The Motives of Forfeiting Money: Experimental Studies in Behavioral Economics

Der Fakultät für Wirtschaftswissenschaften der Universität Paderborn

zur Erlangung des akademischen Grades Doktor der Wirtschaftswissenschaften - Doctor rerum politicarum -

> vorgelegte Dissertation von

Silvia Lübbecke, M.Sc.

geboren am 26.03.1984 in Detmold

März 2018

"Those who love money will never have enough. How meaningless to think that wealth brings true happiness."

Ecclesiastes 5:10

Contents

In	trodu	ction	2				
	Trad	ing monetary for psychological benefits	3				
	The role of biased beliefs						
	Grou	p membership and other-regarding preferences	7				
	Sum	mary of the results	8				
	Refe	rences	9				
1	Don	i't Patronize Me!					
	An H	Experiment on Rejecting Paternalistic Help	12				
	1	Introduction	13				
	2	Experimental Design	16				
	3	Predictions	20				
	4	Procedure and Descriptive Statistics	24				
	5	Results and Discussion	27				
		5.1 The Results	27				
		5.2 Discussion	32				
	6	Conclusion	33				
		References	33				
		Appendix	36				
2	Whe	en Supervisors Start to Meddle:					
	An H	Experiment on the Determinants of Intervention	45				
	1	Introduction	46				
	2	Experimental Design	49				
	3	Hypotheses	55				
		3.1 The influence of incidental affects on the intervention rate	55				
		3.2 The role of anticipated regret on the intervention rate	59				
	4	Procedure and Descriptive Statistics	62				
	5	Results	64				
		5.1 The influence of incidental affects on the intervention rate	64				
		5.2 Anticipating regret: the effect of feedback on the interven-					
		tion rate	67				
	6	6 Discussion					
	7	Conclusion					

		References				
		Appendix				
3	Beli	ieving in Others' Dishonesty:				
	An	Experimental Study on Beliefs about Lying 82				
	1	Introduction				
	2	Experimental Design				
	3	Hypotheses				
		3.1 Theoretical Considerations				
		3.2 Hypotheses				
	4	Procedure and Descriptive Statistics				
	5	Results				
		5.1 Belief Elicitation in In- and Out-groups				
		5.2 Belief Elicitation based on Individual Lying Behavior 104				
		5.3 Further Indications				
	6	Discussion				
		6.1 The absence of discriminating beliefs				
		6.2 The interrelation of beliefs and behavior				
	7	Conclusion				
		References				
		Appendix				
4	4 Cheating for My or for Your Benefit?					
	A F	ield Experiment with Children 123				
	1	Introduction				
	2	Experimental Design				
	3	Descriptive Statistics				
	4	Results and Discussion				
	5	Conclusion				
		References				
		Appendix				
Aj	opene	dix i				
-	Erk	lärung zu Studien in Koautorenschaft				
	Eidesstattliche Erklärung					

Introduction

Since the 1970s, economists have been acknowledging that subjects behave differently from what the neoclassical theory would consider rational. In an endless number of experiments, it has been observed that people consciously deviate from payoff-maximizing decisions.¹ These observations have led to the new interdisciplinary research domain, Behavioral Economics, which merges insights from the field of psychology with economic principles. This domain aims at documenting and predicting human behavior that does not follow the principle of a homo oeconomicus. This dissertation considers established behavioral economic concepts and tests their implications in different contexts: it investigates to what extent behavioral concepts explain why subjects do not make payoff-maximizing choices in the context of rejecting paternalistic help, the willingness to intervene in teams, discrimination, mistrust and cheating. In particular, three behavioral aspects that explain why people forgo money are discussed: the trade-off between monetary and psychological benefits, the influence of biased beliefs, and other-regarding preferences.

In the following chapters, four studies are presented to shed light on these issues:

- 1. "Don't Patronize Me! An Experiment on the Motives behind Rejecting Paternalistic Help" (with Wendelin Schnedler)
- 2. "When Supervisors Start to Meddle: An Experiment on the Determinants of Intervention" (with Wendelin Schnedler)
- 3. "Believing in Others' Dishonesty: An Experimental Study on Beliefs about Lying"
- 4. "Cheating for My or for Your Benefit? A Field Experiment with Children" (with Julia Kramer and Nina Stephan)

The aspect of trading monetary for psychological benefits is discussed in the first two studies. The influence of biased beliefs is examined in the second and third study. The third and fourth study provide input to the theory of other-regarding preferences by suggesting that subjects are more other-regarding towards their

¹(e.g. Gneezy, 2005; Fehr and Gächter, 2000; Camerer and Thaler, 1995; Ledyard, 1995; Roth, 1995; Isaac and Walker, 1991).

peers than towards strangers.

The remainder of the introduction briefly discusses each of the previously named aspects in turn, followed by a description on how they are examined in the respective studies presented in this dissertation.

Trading monetary for psychological benefits

The deviation from the payoff-maximizing option is often a deliberate and conscious decision. Most of our decisions do not only have monetary consequences, but may also affect our psychology. What is payoff-maximizing is not necessarily also psychologically beneficial. Under this condition, we have to trade monetary for psychological benefits which can cause even the most rational agents among us to forfeit money. The pursuit of self-esteem (Bénabou and Tirole, 2002; Ryan and Deci, 2000) and image concerns (Andreoni and Bernheim, 2009) as well as avoiding that we might later regret our decision (Loomes and Sugden, 1982; Bell, 1982) are examples of psychological benefits that are examined in the first and second study.

In "Don't Patronize Me! An Experiment on the Motives behind Rejecting Paternalistic Help", we study why people reject paternalistic help, in a situation where rejection is costly and does not lead to a higher monetary outcome. Consider a father who has just fixed cereals for her daughter. Instead of being grateful, she angrily rejects the cereals and insists on fixing another portion of the exactly same cereals for herself. This child does not seek economic benefits, as she incurs the costs of doing the work herself without improving the outcome. We test whether such an opposing behavior is also found in adults and examine whether it is motivated by the pursuit of essential psychological needs like preserving self-esteem² or inducing a favorable image³. First, rejecting help can serve as the attempt to avoid a feeling of incapacity and worthlessness, especially, since paternalistic help is often perceived as intimidating by the helped person. In this sense, deciding for oneself is conducive to a person's self-esteem. Second, people may reject help in order to influence the image that others have of them. A first aspect is that subjects reject help in order to induce the image of being an autonomous person who decides for herself. This motive is among others the expression that she con-

²The need for self-esteem is also suggested by the model of self-deception developed by Bénabou and Tirole, 2002.

³Andreoni and Bernheim (2009) provide evidence that people are governed by image concerns in their behavior.

siders herself worthy enough to decide on her own fate and that she demands her autonomy to be honored by others. A second aspect is the wish to be recognized as a smart individual. A person, who shows that she has the cognitive competence to oversee her own doings and to take appropriate decisions, can suitably induce the image that she is a smart individual. Our results show that psychological benefits serve as a suitable motivation to reject help, even though this behavior is costly. However, we cannot conclude whether individuals are driven by the pursuit of self-esteem, signaling their autonomy, or signaling their cognitive competence. Yet, a sizable fraction of subjects responds to each motive: arguably, every motive seems to be economically relevant. Additional research is clearly needed to statistically identify the motivational force of each motive investigated in this study.

Another psychological reason that induces people not to take payoff-maximizing choices is the anticipation of regret. Loomes and Sugden (1982) and Bell (1982) designed models allowing people to incorporate the anticipation of regret into the utility of an option, which reduces its attractiveness. Evidence for regret aversion is provided by Zeelenberg et al. (1996). In "When Supervisors Start to Meddle: An Experiment on the Determinants of Intervention", we test this idea in the context where subjects can increase their expected payoff by taking risk-eliminating, though costly, intervention. Imagine the situation where a supervisor is hired by the management of a company in order to increase the productivity of a team. When she assesses the decisions of subordinated workers as too risky, she has the authority to intervene with the workers' decisions and induce an alternative outcome which generates a higher expected profit. In our experiment, we observe that a substantial share of supervisors prefers to stay inactive. We study whether this behavior is due to the anticipation of regret. Two types of regret triggers are designed. First, supervisors receive a post-decisional feedback on the efficiency of their intervention. This feedback can be thought of as an ex-post evaluation of the supervisor's decision by the upper management. If supervisors learn that their intervention has not led to a higher monetary outcome, they may regret having spent resources on an unnecessary intervention, and thus retrospectively would have preferred to stay inactive. As a second regret trigger, supervisors receive ex-post feedback on whether their subordinated workers have disapproved their intervention. Facing social disapproval is psychologically costly which makes intervention less attractive, even though it might be monetarily optimal. Our results show that the anticipation of both types of regret leads to a sizable, however statistically not significant, reduction in intervention. This is consistent with the idea that supervisors are influenced not only by monetary outcomes but also by psychological factors—like the anticipation of regret. However, further research is needed to confirm this idea.

