tuesday until Thursday if you choose the fixed payment.

	Fixed Payment	Lottery
1)	Fixed Payment of euro13	Participation in Lottery
2)	Fixed Payment of euro14	Participation in Lottery
3)	Fixed Payment of euro15	Participation in Lottery
4)	Fixed Payment of euro16	Participation in Lottery
5)	Fixed Payment of euro17	Participation in Lottery
6)	Fixed Payment of euro18	Participation in Lottery
7)	Fixed Payment of euro19	Participation in Lottery
8)	Fixed Payment of euro20	Participation in Lottery
9)	Fixed Payment of euro21	Participation in Lottery
10)	Fixed Payment of euro22	Participation in Lottery
11)	Fixed Payment of euro23	Participation in Lottery
12)	Fixed Payment of euro24	Participation in Lottery
13)	Fixed Payment of euro25	Participation in Lottery
14)	Fixed Payment of euro26	Participation in Lottery
15)	Fixed Payment of euro27	Participation in Lottery
16)	Fixed Payment of euro28	Participation in Lottery
17)	Fixed Payment of euro29	Participation in Lottery
18)	Fixed Payment of euro30	Participation in Lottery
19)	Fixed Payment of euro31	Participation in Lottery
20)	Fixed Payment of euro32	Participation in Lottery
21)	Fixed Payment of euro33	Participation in Lottery
22)	Fixed Payment of euro34	Participation in Lottery
23)	Fixed Payment of euro35	Participation in Lottery
24)	Fixed Payment of euro36	Participation in Lottery
25)	Fixed Payment of euro37	Participation in Lottery

In all 25 decisions you need to select the alternative you prefer. One out of the 25 decisions will be randomly chosen and implemented. If for example you decided to choose the fixed payment of 19 Euro in decision 7 and decision 7 is randomly chosen, then your choice from decision 7 will be implemented and thus you will receive a fixed payment of 19 Euro on Friday the XX.XX.XX.

At the end of the experiment the experimenter will determine in front of you which of your choices will be implemented. Thus you will know at the end of the experiment whether you will participate in the lottery or receive a fixed payment. The other participants will not be able to see this. Thus, no other participant will know which choice was implemented for you.

Make your choices carefully! Note that there are no “right” or “wrong” choices. Make your choice according to your preferences.

Appendix B (Chapter 2)

Uniqueness of equilibrium described in Proposition 1:

First note that in every equilibrium, the types with very low image concerns ($\alpha \to 0$) will always choose the money-maximizing outcome, i.e., those with $p < \overline{r}$ optimally choose $x = No$ and those with $p > \overline{r}$ optimally choose $x = Yes$. Next we show that there cannot be an equilibrium where decision makers with $p > \overline{r}$ do not choose $x = Yes$. Suppose there would be such an equilibrium. Then the image utility from choosing $x = No$ necessarily would need to be greater than the image utility from $x = Yes$. In that case however, all D with $p < \overline{r}$ would also choose $x = No$. This leads to a contradiction because then the public will infer a lower p from $x = No$ than from $x = Yes$ and consequently image utility from $x = Yes$ would be higher. Thus in every equilibrium, some D with $p < \overline{r}$ and low values of α will choose $x = No$ and all D with $p > \overline{r}$ choose $x = Yes$. Also, by assumption $\overline{\alpha}$ is large enough such that some D with $p < \overline{r}$ choose $x = Yes$. From that it is easy to see that every equilibrium has a threshold type α^*, such that decision makers with $p < \overline{r}$ and $\alpha > \alpha^*$ will choose $x = Yes$ and those with $\alpha < \alpha^*$ will choose $x = No$. From equation (2) we see that α^* and consequently the equilibrium described above is unique.

Clarification Proposition 3:

We assume that there exist decision makers with image concerns large enough such that they would choose $x = Yes$ if $\hat{p} < \overline{r}$ and if all other decision makers would simply maximize monetary outcomes. More precisely, $\overline{\alpha}$ is large enough such that the image gains from choosing $x = Yes$, $\overline{\alpha}\beta\left[\frac{\sum_{p>\overline{r}}(1-\phi(p))f(p)p+\sum_{p<\overline{r}}\phi(p)f(p)p}{\sum_{p>\overline{r}}(1-\phi(p))f(p)+\sum_{p<\overline{r}}\phi(p)f(p)} - \frac{\sum_{p<\overline{r}}(1-\phi(p))f(p)p+\sum_{p>\overline{r}}\phi(p)f(p)p}{\sum_{p<\overline{r}}(1-\phi(p))f(p)+\sum_{p>\overline{r}}\phi(p)f(p)}\right]$ outweigh the monetary costs y. Note that this condition is more demanding than that in the case of perfect information. The reason is that type uncertainty reduces the reputational gains from choosing $x = Yes$. Therefore image concerns need to be higher in the case of imperfect knowledge of own type.