The role of biased beliefs

In many situations, people aim at maximizing their monetary benefits, but still fail to do so. This is mostly the case when they are not perfectly informed. When people take decisions under uncertainty, they are forced to rely on beliefs. These beliefs are not necessarily correct and all too often biased. This prevents them from making correct assessments and thus leads to monetarily suboptimal decisions. Biased beliefs often arise because people find it difficult to differentiate between relevant and irrelevant information (Payne et al., 2010). As a result, they may mistakenly consider information that is actually uninformative for their decision as relevant and thus form irrational beliefs. The second and the third study examines whether subjects take inefficient decisions because of drawing spurious inference from irrelevant information.

Beside the role of anticipated regret in intervention decisions, the study "When Supervisors Start to Meddle: An Experiment on the Determinants of Intervention" studies how incidental moods induce supervisors to remain inactive. Consider a supervisor who has just suffered from a depressing incidence in her private life. Back at work, she intervenes with her workers' decisions in order to regain the feeling of being in control. Since incidental moods-like the depressiveness arising from private incidences—are uninformative, because irrelevant, for the intervention decision at work, a rational supervisor should simply ignore them. In our experiment, students act in the role of supervisors and face mood-manipulating treatments. In these treatments, they receive either positively or negatively framed information about the challenges of entering labor market after their graduation. Having received the mood-manipulation, every supervisor interacts with another participant, the worker. While the worker is given the task to solve a logical puzzle and to submit her solution, the supervisor is given an observer role with the opportunity to intervene. Through intervention, supervisors can eliminate the possibility that an incorrect solution is submitted by their worker. Since payoffs for both the supervisor and the worker crucially depend on whether the correct solution is submitted at the end of the experiment, it is rational for the supervisors to intervene given that they sufficiently doubt their worker's ability to solve the puzzle correctly. The information about their own career prospects in the labor market should not alter their beliefs about the worker's ability to perform successfully in the puzzle task, since the two domains are completely unrelated. Nevertheless, they may draw spurious inference from their own career prospects. A subject who became discouraged about her own success in the labor market, may doubt her abilities in general. If she uses her own ability as a reference point, she may also alter her beliefs on whether others perform successfully in a given case. In consequence, a

discouraged supervisor may be more pessimistic regarding her worker's ability to solve the puzzle correctly. Our results show that supervisors in the negative mood treatment do not differ from those in the positive mood treatment with respect to their beliefs about the worker's ability to successfully solve the logical puzzle task. This suggests that incidental moods (triggered by career prospects) do not induce our supervisors to form irrational beliefs. Nevertheless, we observe that supervisors in the bad mood treatment intervene less (but not significantly) than those in the positive mood treatment. A possible explanation for this observation could be that, rather than holding irrational beliefs, supervisors may intervene less because incidental moods cause them to suffer a loss in motivation.

The third study "Believing in Others' Dishonesty: An Experimental Study on Beliefs about Lying" investigates whether beliefs cause subjects to discriminate against their out-group in a setting where the group identity is completely arbitrary and thus uninformative. In particular, the study examines where subjects in so-called minimal groups (groups that are artificially induced by an arbitrary and thus irrelevant group distinction) expect their out-group to be less honest than their in-group. For individuals, it is irrational to hold such discriminating beliefs if they are solely based on minimal group distinction, since the arbitrary group identity does not allow drawing inferences regarding the dishonesty of a group from a rational point of view. However, other experimental studies find that subjects discriminate even in minimal group settings. In contrast, our results show that subjects do not hold discriminating beliefs under arbitrary group distinction. This would suggest that, if subjects discriminate against minimal out-groups, they may be guided by a preference for discrimination, but not because they hold unfavorable beliefs against the out-group. Put differently, subjects may simply prefer to favor members of their in-group over those of their out-group.

In summary, it is possible that subjects draw irrational beliefs because they mistakenly consider irrelevant information as relevant. However, our results from the second and third study suggest that rather than affecting beliefs, irrelevant information may be more likely to influence subjects' preferences, and thus induce them to deviate from the payoff-maximizing choices.

That people draw spurious conclusion from irrelevant information is one explanation for the existence of biased beliefs. In this case, people do not reason rationally. Nevertheless, even when people behave largely rational in the process of reasoning, they may still end up holding biased beliefs. In Tversky and Kahneman (1974), some heuristics are described that lead to biased judgments. One is insufficient adjustment from an anchor information. In *"Believing in Others' Dishonesty. An Experimental Study on Beliefs about Lying"*, a second research question is addressed: whether people's own honesty behavior is in line with their beliefs about others' honesty behavior. Our results show that subjects who behave dishonestly also expect more dishonesty from others compared to those subjects who behave honestly. In the absence of complete information, it is perfectly rational to use one's own behavior as an anchor. Knowing that not all subjects may behave alike, subjects have to adjust their beliefs from this anchor. The inconsistency between beliefs of honest and dishonest subjects suggests that at least one of these groups of subjects—if not both—tends to make insufficient adjustments (Tversky and Kahneman, 1974) leading to biased beliefs and thus suboptimal decisions.

Group membership and other-regarding preferences

A self-interested person is defined as an individual who exclusively cares for her own profit without regarding how her choice affects the payoff of others. At least since Fehr and Schmidt (1999), economists have acknowledged that people are also other-regarding: individuals do not only behave self-interested, but also care about unequal distribution of payoffs. In contrast to purely self-regarding preferences, these preferences lead to decisions which are utility-maximizing but not necessarily money-maximizing. Other distributional motives are social-welfare preferences such as efficiency and maximin concerns (e.g. Andreoni and Miller, 2002; Charness and Rabin, 2000). While the classical theoretical models, like the Fehr-Schmidt-model, regard other-regarding preferences as subject-specific but context-independent, more recent models acknowledge that subjects adjust their preferences to social contexts. For example, Frohlich et al. (2004) provide an extension of the Fehr-Schmidt-model that incorporates the social norm of just deserts.⁴ Ruffle (1998) also finds that subjects alternate their other-regardedness according to the effort that the recipient exerts. The third and fourth study contribute to the growing body of literature that investigates the influence of social distance on other-regarding preferences. The existing literature provides evidence that subjects' willingness to give to others decreases in the social distance to the recipient. These studies define social distance either as "the degree of reciprocity that subjects believe to exist within a social interaction" (Hoffman et al., 1996 (p.654)) or as the identifiability of the recipient (Eckel and Grossman, 1996; Bohnet and Frey, 1999; Small and Loewenstein, 2003). In this dissertation, social distance is measured in terms of group membership: the social distance between two individuals is defined

⁴In their extended model, they account for the production costs that are invested by paired individuals who engage in a joint production. Additionally, they provide evidence for their extended model by running an experiment: Their results match the predictions of their extended model but not the predictions made by the standard Fehr-Schmidt-model.

to be larger when they do not belong to the same social group. The third and fourth study address whether subjects' other-regardedness is higher for in-group members than for members of the out-group.

The results from "Believing in Others' Dishonesty. An Experimental Study on Beliefs about Lying" suggest that discrimination observed in minimal group setting does not arise from discriminating beliefs. In reverse, if subjects discriminate in minimal groups, we may argue that they possibly act out of the desire to favor their in-group, respectively to harm their out-group. Modeling Fehr-Schmidt preferences as a function of group-identity can explain this behavior: subjects display more generosity towards their in-group members and/or more envy towards their out-group members.

The study "Cheating for My or for Your Benefit? A Field Experiment with Children" investigates whether such preferences already exist in young children. In the experiment, children are given the chance to cheat and win a prize either for themselves (control group) or for another child (treatment group). Inter alia, we test whether children in the treatment group who play for a friend cheat more often than those who play for a stranger. We find that children seem to be more gift-giving towards friends and siblings than they are towards strangers. This would suggest that even children possess other-regarding preferences which are influenced by the social distance to others. Further, some children who go empty-handed in the treatment group may be cheating to the disadvantage of other children in order to avoid that those obtain the prize. This would coincide with the idea that children already behave inequality averse. However, further research is needed to confirm this idea.

Summary of the results

First, our results show that people are willing to forfeit monetary benefits in order to pursue psychological benefits. In the context of rejecting paternalistic help, we find that psychological benefits serve as a suitable motivation. However, we cannot identify with statistical significance which psychological needs (preserving self-esteem or signaling autonomy or cognitive competence) motivate people to reject paternalistic help. Regarding the intervention in teams, we observe that supervisors are less willing to intervene with their worker's decision, when the supervisors anticipate that they may ex-post regret their intervention. This is in line with the regret theory of Loomes and Sugden (1982) and Bell (1982). Though, our results fall short of statistical significance, they are consistent with the idea that people trade monetary benefits for psychological benefits—such as avoiding regret.

Second, subjects may tend to take monetarily suboptimal decisions because of biased beliefs. On the one hand, such biased beliefs may form because subjects draw spurious inference from irrelevant information. While this might be possible, we do not observe the existence of such irrational beliefs in the contexts of intervening in teams and discriminating against minimal out-groups. Rather, it seems more likely that irrelevant information may influence subjects' preferences, inducing them to forgo monetary benefits. On the other hand, under incomplete information, biased beliefs may be the result of insufficient adjustments from a relevant information that serves as an anchor—as suggested by Tversky and Kahneman (1974). In line with their observations, we find that people who behave dishonestly are also more mistrusting against others compared to those who behave honestly. The inconsistency between beliefs of honest and dishonest subjects suggests that at least one of these groups holds biased beliefs, which cause them to take inefficient decisions.

Last, we have raised the question whether subjects are more other-regarding towards in-group members than towards out-group members. While other studies find evidence that subjects seem to discriminate against minimal out-groups, our results suggest that discriminating beliefs do not play a role in minimal groups. Given the absence of discriminating beliefs, we may rather conjecture that, if subjects discriminate in minimal groups, they may simply prefer to favor their in-group. This would be in line with the idea that subjects are more generous and/or less envy towards in-group members than towards out-group members. There is an indication that already young children may hold other-regarding preferences which depend on social contexts. We find that children are more gift-giving towards their friends and siblings than they are towards strangers.

References

- Andreoni, J., & Bernheim, B. D. (2009). "Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects." Econometrica, 77(5): 1607-1636.
- Andreoni, J., & Miller, J. H. (2002). "Giving according to GARP: an experimental study of rationality and altruism." Econometrica, 70(2): 737-753.
- Bell, D. (1982). "Regret in Decision Making under Uncertainty." Operations

Research, 30(5): 961-981.

- Bénabou, R., & Tirole, J. (2002). "Self-confidence and personal motivation." The Quarterly Journal of Economics, 117(3): 871-915.
- Bohnet, I., & Frey, B. S. (1999). "The sound of silence in prisoner's dilemma and dictator games." Journal of Economic Behavior & Organization, 38(1): 43-57.
- Camerer, C. F., & Thaler, R. H. (1995). "Anomalies: Ultimatums, dictators and manners." Journal of Economic Perspectives, 9(2): 209-219.
- Charness, G., & Rabin, M. (2000). "Social Preferences: Some Simple Tests and a New Model." University of California at Berkeley. (No. E00-283). Available at: https://cloudfront.escholarship.org/dist/prd/content/qt46j0d6hb/qt46j0d6hb.pdf.
- Eckel, C. C., & Grossman, P. J. (1996). "Altruism in anonymous dictator games." Games and Economic Behavior, 16(2): 181-191.
- Fehr, E., & Gächter, S. (2000). "Cooperation and punishment in public goods experiments." The American Economic Review, 90(4): 980-994.
- Fehr, E., & Schmidt, K. M. (1999). "A theory of fairness, competition, and cooperation." The Quarterly Journal of Economics, 114(3): 817-868.
- Frohlich, N., Oppenheimer, J., & Kurki, A. (2004). "Modeling other-regarding preferences and an experimental test." Public Choice, 119(1): 91-117.
- Gneezy, U. (2005). "Deception: The role of consequences." The American Economic Review, 95(1): 384-394.
- Hoffman, E., McCabe, K., & Smith, V. L. (1996). "Social distance and other-regarding behavior in dictator games." The American Economic Review, 86(3): 653-660.
- Isaac, R. M., & Walker, J. M. (1988). "Group size effects in public goods provision: The voluntary contributions mechanism." The Quarterly Journal of Economics, 103(1): 179-199.
- Ledyard, J. (1995). "Public Goods: A Survey of Experimental Research." In: Kagel, J. H., & Roth, A. E. (Eds.). Handbook of Experimental Economics (pp. 111-194). Princeton: Princeton University Press.
- Loomes, G., & Sugden, R. (1982). "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty." The Economic Journal, 92(368): 805-824.
- Payne, B. K., Hall, D. L., Cameron, C. D., & Bishara, A. J. (2010). "A process

model of affect misattribution." Personality and Social Psychology Bulletin, 36(10): 1397-1408.

- Roth, A. E. (2015). "Bargaining Experiments." In: Kagel, J. H., & Roth, A. E. (Eds.). Handbook of Experimental Economics (pp.253-348). Princeton: Princeton University Press.
- Ruffle, B. J. (1998). "More is better, but fair is fair: Tipping in dictator and ultimatum games." Games and Economic Behavior, 23(2): 247-265.
- Ryan, R. M., & Deci, E.L. (2000). "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being." American Psychologist, 55(1): 68-78.
- Small, D. A., & Loewenstein, G. (2003). "Helping a victim or helping the victim: Altruism and identifiability." Journal of Risk and Uncertainty, 26(1): 5-16.
- Tversky, A., & Kahneman, D. (1974). "Judgment under uncertainty: Heuristics and biases." Science, 185(4157): 1124-1131.
- Zeelenberg, M., Beattie, J., Van der Pligt, J., & De Vries, N. K. (1996).
 "Consequences of regret aversion: Effects of expected feedback on risky decision making." Organizational Behavior and Human Decision Processes, 65(2): 148-158.

Chapter 1

Don't Patronize Me! An Experiment on Rejecting Paternalistic Help

Silvia Lübbecke University of Paderborn

Wendelin Schnedler University of Paderborn, University of Bristol and IZA

Abstract

Children sometimes reject help that they have not asked for only to do the work themselves. Here, we study whether adults also reject such paternalistic help and distinguish between three possible reasons. The person rejecting help may want to preserve her self-esteem, signal her autonomy or signal her cognitive competence to the interfering party (paternalist). By varying the information available to the paternalist, we can isolate these three effects. If all three effects can operate, a substantial fraction rejects paternalistic help. Excluding the opportunity to signal cognitive competence or autonomy to the paternalist through rejection leads to a sizable (but not statistically significant) reduction of rejections.

Keywords: self-esteem, image concerns, autonomy, cognitive competence, paternalism, self-determination

1 Introduction

Anecdotal evidence suggests that children often reject paternalistic help that they have not asked for. Consider a father who has already fixed the cereal for his daughter. Instead of being grateful, the daughter reacts with frustration and insists on preparing the exact same cereal for herself. If there is evidence for such behavior also in adults, this has important implications for management. A paternalistic management might poison the work atmosphere, lead to unnecessary extra effort and thereby reduce productivity. Moreover, knowing the motive behind such rejections can help to find an appropriate remedy. From the recipient's perspective, help may rob her of an opportunity to show her competence. If this is the case, finding other ways for employees to express their competence may prevent unnecessary rejection. She may also see help as an interference into her autonomy and reject help to prove her independence to her employer. Or help may impinge on her self-esteem as an autonomous being, and she wants to assure herself of her autonomy.

This paper examines with an experimental design whether individuals reject paternalistic help. The design allows us to distinguish whether the rejection is driven by a desire to signal cognitive competence, the wish to signal autonomy, or the need to preserve self-esteem. We find that about a third of subjects rejects paternalistic help. (This rejection rate is significantly different from an error rate that we measure by confronting subjects with a very similar decision and counting how many of them select a strictly dominated choice.) Eliminating the opportunity to signal cognitive competence reduces the share of rejections to 20%. About 14% still reject paternalistic help if this can neither signal their cognitive competence nor their autonomy. The reductions, however, are not significant.

Relying on observational data, it is hard to provide clean evidence on the motives for rejecting help. Precisely because it looks childish, adult actors (managers, employees, etc.) have an incentive to mask their behavior by pretending that there are substantial differences between the rejected help and what they eventually did. By using a laboratory experiment, where the subjects remain anonymous, we can eliminate this incentive. Even if rejected help could be observed, disentangling the motives that underpin this observation requires a systematic variation of available information that is hard to obtain outside a controlled experiment.

In our experiment, an agent can solve a logical puzzle and submit a solution. Without knowing whether the agent's solution is correct, a paternalist has to decide whether to invest a small fee to 'help' the agent and replace her solution independently of whether it was right or wrong by the correct solution. Interests between agent and paternalist are aligned: both get the highest payoff if the solution is correct. If the agent's solution is correct, the agent's payoff is unaffected by the paternalist's 'help'. If the agent's solution is wrong, 'help' increases her payoff. Agents learn whether their solution is correct or not before they can decide to reject the offered 'help'. This ensures that the agent is aware of her competence. If the agent rejects help, her own solution is submitted.

Like in the child's example, rejecting help in our design has negative consequences on the material payoff of the agent. Still, the agent may choose to do so in order to signal to the paternalist that she was able to solve the puzzle (cognitive competence signal motive) or that she can determine the solution by herself (autonomy signal motive). Moreover, she may want to preserve her self-esteem by submitting her solution (self-esteem preservation motive).

We suspect the presence of the signal motives because subjects have already been shown to be concerned about their image in other contexts. Andreoni and Bernheim (2009) famously point out that individuals care about being perceived as fair by others even in an anonymous setting. This suggests that subjects may also care about being perceived as cognitively competent or independent. Accordingly, they might be willing to incur costs to signal their cognitive competence or autonomy.