Dependent variable: Relative self-assessment= $\begin{cases} 1 & \text{if better} \\ 0 & \text{if worse} \end{cases}$	
Dummy treatment	1.36
	(1.39)
Gender dummy	-0.19
	(0.16)
I_Treatment*Gender dummy	-0.23
	(0.28)
Quiz performance	0.06*
	(0.03)
I_Treatment*Quiz performance	0.05
	(0.06)
Controls	included
N	95
-LL	42

Notes: Probit estimates. Marginal effects reported; robust standard errors are in parentheses. Significance at the 1, 5, and 10 percent level is denoted by ***, **, and *, respectively. *Dummy treatment* =1 if audience treatment and 0 if private treatment. *Dummy gender* =1 if female. *Controls* include the survey based risk measure, image concerns, age, relationship status, and interactions of the *Dummy treatment* with each variable.

Table 5.1: Probit regression with treatment interactions.

Instructions

Instructions, translated into English. General instructions and instructions for the first part of the experiment were identical across treatments. Instructions for the second part of the experiment differed across treatments - below you find the versions from the audience and the private treatment. The second part of instructions in the feedback treatment was identical, apart from the additional explanation that the experimenter would reveal the correct self-assessment after subjects publicly reported their self-assessment.

GENERAL INSTRUCTIONS

You are taking part in a decision-making experiment in which you have the opportunity to earn money. The amount of money you earn is paid to you upon completion of the experiment. Please read the instructions carefully. The instructions are identical for all participants. If you have any questions, please raise your hand. The experimenter will answer your question at your place. During the experiment, you have to remain silent. Violation of this rule leads to immediate exclusion from the experiment and all payments.

All monetary units in the experiment are measured in tokens, and 100 tokens = 1 Euro.

This experiment consists of two parts. In both parts, you can earn money. Your payoff from the experiment results from the sum of your payoffs in both parts. In the following we will go through the instructions for the first part of the experiment. After the first part is completed, we will provide you with the instructions of the second part.

INSTRUCTIONS FOR THE FIRST PART OF THE EXPERIMENT

In the first part of the experiment you will be asked 20 quiz questions. You will always be offered 4 possible answers of which exactly one will be correct. Please always select one of the four possible answers. For each correct answer you get 40 tokens. After you have answered the first 10 questions, please click on the OK button. Then a new screen with 10 more questions will appear. Please confirm your responses again with the OK button.

Do you have any questions?

INSTRUCTIONS FOR THE SECOND PART OF THE EXPERIMENT - *(Private Treatment)*

All participants have answered 20 quiz questions in the first part of the experiment. In this part of the experiment, you need to assess whether your quiz result is better or worse than the average result of another group of participants. If your assessment is correct, you get 500 tokens; if your assessment is wrong, you get 0 tokens. This will be further explained below in more detail.

The quiz questions you were asked in the first part of the experiment, were also answered by a group of 95 participants (all of which (like you) participated in an experiment in the BonnEconLab) some time ago. You now need to assess whether your performance in the quiz was better or worse than the average performance of the group of 95 participants. You get 500 tokens for a correct assessment, otherwise you get 0 tokens.

Please read these instructions again carefully.

An input box appears soon on your screen into which you can enter your decision.

Do you have any questions?

INSTRUCTIONS FOR THE SECOND PART OF THE EXPERIMENT - *(Audience Treatment)*

All participants have answered 20 quiz questions in the first part of the experiment. In this part of the experiment, you need to assess whether your quiz result is better or worse than the average result of another group of participants. If your assessment is correct, you get 500 tokens; if your assessment is wrong, you get 0 tokens. This will be further explained below in more detail.

Note the following: After all participants entered their assessment into the computer, all participants must report their assessment to the other participants. Every participant will be called up individually one after the other. Once it is your turn, you have to stand up, say your name and report your assessment.

So if you stated that you think your quiz result was better than the average of the other group, then you have to stand up after you were called and say: "My name is ... and I think I was better than the average of the other group."

If you stated that you think your quiz result was worse than the average of the other group, then you have to stand up after you were called and say: "My name is ... and I think I was worse than the average of the other group."

Below we will explain your decision in more detail.