The self-esteem preservation motive is rooted in the likewise famous idea among psychologists that individuals want to perceive themselves as autonomous (see e.g. Ryan and Deci, 2000). Bénabou and Tirole (2002, 2003) formalize this notion in economics in terms of individuals wanting to preserve their identity by self-signaling. Closely related is the notion that meaningful work can boost self-esteem (Frankl, 1992). Ariely et al. (2008) show that people suffer a lack of motivation if their work is destroyed in front of their eyes immediately after its completion. They conclude that individuals draw a value from their work being meaningful. If the agent rejects help, it may be easier for her to regard her previous effort as meaningful and thus preserve her self-esteem. All this suggests that individuals may have a desire to reject help, even if no one learns about this.

Our design separates out these motives by systematically varying the available information to the paternalist. In our **Rejection Info (RI)** treatment, the paternalist only learns whether the agent rejected 'help' or not and all three motives are potentially relevant. In our **Full Info (FI)** treatment, the paternalist also learns whether the agent's solution was correct. Moving from Rejection Info to the Full Info treatment thus eliminates the opportunity for signaling cognitive competence. Finally, in the **Correctness Info (CI)** treatment, the paternalist only learns whether the puzzle was solved correctly but not whether the agent rejected help. (The pater-

nalist cannot deduce this from his payoff either.) Since rejection is not observed by the paternalist, it cannot be used as a signal to the paternalist at all. In contrast to the Full Info treatment, where motives to reject are signaling autonomy and preserving self-esteem, preserving self-esteem is the only motive that remains in the Correctness Info treatment.

Our paper is related to the notion that people may want to maintain control or prevent others from controlling them. The motives that we examine, however, are different from those typically studied. In a series of experiments, economic agents gain from deciding themselves rather than delegating the decision (Fehr et al., 2013, Dominguez-Martinez et al., 2014, Bobadilla-Suarez et al., 2016). Neri and Rommeswinkel (2014) show that subjects are interference-averse in the sense that they prefer others not to affect their payoff. In all these experiments, subjects actually materially gain from maintaining control or preventing interference. The motives present in this literature are eliminated in our design: the agent has no material gains but only losses from rejecting help. Owens et al. (2014) as well as Bartling et al. (2014) identify that the desire to control is valued, even if it does not lead to higher payoffs. In contrast, the agent in our design cannot control her payoff after rejecting interference.

Probably, the closest paper to ours is Sloof and von Siemens (2017), where the agent also cannot control her payoff. In their experiment, subjects prefer to choose between two known tasks, although they do not know how their choice translates into the actual task assignment. Put differently, subjects are willing to pay to decide between two identical lotteries. Based on elicited beliefs, Sloof and von Siemens argue that this behavior results from an 'illusion of control'. Taking control enables subjects to make a (meaningless) choice. Subjects in our design do not even control anything meaningless: there is no further decision after help is rejected, which arguably eliminates any room to maintain an 'illusion of control'. However, subjects still can signal cognitive competence or autonomy or preserve their self-esteem, which is what we intend to study.

The remaining of the paper is organized as follows: the design of the experiment is presented in Section 2. Section 3 derives the predictions. In Section 4, the experimental procedure and descriptive statistics are described followed by the results and the discussion thereof in Section 5. Section 6 concludes.

2 Experimental Design

Paternalistic situations usually involve a paternalist (here: he) who can help an economic agent (here: she) at some costs, while the agent can reject that help but also incurs costs when doing so. The experimental design tries to capture such a situation, while maintaining a maximum of control on possible motives of the agent. We pair subjects and assign them the role of paternalist and agent. In the experiment, these roles where more neutrally framed as observer and decision maker in order to avoid demand effects purely on the basis of the word 'paternalist'. The main task lies with the agent. She is given a logical puzzle, while the paternalist only gets a short glimpse at the puzzle. The puzzle was constructed to seem rather complicated at the first glance but easily solvable.

Interests of paternalist and agent are aligned with respect to the material outcome. In the initial example, both father and daughter are interested in the cereal being eaten. We reflect this in the design by giving a bonus payment in case that a correct solution to the logical puzzle is submitted. The agent receives $4 \in$ and the paternalist $2.50 \in$, but only if the solution is correct. The agent is thus the main stake holder and cares most about the correct solution being submitted. In order to further strengthen that it is the agent's task to ensure the correct solution, we give the subject in the role of the agent an endowment of only $2 \in$, while the paternalist's endowment is $5 \in$. For the agent, the bonus payment is hence a substantial part of her overall payout. On the other hand, the paternalist is already relatively comfortably endowed. This further strengthens his role as the observer who does not necessarily have to get involved.

i. Logical puzzle

In the first stage of the experiment, the agent can contemplate the solution to a logical puzzle. Every agent can take as much time to solve the puzzle as needed. Agents are supplied with a pen and a note pad—for more on the puzzle see the appendix.

The logical puzzle is designed in such a way that it appears difficult at first sight but is actually easy to solve. We wanted a maximal number of correct answers because we are only interested in subjects who are competent but still receive help. The task also has to appear difficult because we want the paternalist to be uncertain about the agent's ability to find the correct solution. In order to strengthen the impression that the task is challenging, the paternalist has only 45



Figure 1.1: Illustration of decision process

seconds to read the explanation of the puzzle. In this time, it is virtually impossible to get to the point in the explanation that reveals the crucial clue for the solution.

The first stage ends when all agents have submitted a proposal for a solution to the puzzle.

ii. Paternalist's decision to help

Without knowing this proposal or whether it is correct, the paternalist can replace the proposal at the cost of $0.50 \in$ by the correct solution. This reflects that the father does not know whether the daughter would have been able to fix her cereal and generates a motive to signal cognitive competence. The costs of help are only borne by the paternalist; the agent does not lose out materially from being helped. This represents that the father inflicts no material harm on the child by fixing the cereal.

If the paternalist knew that the agent's solution were correct, he would also know that helping generates no benefit. Since it does entail costs, he would not help. If he knew the agent's solution to be wrong, helping would generate $2.50 \in$ at the price of $0.50 \in$ and help would be thus optimal. Since the principal does not know whether the agent's solution is correct, helping is only rational for the paternalist if he has sufficient doubts about the agent's ability to solve the task correctly. Being helped can thus indicate that the paternalist lacks confidence in the agent's cognitive competence. In other words, we create a situation in which being helped may reflect on the agent, so that the agent has a motive to respond by rejection.

iii. Agent's decision to reject help

In the second and last stage, the agent can reject the paternalist's help. In particular, the agent can incur costs of $0.10 \in$ so that her solution rather than the correct solution of the paternalist is used to determine whether she gets the bonus payment. On the other hand, accepting help does not cost anything and the agent receives the bonus.

In order to maintain more control, we want to remove any uncertainty that the subject in the role of the agent might have about her own ability. This is why we inform these subjects whether their solution was correct or not, before they have to decide on whether to reject help or not. The rejection decision thus becomes independent from risk aversion. Moreover, agents are certain about the negative

material consequences of rejecting help.

In principle, rejecting help can be motivated by the agent's desire to punish the paternalist, e.g., for not trusting in her cognitive competence. We eliminate this reason by not allowing rejection to affect the paternalist's payoff. Rejection has no effect on the paternalist's bonus payment: a paternalist who decided to help gets his payment regardless of whether the agent rejects the help or not.

Since we are interested in studying the rejection of unnecessary help, we focus the analysis on subjects who correctly solved the logical puzzle. Thus, our data set consists only of subjects in the role of the agent with the correct solution. This meant we had to ensure that we observe the rejection decision of sufficiently many of these subjects (without making the puzzle easier and hence less meaningful). We ensured that we observe subjects of interest in the following way.

First, the paternalist has to choose whether to help or not without knowing whether the agent's solution is correct. Had the paternalist known the correctness of the agent's solution before deciding to help, help would only have been received by agents with the wrong solution, and we would only have observed very few or no subjects of interest.

Second, the agent's decision is elicited using the strategy method. More precisely, the agent is asked whether she wants to reject the paternalist's help before she learns whether the paternalist actually helped. If the paternalist does not help, the agent's decision to accept or reject help is ignored. If the paternalist helps, this decision determines the agent's actual payoff in the described way. This was done to generate more observations.

A typical criticism of the strategy method is that it creates a demand effect by suggesting to the agent to act differently under different contingencies. Notice that this criticism does not apply here, because the agent only faces one contingency namely 'being helped'; subjects never have to decide or even contemplate what they would have done had the paternalist not helped. Given that the task appears difficult at first sight and the agent is aware of this, it likely that the paternalist helps and thus increases salience of the rejection decision.

Third, the paternalist's decision to help follows the agent's decision whether to put effort into solving the logical puzzle. Suppose, the paternalist could decide to help, before the agent were given the opportunity to solve the puzzle. Then, even capable agents who are being helped may be discouraged and put in little or no effort to solve the puzzle. This would have resulted in very few or no agents with the correct solution.