The quiz questions you were asked in the first part of the experiment, were also answered by a group of 95 participants (all of which (like you) participated in an experiment in the BonnEconLab) some time ago. You now need to assess whether your performance in the quiz was better or worse than the average performance of the group of 95 participants. You get 500 token for a correct assessment, otherwise you get 0 token.

Please read these instructions again carefully.

An input box appears soon on your screen into which you can enter your decision.

Do you have any questions?

Appendix C (Chapter 3)

Figure 5.3: Picture of bowl with peas: This bowl was shown to subjects in the estimation task of experiment 2.

Forward Looking Behavior

We show that the equilibrium described in section 2.2 is also an equilibrium if decision-makers are forward-looking and anticipate the period 2 choice. We do so via backward-induction.

First consider period 2 choices conditional on equilibrium period 1 choices, i.e., $x_1 = m_1$ for low types and $x_1 = \mu$ for high types. This is precisely the situation we already solved in section 2.2. High types will choose consistently and select $x_2 = \mu$. Low types will always behave consistently as well, choosing $x_2 = x_1 = m_1$ regardless of the signal the receive in period 2.

Now we derive period 2 behavior if period 1 choices differ from the equilibrium path. First consider a high type and suppose that w.l.g. D_H chose $x_1 = Blue$ although $\mu = Red$. In period 2 she now has the choice between being consistent by choosing $x_2 = Blue$ (thereby sacrificing standard utility) and being inconsistent by choosing $x_2 = Blue$ (thereby sacrificing image utility). For our purposes it is sufficient to note that whatever D_H is choosing, period 2 utility will always be lower than in the case where D_H chose $x_1 = Red$ and then consistently chooses $x_2 = Red$.

Next consider low types who w.l.g. chose $x_1 = Blue$ although $m_1 = Red$. Here we need to distinguish behavior depending on the period 2 signal. Suppose the period 2 signal differs from the period 1 signal, i.e., $m_2 = Blue$. Then D_L maximizes both standard and reputational utility by choosing consistently $x_2 = Blue$. If $m_2 = Red$, D_L faces a trade-off between standard and image utility. Depending on the strength of image concerns, they either optimally choose consistently ($x_2 = Blue$) or inconsistently ($x_2 = Red$).

Having solved period 2 behavior, we can consider behavior in period 1. First, w.l.g. consider a high type with $\mu = Red$. We have seen that in terms of period 2 utility she will be better of choosing $x_1 = Red$. Since $x_1 = Red$ also maximizes period 1 utility, types D_H optimally choose $x_1 = Red$. The same argument can be made for low types. Again w.l.g. assume that $m_1 = Red$. Consider the expected period 2 utility if $x_1 = Blue$. We focus on the case where image concerns are sufficiently large such that D_L optimally chooses $x_2 = Blue$, regardless of m_2.[2] Compare this to the expected period 2 utility along the equilibrium path, i.e., if $x_1 = Red$. While image utility in both cases is identical (both

[2]Note that the case where image concerns are not sufficiently large such that decision-makers prefer to choose $x_2 = Red$ if $m_2 = Red$ can be solved analogously.

always behave consistently), expected standard utility is lower if $x_1 = Blue$. The reason is that decision-makers for $x_1 = Blue$ need to sacrifice a lot of standard utility for being consistent in the case where $m_1 = m_2 = Red$. Obviously expected period 1 utility is also maximized by $x_1 = Red$. Consequently types D_L optimally choose $x_1 = Red$.

Thus we have shown that the equilibrium developed in section 2.2 is also an equilibrium if decision-makers are forward looking.

Prediction Experiment 2

In both treatments types D_L receive a private signal m_1 and a public signal (the valuable information) m_2^p and shigh types (D_H) perfectly learn the true outcome.[3] We assume that the public signal is of greater strength, i.e., $p_2^p > p_1$ and without loss of generality that $m_2^p = Blue$. In the main treatment, low types commit to their private signal, i.e., $x_1 = m_1$ and high types commit to the choice they know with certainty to be true.

For subject's final choices we proceed in two steps. First we consider an equilibrium where low types do not respond to the public signal. We show that in the main treatment, if reputational concerns are large enough, there exists an equilibrium where low types do not respond to the public signal whereas in the control treatment such an equilibrium does not exist. We then consider the equilibrium where decision-makers fully respond to the public signal. We derive conditions for existence of this type of equilibrium and show that these conditions are either more likely or impossible to be fulfilled in the control treatment than in the main treatment.

First consider the equilibrium where D does not respond to the public signal. In the main treatment this equilibrium is characterized as follows. High types will choose the outcome they know to be true, both in the first decision and the final decision. Low types will first choose following their private signal m_1, i.e., some will choose $x_1 = Red$ and some $x_1 = Blue$. Final decisions of the low types will all be consistent, i.e., D_L's who chose $x_1 = Red$ will choose $x_2 = Red$, neglecting the public signal. For this to be an equilibrium it suffices to check incentive compatibility for low types who received a private signal $m_1 = Red$. Costs in terms of standard utility from choosing *Red* over *Blue* are $1 - 2l(m_2^p = Blue, m_1 = Red)$. Reputational gains are

[3] High types of course also receive the public signal, but do not need to respond to it since they already know the correct answer with certainty.

$\beta\frac{\alpha(1-p_2^p)}{\alpha(1-p_2^p)+(1-\alpha)(1-p_2^p)p_1+(1-\alpha)p_2^p(1-p_1)}$, since the public infers that the decision-maker is a low type if she chooses inconsistently, i.e., $Pr(type = D_L|Blue, Red) = 1$ and that $Pr(type = D_L|Red, Red) = \frac{(1-\alpha)(1-p_2^p)p_1+(1-\alpha)p_2^p(1-p_1)}{\alpha(1-p_2^p)+(1-\alpha)(1-p_2^p)p_1+(1-\alpha)p_2^p(1-p_1)}$.[4] Thus, if the following condition is fulfilled, there exists a Perfect Bayesian Equilibrium in the main treatment where subjects do not respond to the valuable public signal.

$$\beta\frac{\alpha(1-p_2^p)}{\alpha(1-p_2^p)+(1-\alpha)(1-p_2^p)p_1+(1-\alpha)p_2^p(1-p_1)} \geq 1-2*l(m_2^p = Blue, m_1 = Red).$$

In the control treatment this equilibrium is characterized by high types choosing the outcome they know to be true, low types who received a private signal $m_1 = Red$ choosing $x = Red$ (thereby ignoring the public signal) and low types who received a private signal $m_1 = Blue$ choosing $x = Blue$. For this to be an equilibrium we again check incentive compatibility for low types who received a private signal $m_1 = Red$.[5] Their costs in terms of standard utility from choosing *Blue* are identical to above, i.e., $1-2l(m_2^p = Blue, m_1 = Red)$. The reputational gains differ, however. Without commitment, the public cannot infer that D is a low type from $x = Blue$. Instead we have that $Pr(type = D_L|Blue) = \frac{(1-\alpha)p_2^p*p_1+(1-\alpha)(1-p_2^p)(1-p_1)}{\alpha*p_2^p+(1-\alpha)p_2^p*p_1+(1-\alpha)(1-p_2^p)(1-p_1)})$. Therefore reputational gains from choosing *Red* are smaller and amount to $\beta(\frac{\alpha(1-p_2^p)}{\alpha(1-p_2^p)+(1-\alpha)(1-p_2^p)p_1+(1-\alpha)p_2^p(1-p_1)} - \frac{\alpha*p_2^p}{\alpha*p_2^p+(1-\alpha)p_2^p*p_1+(1-\alpha)(1-p_2^p)(1-p_1)})$. Since costs in terms of standard utility are positive, positive reputational gains are a prerequisite for existence of this equilibrium. However, simplification of the inequality $\beta(\frac{\alpha(1-p_2^p)}{\alpha(1-p_2^p)+(1-\alpha)(1-p_2^p)p_1+(1-\alpha)p_2^p(1-p_1)} - \frac{\alpha*p_2^p}{\alpha*p_2^p+(1-\alpha)p_2^p*p_1+(1-\alpha)(1-p_2^p)(1-p_1)}) \geq 0$ yields $1-p_2^p \geq p_2^p$, which is impossible to be fulfilled since $p_2^p > \frac{1}{2}$.

Thus, in the main treatment, if reputational concerns are large enough, there exists a Perfect Bayesian Equilibrium where types D_L who received a signal $m_1 = Red$ choose consistently and thereby ignore the public signal. This type of equilibrium does not exist in the control treatment. Intuitively, in the main treatment low types who receive a private signal $m_1 = Red$ can only respond to the public signal by fully revealing that they

[4]In line with the baseline model we assume off-equilibrium beliefs to be such that the public infers a low type from inconsistent behavior.