3 Predictions

We want to study whether people reject paternalistic help. Accordingly, a key statistic will be how many of the subjects in the role of agent reject help. The share of rejecting agents can then be compared to some suitable benchmark. First, we will establish this benchmark. Then, we introduce our first treatment and predict that subjects who have preferences over material outcomes do not reject help. Next, we draw on the psychological and economical literature and bring forward three motives why subjects might reject help. Finally, we introduce two more treatments to tease out which motive can best explain why subjects reject if they reject.

i. Error rate: a benchmark for the rejection rate

If our null hypothesis is that help is not rejected, a single rejection of help suffices to refute this null hypothesis. In other words, any rate different from zero would be significant. Even though, we aim to minimize error by giving clear instructions, control questions, and transparent screen layouts, it seems heroic to assume that all subjects always fully understand the consequences of their actions and are free from mistakes. More plausibly, some share, say γ , of subjects errs. Then, one could say that a statistically significant share of agents rejects help if this share is significantly different from the error rate γ .

But what is an adequate value for γ ? If we as authors set γ to a particular level, say 1%, we might expose ourselves to the criticism of having chosen this level after knowing the data. (You as a reader are of course free from such a suspicion and free to fix such a level for yourself, now.)

Rather than setting the value, we obtain an estimate from the data. For the second half of our sessions, we added a respective feature to our design. We give subjects the choice between two options that seemingly differ in terms of the denomination of coins in which they receive their payoff. The two payoff options are presented as two lists and are actually perfectly identical—only the ordering of entries in the list differs. The first option is for free, the second option is priced at $0.10 \in$. The costs and location on the screen are exactly identical to the choice of rejecting help (compare the screen shots in Figure 1.3 and Figure 1.4, Appendix). Monetary benefits from the costly option are the same as in the case of rejecting help: there are none. If a subject chooses the costly second option, she would get the same

payoff as under the first option minus the costs of $0.10 \in$. By examining the answers, we can thus determine how many subjects are willing to 'pay for nothing', which gives us the desired estimate of an error rate.

The estimated error rate servers as an upper bound because it is likely to be inflated for several reasons. First, the complexity of the decision is possibly larger because subjects have to take in more numbers on the decision screen than when deciding whether to reject. Second, the decision about the denomination of coins is at the end of the experiment, where subjects may be more tired and error prone. Third, while subjects were prepared for the consequences of the decision to reject help in the instructions, the decision in which coins they want their payout comes as a surprise. Finally, the measured error rate is based on all subjects including those who were not able to solve the logical puzzle and are perhaps more prone to error than those who correctly solved the puzzle.

ii. Prediction for outcome-oriented subjects

In our first treatment, the **Rejection Info treatment** (**RI treatment**), a helping paternalist learns whether the agent rejected help, while he remains ignorant about whether the agent's solution to the puzzle was correct. As will become clear later, all three motives (signaling cognitive competence, signaling autonomy and preserving self-esteem) are potentially relevant in this treatment.

In this treatment, a purely money-maximizing agent will not reject help even if her solution to the logical puzzle is correct. Doing so would lead to monetary costs without generating a monetary benefit. Standard reciprocity theories (e.g. Dufwenberg and Kirchsteiger, 2004; Charness and Rabin, 2000; Falk and Fischbacher, 1999; Rabin, 1993) also predict no rejection because the agent's rejection decision does not affect the paternalist's payoff; the agent can thus not punish or reward the paternalist even if she wanted to. An agent with fairness preferences in the sense of Fehr and Schmidt (1999) does not reject help either. Whether the agent rejects or not, the paternalist always has a higher monetary payoff. Since rejecting help entails a cost to the agent, the payoff gap between paternalist and agent only widens. An envious agent would thus not reject help. Irrespective which of these preferences over material outcomes the agent has, she will not reject.

Null Hypothesis. The rejection rate in the RI treatment is smaller or equal to the error rate.

If the agent cares about preserving self-esteem, signaling self-esteem or signaling

cognitive competence, she may reject help. Such motives are well-founded in psychology and have also made some inroads into economics.

iii. Preserving self-esteem and image concerns

According to Maslow (1943), self-esteem is an essential need of individuals. The Cognitive Evaluation Theory¹ (Deci, 1975; Deci and Ryan, 1985) is based on the idea that individuals want to perceive themselves as competent and autonomous. Bénabou and Tirole (2002, 2003) formalize these ideas in a series of economic models. Bénabou and Tirole (2002) describe a three-fold value of having a high self-confidence²: a positive evaluation of oneself makes people happier (consumption value); believing in one's competence increases the ease to convince others of one's qualities (signaling value); self-confidence increases people's motivation to undertake tasks (motivational value).

Summarizing, self-esteem is valuable to individuals, and they have an incentive to maintain it. Being offered help may reduce this value. Bénabou and Tirole (2003) state that "help offered by others may be detrimental to one's self-esteem and create a dependence" (p.492). If help is damaging to self-esteem, rejecting it may be a way to preserve self-esteem.

By rejecting help, the agent can also preserve the value of her work, which in turn may nourish her self-esteem. The agent had solved the puzzle, before the paternalist decided to interfere. Hence, interference nullifies the effort that the agent invested in finding the correct solution to the puzzle. Ariely et al. (2008) find that people suffer a loss of motivation when seeing the value of their work being destroyed. Further, Norton et al. (2012) state that working for a product increases its value for a subject. In line with this evidence, a person should prefer

¹The cognition evaluation theory (CET) is a sub-theory within the Self-determination Theory (SDT) (Deci and Ryan, 1985). The CET focuses on the role of autonomy and competence for a person's intrinsic motivation, while the SDT adds the third psychological need of relatedness for intrinsic motivation.

²Self-confidence can be regarded as the assurance in one's ability and power, while self-esteem describes a broader concept of a positive evaluation of one's self as a whole. In psychological literature, a first definition of self-esteem was introduced by James (1890). His concept is rather self-confidence based: he described self-esteem in terms of competence as the ratio of a person's achievement and her aspirations in domains that are of personal relevance. Later, self-esteem was used for a more general feeling of worthiness (e.g. Rosenberg, 1965) that exceeds a person's self-confidence. Branden (1969) adds the interrelation of competence and worthiness and argues that self-esteem is "integrated sum of self-confidence and self-respect" (p.110). Here, we use the terms self-confidence interchangeable.

a self-determined solution that he worked hard for over a solution from another source.³

Apart from self-esteem, Maslow (1943) also lists other-esteem as a fundamental need.⁴ Put differently, individuals care about how they are valued by others. In economics, Andreoni and Bernheim (2009) show that people have image concerns: they want to be perceived as fair—even in an anonymous context. Likewise, subjects are expected to care about being perceived as autonomous. An autonomous person proves to have the competence to take responsibility for herself, as she can fix decisions and foresee the consequence of her doings. Further, autonomy is also the expression of worthiness that the person assigns to herself: a person only acts autonomously, if she considers herself as worthy to determine her fate. According to Sennett (2003), denying a person control over her own live is a "peculiar lack of respect which consists of not being seen, not being accounted as full human beings" (p.13). Therefore, a rejecting agent builds up the image of being a strong, self-determined personality who does not allow others to take decisions on her behalf.

Further, rejecting help is also suitable to build up the image of being smart. Since rejecting is rather expensive for an agent who did not solve the puzzle, the agent can use rejection to signal that she was able to solve the puzzle and hence must have a certain cognitive competence at her disposal.

Summarizing, for any of these motives subjects may not just err, but intentionally reject help.

Alternative Hypothesis. In the RI treatment, the rejection rate is larger than the error rate.

iv. Treatments and motives to reject help

Our paper does not only want to examine whether subjects reject help, but also tries to disentangle why. One of the derived motives was that the agent uses rejection to signal her cognitive competence. In order to test whether this is the case, we use a

³Our design will not allow for a distinction between the desire to be autonomous and the wish to preserve one's effort. Therefore, we generally refer to these self-esteem relevant motives as the wish to preserve self-esteem.

⁴While image is defined as the impression that others have of a person, other-esteem should be seen as the evaluation thereof. For the purpose of this paper, this differentiation can be neglected, as an individual can only gain or lose from her image if it is translated into an evaluative action by others.

new treatment, the **Full information treatment (FI treatment)**. This treatment is identical to the RI treatment with the exception that we eliminate the opportunity to signal cognitive competence by rejection. We do so by directly informing the paternalist whether the agent's solution was correct or not.

Hypothesis 1. If subjects reject help to signal cognitive competence, the rejection rate in RI treatment is higher than in the FI treatment.

If the paternalist knows that the agent solved the puzzle correctly, rejection may still be used as a signal because the agent may care about being perceived as an autonomous being. At the price of $0.10 \in$, she can signal the paternalist that she is the one who has determined her bonus payment, not him—even though this bonus payment is paid in any case. Rejection cannot be used as a signal of autonomy if the paternalist does not learn whether the agent rejected help. Our third treatment, the **Correctness information treatment (CI treatment)**, is identical to the FI treatment, but the paternalist does not know whether the agent rejected his help. By design, the paternalist's payoff is independent from rejection, so he cannot infer from his payoff either whether the agent rejected. When comparing the CI with the FI treatment, the agent can thus no longer signal autonomy.

Hypothesis 2. If subjects reject help to signal autonomy, the rejection rate in FI treatment is higher than in the CI treatment.