[5]Note that here we also need to check incentive compatibility for low types that received a private signal $m_1 = Blue$ and high types who know the true outcome to be *Blue*. For our purposes, however it is sufficient to realize that these conditions only make the existence of this type of equilibrium in the control treatment even less likely.

are a low type while low types in the control treatment face lower reputational costs since they can "hide" behind the high types that are choosing *Blue*.

Now consider the equilibrium where the public signal is fully taken into account. In the main treatment this equilibrium is characterized as follows. High types will choose the outcome they know to be true, both in the first and in the final decision. Low types will take a first decision following their private signal m_1, i.e., some will choose $x_1 = Red$ and some $x_1 = Blue$. Since the public signal $m_2^p = Blue$ is more informative than the private signal, all low types will take the final decision $x_2 = Blue$, taking the public signal fully into account. Thus, the low types who received as private signal $m_1 = Red$ will behave inconsistently. For this to be an equilibrium it suffices to check incentive compatibility for low types who received as private signal $m_1 = Red$.[6] Their reputational cost from choosing *Blue* over *Red* is β, since the public infers that the decision-maker is a high type if she chooses $x_1 = x_2 = Red$, i.e., $Pr(type = D_L|Red, Red) = 0$ and that the decision-maker is a low type if she chooses $x_1 = Red$ and $x_2 = Blue$, i.e., $Pr(type = D_L|Blue, Red) = 1$. Gains in term of standard utility from choosing *Blue* over *Red* are $1 - 2l(m_2^p = Blue, m_1 = Red)$. Thus, if the following condition is fulfilled, there exists a Perfect Bayesian Equilibrium in the main treatment where subjects take the public signal fully into account.

$$\beta \leq 1 - 2 * l(m_2^p = Blue, m_1 = Red).$$

In the control treatment where subjects only take one decision this equilibrium is characterized by high types choosing the outcome they know to be true and low types choosing $x = Blue$ regardless of their private signal m_1. For this to be an equilibrium it again suffices to check incentive compatibility for low types who received a private signal $m_1 = Red$. Their gains in terms of standard utility from choosing *Blue* are identical to above, i.e., $1 - 2l(m_2^p = Blue, m_1 = Red)$. The reputational costs, however, differ. While the public infers a high type from $x = Red$, i.e., $Pr(type = D_L|Red) = 0$, it cannot infer that D is a low type from $x = Blue$. Instead we have that $Pr(type = D_L|Blue) = \frac{1-\alpha}{(1-\alpha)+p_2^p*\alpha}$. We end up with the following condition for existence of a Perfect Bayesian Equilibrium in the control treatment where subjects take the public signal fully

[6] For a full characterization of the equilibrium we also need to specify off-equilibrium beliefs. Again we assume that the public infers low strength from choices $x_1 = Blue$, $x_2 = Red$.

into account.

$$\beta(1 - \frac{p_2^p * \alpha}{(1-\alpha) + p_2^p * \alpha}) \leq 1 - 2 * l(m_2^p = Blue, m_1 = Red).$$

Comparing equilibrium conditions for both treatments, one can easily see that the condition in the control treatment is easier to be fulfilled since $1 - \frac{p_2^p * \alpha}{(1-\alpha)+p_2^p * \alpha} < 1$. Thus, an equilibrium where information is fully taken into account is less likely in the main treatment than in the control treatment.

Instructions - Experiment 1: Consistency as a Signal of Ability

Translated into English - Instructions for principal-agent treatment.

All subjects received general instructions. Then, depending on their role (principal or agent), they received additional more specific instructions (which we simply call Instructions).

Subjects in the agent treatment also first received general instructions and then more specific ones. Instructions were identical to those in the principal-agent treatment (for the agents) except for all parts describing the principal-agent structure.

Subjects in the principal treatment received instructions almost identical to the instructions principals in the principal-agent treatment received. The main difference was that subjects now were explained that they had to select subjects who performed the estimation tasks in a past session and that they would be paid according to the first estimate of the selected subject. (In the additional control treatment we conducted, instructions were identical to the principal treatment, except the parts describing which kind of information subjects would receive (absolute deviations between estimates vs all four estimates).)

General Instructions

You are participating in a decision-making experiment. You receive a show-up fee. Depending on your decisions you can earn additional money.

You will receive your earnings from the experiment in the experimenter's office on:

Monday, XX.XX.XX, 9am-12pm and 2pm-4pm

(Note: the day varied between sessions)

In case you are not able to come at that day, you can arrange a different time with the experimenter afterwards. You receive additional information on how and when to collect your earnings at the end of the experiment.

Please read the following explanations carefully.

During the experiment you are not allowed to talk to other participants. In case you violate this rule you are excluded from the experiment and lose all your payments. If you have a question at any time, raise your hand and the experimenter will come to your place to answer it.

As you could see, one group of participants of this experiment has been seated in the computer room of the BonnEconLab, while another group has been seated in the room next to the computer room. From now on we call participants in the computer room participants “A”, and participants in the other room participants “B”. This part of the instructions (General Instructions) is identical for all participants. Later you will receive more specific instructions, depending on your role in the experiment.

Participants A in this experiment have to solve an estimation task. We selected two texts in which a certain letter appeared exactly the same number of times. The task of participants A is to estimate how many times the letter appears in each text. (On the day you receive your payments you can verify that the letter really appears the same number of times in both texts.)

Each participant A will see one of the two texts for 60 seconds. Then he or she has to give an estimate. Afterwards the other text is shown for 60 seconds. As already mentioned, the letter to be estimated appears the same number of times in both texts. Then he or she has to estimate again.

The earnings of participants A depend on the precision of their two estimates. Thus, for both estimates the following holds: the closer the estimate is to the true value, the higher are the earnings.

The task of participants A is to estimate how many times the letter "e" appears. Earnings increase in the precision of estimates. The rule for earnings of participants A is the following: Maximum earnings for each estimate are 5 Euro. 5 Euro are received if the estimate deviates less than 1 percent (in both directions) from the correct number of times the letter "e" appears in the text. For every percentage point the estimate deviates from the true value (in both directions), 10 Cents are deducted. This means, if the estimate deviates X percent from the true value, X*10 Cents are deducted from the earnings. If the estimate deviates more than 50 percent from the true value, earnings from the estimate are 0 Euro. Negative earnings from the estimation task are not possible.

After participants A have given their two estimates, two participants A are randomly assigned to one participant B. The task of the participant B is to select one of the two participants A for a third estimation task. The estimation task will again be to estimate how many times a certain letter appears in a text. For the third estimate both, the participant B and the participant A will be paid according to precision of the estimate. Participants A that are not selected receive no additional earnings from the experiment.

Therefore, participants A have incentives to be selected, since this allows them to increase their earnings from the experiment. Participants B have incentives to select the participant A they think is most able to solve the estimation task.

Note that participants B do not know the true estimation result. Therefore participants B cannot select participants A on the basis of how close their estimates were to the true value. Participants B know that the number of times the letter appears is the same in both texts. In addition, participants B receive information about the absolute difference between the first and the second estimate for both participants A assigned to them. Each participant B will know the distance between the two estimates of the first participant A assigned to him or her and the distance between the two estimates of the second participant A assigned to him or her.

After participants A have given both their estimates, each participant B learns the absolute difference between the first and the second estimate for both participants A assigned to him or her. Then each participant B selects one participant A. The participant B and the participant A selected by him or her can increase their earnings depending on how well the selected participant A solves the third estimation task. (All participants A need to give a third estimate, but only those who were selected by a participant B receive additional earnings from their third estimate.)

Do you have any questions?

Shortly we will distribute additional instructions.

Instructions *(for agents)*

You receive a show-up fee of 4 Euro.

In this experiment you are in the role of a participant A.

In this experiment you have to solve an estimation task. We selected two texts in which the letter "e" appears exactly the same number of times. Your task is to estimate the number of times the letter "e" appears.

First you will see one of the two texts for exactly 60 seconds on your computer screen. For this text you have to estimate how often the letter "e" appears. After the 60 seconds there will appear a new screen on which you have to enter your estimate within 60 seconds. After a very short break you will see the other text for 60 seconds. Note again that this text is selected such that the number of times the letter "e" appears is identical to the text before. Now you have to estimate again. Again there will appear a new screen where you have to enter your estimate within 60 seconds. Thus, in total you will give two estimates.

Your earnings from this experiment depend on the precision of your two estimates.

For both estimates the following holds: the closer your estimate is to the true value, the higher are your earnings.

For both estimates your earnings are determined according to the following rule. Maximum earnings per estimate are 5 Euro, which you receive if your estimate deviates less than 1 percent (in both directions) from the true value. For every percentage point your estimate deviates from the true value (in both directions) your earnings are reduced by 10 Cent. If your estimate deviates X percent from the true value your earnings are reduced by X*10 Cent. If your estimate deviates more than 50 percent from the true value your earnings from this estimate will be 0 Euro. Negative earnings from the estimation tasks are not possible. Deviations are always rounded to the next whole-number percentage value.