Since the paternalist never learns whether the agent rejected help or not, the agent cannot use rejection as a signal. The CI treatment is hence stripped off any of the signaling motives that we have discussed above. What remains is the class of motives that directly have to do with the agent's own perception of her work and hence ultimately herself. She may still reject help to preserve her self-esteem.

Hypothesis 3. If subjects reject help only to preserve self-esteem, the rejection rate in CI treatment is higher than the error rate.

Table 1.1 gives an overview over the treatments and the three types of motives.

4 Procedure and Descriptive Statistics

The experiment was run in the BaER-Lab at the University of Paderborn in March and May 2016. 12 sessions were conducted with each session being devoted to one

pater	malist receives	information about rejection		
		YES	NO	
ution		Full Info treatment:	Correctness Info treatment:	
of solu		preserve self-esteem	preserve self-esteem	
less c	YES	+ signal autonomy		
orrectr				
out co		Rejection Info treatment:		
on abo		preserve self-esteem	(preserve self-esteem)	
matic	NO	+ signal autonomy		
infor		+ signal cognitive competence		
:				

Table 1.1: Motives in Treatments

of the three treatments. Every treatment was run 4 times. Six sessions were held in March 2016 and further six sessions in May 2016. The error benchmark "paying for nothing" was only elicited in the second wave. A session took between 48-61 minutes excluding the time needed to pay the participants. The experiment was computerized via z-Tree (Fischbacher, 2007). All students were recruited from the same subject pool via ORSEE (Greiner, 2015) and each subject only participated in one treatment.

In total, 244 students took part in the experiment. As subjects were matched in pairs, 122 students acted in the role of an agent. Recall that our analysis focuses on subjects who correctly solved the puzzle. Excluding all observations in which the subject failed to solve the puzzle leaves us with a total number of 92 observations: 31 subjects participated in the RI treatment, 32 in the FI treatment, 29 in the CI treatment.

During the experiment, subjects received their payoffs in ECU which were converted into euros (1 ECU = 5 cents) and paid out in cash immediately after each session, additional to a fixed show-up fee of $2.50 \in$. On average, participants received a total payoff of $8.49 \in$. Subjects were provided with a printed version of the instructions. Instructions consisted of a written part explaining the experiment in detail and a graphical part illustrating the sequence of the experiment in order

to facilitate understanding. The instructions were identical for all subjects and differed across treatments but only with respect to the treatment variations. For further information, see the appendix (Figures 1.5, 1.6 and 1.7).

Prior to the experiment, participants answered comprehension questions. We took extra time and care to ensure that subjects fully understand the consequences of their choices. Answering a comprehension question wrongly meant that the screen of the respective subject was blocked. The respective subject then had to call the experimenter for unblocking the screen. At this opportunity, the experimenter reviewed the relevant material for the answer with the subject without suggesting an answer. The experimenter then unblocked the screen and left the subject alone for the decision. If the answer was wrong again, the procedure started from the beginning. The experiment only started when all subjects had answered all questions correctly.

At the end of every session, subjects answered socio-demographic questions and questions concerning the experiment. Participants came from different study backgrounds: approximately 40% studied in the field of business administration and economics, 30% intended to get a teaching degree and the remaining participants were enrolled in different fields of study, such as engineering, cultural sciences, computer sciences and natural sciences. Overall, randomization worked well—see Table 1.2. There could be a slight concern that female participants are over-represented in the Correctness Info treatment—an issue to which we return when analyzing the results.

	Correctness Info	Full Info	Rejection Info	Total
Observations (correct solution)	29	31	32	92
Female Participants (in %)	75.86	48.39	46.88	56.52
Business Students (in %)	48.28	38.71	34.38	40.22
Payoff in €	7.97	7.88	8.129	8.49
Age	23.48	24.68	23.16	23.77

Table 1.2: Summary Statistics

		Correctness Info	Full Info	Rejection Info	Total
Help rejected?					
	YES	13.79%	19.35%	28.13%	20.65%
		(4/29)	(6/31)	(9/32)	(19/92)
	NO	86.21%	80.65%	71.87%	79.35%
		(25/29)	(25/31)	(23/32)	(73/92)

Table 1.3: Rejection Rates in Treatments

5 Results and Discussion

In this section, we study whether people are willing to incur costs for rejecting paternalistic help and whether such behavior can be traced back to self-esteem or image concerns. First, we investigate whether the rejection rate in the RI treatment significantly differs from our error benchmark rate. Second, we compare the three treatments in pairs in order to test whether signaling autonomy or signaling cognitive competence motivates subjects to reject help. Third, we test whether the rejection rate in the CI treatment differs from the error benchmark rate in order to find out whether self-esteem concerns matter. Last, a Linear Probability Model is run where we control for the imbalance of female participation and the proportion of business students across treatments.

5.1 The Results

Our analysis is based only on subjects in the role of agents who solved the logical puzzle correctly. Among these, a sizable proportion of agents reject help in all treatments (see Table 1.3 and Figure 1.2). In the RI treatment where agents can satisfy all three psychological needs (preserving self-esteem, signaling autonomy and cognitive ability), 28.13% (9 out of 32) reject paternalistic help. Subjects who failed to solve the logical puzzle give some indication that the consequences of rejection were well-understood: none of them rejected help. Assuming that subjects who were capable to solve the puzzle have a higher cognitive ability, one would expect them to have even better understood these consequences. This suggests that the rejection decision is not random.



Figure 1.2: Rejection Rates in Treatments

In order to test whether rejection is caused by randomness, we compare the rejection rate in the RI treatment to our error benchmark based on the "paying for nothing" option. The error rate is about 5.13%: 2 out of 39 subjects choose the costly payoff option which delivers no additional value. This error rate is significantly different from the rejection rate of 28.13% in the RI treatment (Fisher's exact test; p=0.009). This finding suggests that people do not reject help erroneously. Subjects seem to be driven by more than their material outcomes—be it in relative or absolute terms. The psychological needs to reject help seem to outweigh the obvious material costs.

Result 1. Subjects are willing to reject help, even though it is costly.

Next, we study whether the willingness to incur cost by rejecting help can be ascribed to any of the psychological needs studied in this paper. First, we address the desire to be recognized as smart by signaling cognitive competence to the paternalist. While subjects in the RI treatment can use rejection as a signal for cognitive competence, subjects in the otherwise identical FI treatment cannot. Our results show that subjects in the RI treatment reject help 8.78% more than those in the FI treatment. This difference is sizable and has the predicted direction. However, it falls short of statistical significance (Fisher's exact test, p=0.302).

Result 2. When a subject cannot signal her cognitive competence through rejecting help (moving from the RI treatment to the FI treatment), the rejection rate decreases by almost 9%. This decrease is not statistically significant.

In order to test whether subjects are driven by the desire to signal autonomy, we compare the rejection rate in the FI treatment to that in the CI treatment. Recall that paternalists in the FI treatment receive the information whether the agent has rejected help, while those in the CI treatment do not. Rejection can thus no longer be used to signal anything. In particular, the agent cannot signal her autonomy. (Since the paternalist receives the information about whether the solution is correct in both treatments there is no variation in terms of signaling cognitive competence). We find that the rejection rate is 5.56% lower in the CI treatment than in the FI treatment. However, the difference is not statistically significant (Fisher's exact test, p=0.410).

Result 3. When a subject cannot signal autonomy through rejecting help (moving from the FI treatment to the CI treatment), rejection rate drops by almost 6%. This decrease is not statistically significant.

Final, we test for preserving self-esteem as a motive to reject help. In the CI treatment, agents cannot send any signal to their paternalist. Thus, rejection only serves the purpose of preserving self-esteem. Even when image concerns are not relevant, a substantial proportion of 13.79% still reject help. By comparing the rejection rate in the CI treatment to our benchmark error rate of 5.13%, we can test whether the share is statistically significant. We find that the rejection and error rate do not significantly differ (Fisher's exact test, p=0.207).

Result 4. When a subject can preserve her self-esteem through rejecting help, a sizable proportion of subjects decide to reject help. The rejection rate, however, is not statistically different from the error rate.

Recall the worry that the share of females was higher in the CI treatment. If women are more likely to reject help, this may bias our result. This is why we estimate a Linear Probability Model (LPM) using the FI treatment as the baseline treatment with and without controls for gender and study field. The results are shown in Table 1.4.

We first estimate the treatment effect without inclusion of controls—see specification (1). The results confirm our findings from the non-parametric tests. The constant is highly significant, confirming our finding that subjects do not randomly reject help. The coefficients for the treatments coincide with the differences between the treatments. Subjects in the CI treatment are approx. 5.56% less likely to reject help than subjects in the FI treatment who can signal autonomy. Subjects in the RI treatment are approx. 8.77% more likely to reject help than those in the FI treatment who cannot signal cognitive competence. Significance levels are similar to those in the non-parametric tests.

Next, we estimate a second specification (2) where we control for study field and most importantly gender. In psychology, there is evidence that males have a higher desire for control than females (e.g. Burger and Cooper, 1979; Smith et al., 1988; Burger and Solano, 1994). The inclusion of the control variable *female* is warranted.