Examples: If for instance your estimate deviates 10 percent from the true value, 10*10 Cent are deducted from the maximum earnings of 5 Euro for this estimate. Thus your earnings from this estimate are 4 Euro. If your estimate deviates 6 percent from the true value, 6*10 Cent are deducted and you receive 4.40 Euro from this estimate. Please also consider the following two examples. If your estimate deviates 24 percent from the true value, 24*10 Cent are deducted and you receive 2.60 Euro from this estimate. If your estimate deviates 43 percent from the true value, 43*10 Cent are deducted and you receive 0.70 Euro from this estimate. All earnings are in addition to the show-up fee.

Please note again that you will be paid for both estimates and that the two texts are selected such that the number of times the letter “e” appears is the same in both texts.

As already mentioned there are other participants in the room next to you. We call them participants B. After you gave your two estimates you will be randomly, together with another participant A, assigned to a participant B. The participant B will receive information about your two estimates and those of the other participant A. Participant B has to choose between you and the other participant A. If the participant B selects you, you will have to do a third estimation task and you will receive additional 10 Euro. Furthermore, you and the participant B will receive earnings depending on your precision in the third estimation task. In case you are not selected by the participant B you will not receive any additional earnings from the experiment, but you will nonetheless face the third estimation task. Since the participant B is also paid according to your precision in the third estimation task, he or she has incentives to select the participant A he or she thinks is most able to solve the task.

As mentioned before, you receive 10 Euro if you are selected by participant B. You receive these 10 Euro independent of your performance in the third estimation task. In addition you receive earnings dependent on the precision of your third estimate. These earnings are determined exactly in the same way as for the first two estimation tasks. In case you are not selected by the participant B, you do not receive any additional earnings. Thus, if you are selected by participant B, your earnings from the experiment will increase significantly.

Note that the participant B does not know the correct number of times the letter to be estimated appears in the text. Thus, the participant B cannot select you on the basis of how close your estimates were to the true value. He or she does know that the true number of times the letter appears is the same in both texts. In addition, the participant B learns the absolute difference between your first and your second estimate and the absolute difference between the first and the second estimate of the other participant A before he or she makes his or her choice. Thus, participant B will know the distance between your two estimates and the distance between the two estimates of the other participant A assigned to him or her.

After you and the other participant have given both estimates, the participant B will receive information about the distance between your two estimates and the distance between the two estimates of the other participant.

To make it easier for you to imagine the situation the participants B are in, you now see two examples of which information participants B might receive.

Participant cabin number: X	Participant cabin number: Y
Difference between his/her estimates: 5	Difference between his/her estimates: 35

Which of the two participants do you select? “I select participant ___ ”

Participant cabin number: X	Participant cabin number: Y
Difference between his/her estimates: 77	Difference between his/her estimates: 13

Which of the two participants do you select? “I select participant ___ ”

Do you have any questions?

Instructions *(For principals)*

You receive a show-up fee of 8 Euro.

In this experiment you are in the role of a participant B.

Two participants from the other room (participants A) will be randomly assigned to you. Both have to solve two estimation tasks. They need to estimate how many times a certain letter appears in a text. We selected two texts in which the letter appears exactly the same number of times.

Each participant A gets to see one of the two texts for 60 seconds, then he or she has to estimate how many times the letter appears in the text. Afterwards he or she gets to see the other text for 60 seconds. In this text, as already mentioned, the letter appears exactly the same number of times as in the text before. Then he or she has to estimate again.

Your task is to select one of the two participants A for an additional estimation task. Your earnings from this experiment depend on how precisely the selected participant estimates in this estimation task. Thus you should select the participant who you think is most able to solve the estimation task.

Your earnings are determined as follows: Maximum earnings are 10 Euro, which you receive if the estimate of the participant A you selected deviates less than 1 percent (in both directions) from the true value. For every percentage point the estimate deviates from the true value (in both directions), your earnings are reduced by 20 Cent, thus if the estimate deviates X percent from the true value your earnings are reduced by X*20 Cent. If the estimate deviates more than 50 percent from the true value your earnings from this estimate will be 0 Euro. Negative earnings from the estimation tasks are not possible. Deviations are always rounded to the next whole-number percentage value.

Examples: If for instance the estimate of the participant A you selected deviates 10 percent from the true value, 10*20 Cent are deducted from the maximum earnings of 10 Euro, thus your earnings are 8 Euro. If the estimate deviates 6 percent from the true value, 6*20 Cent are deducted and you receive 8.80 Euro. Please also consider the following two examples. If the estimate of the participant you selected deviates 24 percent from the true value, 24*20 Cent are deducted and you receive 5.20 Euro. If the estimate deviates 43 percent from the true value, 43*20 Cent are deducted and you receive 1.40 Euro. All earnings are in addition to the show-up fee.

The participant you selected can also increase his or her earnings the more precise the estimate is, analogous to the first two estimation tasks. In addition he or she receives a flat payment for being selected.

For your choice you will not know the true number of times the letter appears. Thus you cannot make your choice according to how close the estimates of the two participants assigned to you were to the true value. You know that the correct number is the same in both texts. In addition, you learn the absolute difference between the first and the second estimate for both participants A assigned to you before you make your choice. Thus, you will know the distance between the two estimates of the first participant A assigned to you, and the distance between the two estimates of the second participant A assigned to you.

Now you see two examples of which information about the two participants you might receive.

Participant cabin number: X	Participant cabin number: Y
Difference between his/her estimates: 5	Difference between his/her estimates: 35

Which of the two participants do you select? “I select participant ___ ”

Participant cabin number: X	Participant cabin number: Y
Difference between his/her estimates: 77	Difference between his/her estimates: 13

Which of the two participants do you select? “I select participant ___ ”

Do you have any questions?

Instructions - Experiment 2: The Role of Commitment

Translated into English - Instructions were identical in both treatments. After the instructions, when the bowl with peas was shown, the experimenter told subjects to write down their estimate on their answer sheet. In the main treatment the experimenter also told subjects to raise their hand once they had written down an estimate. (In the additional control treatment instructions were identical as well. In that treatment the experimenter also told subjects to raise their hand once they had decided on an estimate.)

Instructions

You are participating in a decision-making experiment. You receive a show-up fee of 5 Euro. Depending on your decisions you can earn additional money.

You receive your earnings from the experiment at the experimenter's office on:

Monday, XX.XX.XX, 9am-12pm and 2pm-5:30pm

In case you are not able to come at that day, you can arrange a different time with the experimenter afterwards. You receive additional information on how to collect your earnings at the end of the experiment.

Please read the following explanations carefully.

During the experiment you are not allowed to talk to other participants. In case you violate this rule you are excluded from the experiment and lose all your payments. If you have a question at any time, raise your hand and the experimenter will come to your cabin to answer it.

Your task in this experiment is to estimate the number of peas in a bowl. After these instructions we will show a bowl with peas, which will be removed after approximately 8 minutes. The bowl will be placed on a table in the middle of the lab.

Your earnings from this experiment depend on how precisely you estimate. Thus, the closer your estimate is to the true value, the higher are your earnings. If your estimate is for example 5 percent (or less) higher or lower than the true number of peas, you receive 10 Euro. If your estimate is more than 5 percent but less than 10 percent higher or lower than the true value, you receive 9.50 Euro. The further away your estimate is from the true value, the lower are your earnings.

The following table shows exactly how your earnings depend on the precision of your estimate.

Deviation of your estimate from the true value of peas up- or downwards in %	**Your Payment**
0% up to 5%	10.00
More than 5% up to 10%	9.50
More than 10% up to 15%	9.00
More than 15% up to 20%	8.50
More than 20% up to 25%	8.00
More than 25% up to 30%	7.50
More than 30% up to 35%	7.00
More than 35% up to 40%	6.50
More than 40% up to 45%	6.00
More than 45% up to 50%	5.50
More than 50% up to 55%	5.00
More than 55% up to 60%	4.50
More than 60% up to 65%	4.00
More than 65% up to 70%	3.50
More than 70% up to 75%	3.00
More than 75% up to 80%	2.50
More than 80% up to 85%	2.00
More than 85% up to 90%	1.50
More than 90% up to 95%	1.00
More than 95% up to 100%	0.50
More than 100%	0

If your estimate is for example 23 percent higher than the true value, you receive 8 Euro. If your estimate is for example 36 percent below the true value, you receive 6.50 Euro. All payments are in addition to your show-up fee.

Note again that you are not allowed to talk to other participants during the experiment. In case you violate this rule you are excluded from the experiment and lose all your payments.

Do you have any questions?

Eidesstattliche Erklärung

Hiermit versichere ich an Eides Statt, dass ich die vorliegende Dissertation ohne unerlaubte fremde Hilfe angefertigt und andere als die in der Dissertation angegebenen Hilfsmittel nicht benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen sind, habe ich als solche kenntlich gemacht.

Bonn, den 08.07.13