Including these controls leads only to minor changes. The gender of subjects does not seem to have impaired our findings. The coefficient for *female* is not significant. If anything, the rejection rate in the CI treatment is higher after controlling for gender: 0.3017-0.0248=27.69% instead of 0.1935-0.0556=13.79% (see Table 1.4).

For business students, we observe that they seem to have a lower tendency to reject help than students from other fields. This effect is highly significant. The

Dependent variable	1 if agent rejected help		
	(1)	(2)	
CI treatment	0556	0248	
	(.0972)	(.1044)	
FI treatment	Ref.	Ref.	
RI treatment	.0877	.0766	
	(.1083)	(.1008)	
business student		2470***	
		(.0693)	
female		0260	
		(.0919)	
constant	0.1935***	.3017***	
	(.0721)	(.0904)	
N. obs.	92	92	
R^2	0.0212	0.1091	

Table 1.4: Linear probability regression on the decision to reject help

Notes: Linear probability estimates. Robust standard errors are in parentheses. *business student* is equal 1 for a student studying in the field of Business Administration and Economics, (=0) for students in other study fields. * significant at p<0.10; ** significant at p<0.05; *** significant at p<0.01

finding suggests that business students are more outcome-oriented and perhaps less influenced by psychological motives than other students.

5.2 Discussion

Our results clearly indicate that people reject paternalistic help, even though it is costly. Indeed, almost every third agent rejects help. When it comes to the motivational forces behind the results, our results are indicative, but still leave room for interpretation.

Across treatments, rejection rates change in line with the predictions and in economically significant ways. The differences are not statistically significant, but this may be due to too few observations: in our largest treatment, we have only 31 observations. Unfortunately, the available subject pool prohibited us to further increase sample size. Indeed, for this study we exhausted the sizable subject pool of the BaER-Lab in the sense that every subject from this pool received an invitation e-mail. Given the response rate at the time of conducting the experiment, we were able to recruit 244 students in total with half of them acting in the role of the agent. The (desired) side effect that not all subjects in the role of agent were able to solve the puzzle correctly further reduced our sample size.

Apart from the economically sizable effects, there is further indication that the lack of significance is caused by too low observation numbers. While the rejection rate in the RI treatment is significantly different from the error rate, the differences between the treatments are not significant. This would logically imply that in all treatments, the rejection rate must be different from the error rate. This, however, is not what we find: in the CI treatment, we cannot reject that the rejection rate is equal to the error rate. Since we know that the effect must come from somewhere, the only logical explanation is that we do not have enough observations to identify where it originates.

When comparing the p-values related to the three motives, the difference between rejection rate in the CI treatment and the error rate is least likely to be caused by randomness. Recall that the error rate is arguably an upper bound for the true error rate among subjects who solved the puzzle correctly. All this suggests that the desire to preserve self-esteem is the most likely candidate for why subjects reject paternalistic help.

6 Conclusion

In this paper, we study whether people are willing to incur costs in order to reject help. Further, we investigate the motivational force of three psychological needs that may cause a person to reject paternalistic help: preserving self-esteem, signaling autonomy and signaling cognitive competence.

Our findings indicate that people have a positive willingness to pay in order to reject help. Almost every third subject rejects help when this allows them to preserve self-esteem or to signal their autonomy or their cognitive competence. Our data, however, does not allow us to identify a specific motive, be it self-esteem or image related, as a significant driving force. Yet, every motive seems to be economically relevant given that a sizable fraction of subjects respond to each motive. Additional research is clearly needed to further disentangle the motives.

References

- Andreoni, J., & Bernheim, B. D. (2009). "Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects." Econometrica, 77(5): 1607-1636.
- Ariely, D., Kamenica, E., & Prelec, D. (2008). "Man's search for meaning: The case of Legos." Journal of Economic Behavior & Organization, 67(3): 671-677.
- Bartling, B., Fehr, E., & Herz, H. (2014). "The intrinsic value of decision rights." Econometrica, 82(6): 2005-2039.
- Bénabou, R., & Tirole, J. (2002). "Self-confidence and personal motivation." The Quarterly Journal of Economics, 117(3): 871-915.
- Bénabou, R., & Tirole, J. (2003). "Intrinsic and extrinsic motivation." The Review of Economic Studies, 70(3): 489-520.
- Bobadilla-Suareza, S., Sunsteinb, C. R., & Sharota, T. (2016). "Are Choosers Losers?" Haward Law School. Discussion Paper No. 868. ISSN 1936- 5357. Available at: http://www.law.harvard.edu/programs/olin_center/papers/pdf/Sunstein_868.pdf
- Branden, N. (1969). "The psychology of self-esteem." New York: Bantam.
- Burger, J. M., & Cooper, H. M. (1979). "The desirability of control." Motivation and Emotion, 3(4): 381-393.
- Burger, J. M., & Solano, C. H. (1994). "Changes in desire for control over time: Gender differences in a ten-year longitudinal study." Sex Roles, 31(7): 465-472.
- Charness, G., & Rabin, M. (2000). "Social Preferences: Some Simple Tests and a New Model." Mimeo, University of California at Berkeley.
- Deci, E. L. (1975). Intrinsic motivation. New York: Plenum.
- Deci, E. L., & Ryan, R. M. (1985). "Intrinsic motivation and self-motivation in human behavior." New York: Plenum Press.
- Dominguez-Martinez, S., Sloof, R., & von Siemens, F. A. (2014). "Monitored by your friends, not your foes: Strategic ignorance and the delegation of real authority." Games and Economic Behavior, 85: 289-305.
- Dufwenberg, M., & Kirchsteiger, G. (2004). "A theory of sequential reciprocity." Games and Economic Behavior, 47(2): 268-298.
- Falk, A., & Fischbacher, U. (1999). "A Theory of Reciprocity." Institute for Empirical Research in Economics. University of Zurich, Working Paper, 6.
- Fehr, E., Herz, H., & Wilkening, T. (2013). "The lure of authority: Motivation and incentive effects of power." The American Economic Review, 103(4): 1325-1359.
- Fehr, E., & Schmidt, K. M. (1999). "A theory of fairness, competition, and cooperation." The Quarterly Journal of Economics, 114(3): 817-868.
- Fischbacher, U. (2007). "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." Experimental Economics, 10(2): 171-178.
- Frankl, V.E. (1963 [1992]). "Man's Search for Meaning." 4th edition, translated by Ilse Lasch. Boston: Beacon Press.
- Gecas, V. (1982). "The self-concept." Annual review of sociology, 8(1): 1-33.
- Greiner, B. (2015). "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE". Journal of the Economic Science Association, 1(1): 114-125.
- James, W. (1890). "The Principles of Psychology." New York: Holt.
- Maslow, A. H. (1943). "A theory of human motivation." Psychological Review, 50(4): 370-396.
- Neri, C., & Rommeswinkel, H. (2014). "Freedom, Power and Interference: An Experiment on Decision Rights". Available at SSRN: http://ssrn.com/abstract=2485107 or http://dx.doi.org/10.2139/ssrn.2485107.

- Norton, M. I., Mochon, D., & Ariely, D. (2012). "The IKEA effect: When labor leads to love." Journal of Consumer Psychology, 22: 453-460.
- Owens, D., Grossman, Z., & Fackler, R. (2014). "The control premium: A preference for payoff autonomy." American Economic Journal: Microeconomics, 6(4): 138-161.
- Rabin, M. (1993). "Incorporating Fairness into Game Theory and Economics." The American Economic Review, 83(5): 1281-1302.
- Rosenberg, M. (1965). "Society and the adolescent self-image." Princeton, N.J.: Princeton University Press.
- Ryan, R. M., & Deci, E.L. (2000). "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being." American Psychologist, 55(1): 68-78.
- Sennett, R. (2003). "Respect in a World of Inequality." New York: Norton & Company.
- Sloof, R., & von Siemens, F. A. (2017). "Illusion of Control and the Pursuit of Authority." Experimental Economics, 20(3): 556-573.
- Smith, R. A. P., Woodward, N. J., Wallston, B. S., Wallston, K. A., Rye, P., & Zylstra, M. (1988). "Health care implications of desire and expectancy for control in elderly adults." Journal of Gerontology, 43(1): 1-7.

Appendix

Logical puzzle

"Please, read through the complete task before solving it.

The dice line

Five ordinary dice are lined up. (The numbers on two opposing sides of an ordinary die sum up to seven.)

The <u>first die</u> faces the tabletop with the number 1. The <u>second</u> die has twice as many points on the upper side as the third die. The <u>third die</u> faces the tabletop with the number 4. On the upper side of the <u>fourth die</u> a number is shown that is equal to the number on the upper side of the second die reduced by the number shown on the upper side of the third die. The fifth die faces up with the number that the first die faces the tabletop.

What is the number of the fifth die facing up?"



Further Material: Screen Shots and Instructions

Figure 1.3: Screen shot of the agent's rejection decision



Figure 1.4: Screen shot of the error benchmark 'paying for nothing'

Allgemeine Erklärungen für die Teilnehmer

Sie nehmen nun an einem wirtschaftswissenschaftlichen Experiment teil. Diese Anleitung ist für alle Teilnehmer gleich, bitte lesen Sie sie genau durch. Es wird Ihnen alles erklärt, was Sie für die Teilnahme am Experiment wissen müssen. Falls Sie Fragen haben, melden Sie sich bitte per Handzeichen. Ihre Frage wird dann an Ihrem Platz beantwortet. Ansonsten gilt während des ganzen Experiments ein absolutes Kommunikationsverbot.

Für die Teilnahme am Experiment erhalten Sie ein Entgelt von 2,50 €. Im Verlauf des Experiments können Sie zusätzliche Taler dazu verdienen, die zunächst auf ein Konto eingezahlt werden. Ihr Kontostand wird Ihnen am Ende des Experiments in Euro umgerechnet ausgezahlt. Hierbei gilt, dass:

1 Taler = 5 Cent

Das umgerechnete Einkommen wird am Ende des Experiments zusammen mit dem Teilnahmeentgelt bar ausgezahlt.

Ablauf des Experiments

- Sie und ein anderer zufällig ausgewählter Teilnehmer bilden eine Gruppe. Sie erfahren aber nicht, mit wem Sie in einer Gruppe sind. In jeder Gruppe gibt es zwei Rollen: einen Entscheider (E) und einen Beobachter (B). Der Entscheider E hat anfänglich einen Kontostand von 40 Talern und der Beobachter B von 100 Talern.
- 2. Sie erfahren, ob Sie Entscheider E oder Beobachter B sind.
- 3. Der Entscheider E löst eine Aufgabe und schickt die Lösung für sich und B ab.
- 4. Bevor die abgeschickte Lösung eingeht, wird dem Beobachter sowohl die Aufgabe als auch die richtige Lösung zur Aufgabe kurz angezeigt. B erfährt aber nicht, welche Lösung der Entscheider abgeschickt hat.
 - a. Der Beobachter kann den Entscheider **gewähren lassen**. Dann geht die Lösung des Entscheiders für beide ein.
 - b. Der Beobachter kann sich auch zum Preis von 10 Talern **einmischen**. Dann wird die Lösung des Entscheiders durch die richtige Lösung ersetzt.
- 5. Beobachter und Entscheider erfahren beide, ob die von E abgeschickte Lösung richtig ist.[*FI treatment/CI treatment*] [*RI treatment: Nur der Entscheider erfährt, ob die von ihm abgeschickte Lösung richtig ist. Der Beobachter erfährt dies nicht.*]

- 6. Der Entschieder E weißzunächst nicht, ob sich B eingemischt hat, mußaber für den Fall der Einmischung eine Entscheidung treffen.
 - a. E kann die **Einmischung akzeptieren**. Dann bestimmt der Beobachter und es geht die richtige Lösung für beide ein.
 - b. E kann sich **der Einmischung** zum Preis von 2 Talern **widersetzen**. Dann bestimmt der Entscheider E für sich: Für E geht seine eigene Lösung ein und für B die richtige Lösung.

Dem Beobachter wird mitgeteilt, ob sich der Entscheider widersetzt hat. [*RI treatment/FI treatment*] [*CI treatment: Dem Beobachter wird nicht mitgeteilt, ob sich der Entscheider widersetzt hat.*]

Auszahlung des Entscheiders E

Der Entscheider erhält ein Startguthaben von **40 Talern**. Zusätzlich erhält der Entscheider **80 Taler**, wenn die richtige Lösung für ihn eingeht.

Die richtige Lösung geht für ihn in drei verschiedenen Situationen ein ...

- ... wenn sich B nicht eingemischt hat und die von E abgeschickte Lösung richtig ist,
- ... wenn sich B eingemischt hat und E sich nicht widersetzt hat oder
- ... wenn sich B eingemischt hat, E sich widersetzt hat und die von E abgeschickte Lösung richtig ist.

Im Fall, dass sich B eingemischt hat und sich E widersetzt hat werden E 2 Taler abgezogen.

Auszahlung des Beobachters B

Der Beobachter erhält ein Startguthaben von **100 Talern**. Zusätzlich erhält der Beobachter **50 Taler**, wenn die richtige Lösung für ihn eingeht.

Die richtige Lösung geht für ihn in zwei verschiedenen Situationen ein ...

- ... wenn sich B nicht eingemischt hat und die von E abgeschickte Lösung richtig war oder
- ... wenn sich B eingemischt hat.

Falls B sich einmischt werden B 10 Taler abgezogen.

Alle Entscheidungen und Auszahlungen sind in der ausgeteilten Grafik noch einmal dargestellt. [Please: see the following graphics illustrating the instructions.]

Organisatorisches nach dem Experiment

- i Sie beantworten demographische Fragen und Fragen zum Experiment.
- ii Sie warten an Ihrem Platz, bis Ihre Sitzplatznummer aufgerufen wird.
- iii Sie werden aufgerufen und erhalten Ihre Auszahlung.

Bitte beachten Sie:

- Während des gesamten Experiments ist keine Kommunikation mit anderen Teilnehmern gestattet.
- Alle Handys müssen während der kompletten Experimentdauer ausgeschaltet sein.
- Wenn Sie eine Frage haben, bleiben Sie bitte an Ihrem Platz sitzen und heben die Hand. Stellen Sie bitte Ihre Frage so, dass kein anderer Teilnehmer Ihre Frage mithören kann.
- Sämtliche Entscheidungen erfolgen anonym, d.h. keiner der anderen Teilnehmer erfährt die Identität desjenigen, der eine bestimmte Entscheidung getroffen hat.
- Auch die Auszahlung erfolgt anonym, d.h. kein Teilnehmer erfährt, wie hoch die Auszahlung eines anderen Teilnehmers ist.
- Bitte bleiben Sie bis zum Ende des Experiments an Ihrem Platz sitzen, Sie werden zur Auszahlung mittels der Ihnen zugeordneten Platznummer aufgerufen.

Viel Erfolg und Danke für Ihre Teilnahme an diesem Experiment!

Graphics illustrating the Instructions (translated)



Figure 1.5: Correctness Info Treatment



Figure 1.6: Full Info Treatment



Figure 1.7: Rejection Info Treatment

Chapter 2

When Supervisors Start to Meddle: An Experiment on the Determinants of Intervention

Silvia Lübbecke University of Paderborn

Wendelin Schnedler University of Paderborn, University of Bristol and IZA

Abstract

In large companies, supervisors are hired to control their subordinates' performance and intervene with risky decisions in order to increase productivity. However, their decision to intervene may not always be profit-orientated. This paper studies whether the decision to intervene in a worker's decision is influenced by psychological factors that are unrelated to the profitability of intervention. In particular, we examine the role of incidental moods and the anticipation of regret triggered by ex-post evaluation of the decision. Intervention behavior is analyzed in a factorial design controlling for two mood conditions (positive, negative) and the presence or absence of feedback on either the efficiency of intervention or on its social (dis)approval by the supervised worker. We observe that supervisors in the negative mood condition intervene less often (approx. 13%) than those in the positive mood condition. Further, when supervisors are later evaluated, they intervene less (approx. 16%). Our observations are consistent with the idea that supervisors' decision are not only driven by payoff but also by incidental moods and regret anticipation. The effects, however, are not statistically significant.

Keywords: intervention, incidental affects, anticipation of regret, decision under uncertainty, group decision-making

1 Introduction

The management of many large companies face challenges to efficiently communicate statues, visions and targets to their work force and to control for compliance. Companies usually act to this problem of asymmetric information by hiring supervisors: since they have a closer hand on their subordinates, they serve as a link between the upper management that take decisions and the work force that carries out these decision. An effective supervisors uses her intervention authority as a mean of productivity enhancement: inter alia, her close relationship to subordinates allows her to promptly identify risky actions of workers and to instantly take efficient corrective measures. However, often companies lack policies ensuring that supervisors do not abuse their power to satisfy personal, often psychological, needs like the desire for control. Consider a manager who played football lost an important match on the week-end. On Monday morning back at the office, still feeling depressed, he starts bullying around with his subordinates and intervene in their decisions, whether justified or not, in order to regain the feeling of being in control. Such behavior is not productivity-enhancing and undesired by both the upper management and the workers.

This paper studies whether supervisors are guided in their intervention decision by psychological motives which do not follow profitability concerns. Two psychological cal channels that may affect the decision to intervene are examined. First, we study the influence of negative incidental moods¹ which do not arise from the decision to intervene but from outside circumstances. Second, we test whether supervisors shy away from intervention when being confronted with an ex-post evaluation of their decision. Such an evaluation may induce supervisors to anticipate that they may later regret their decision and, thus, cause them to rethink an intervention. To be precise, two types of post-decisional feedback are examined: feedback on the efficiency of intervention (efficiency feedback), which can be thought of as feedback from the upper management, and feedback on the social disapproval of intervention (disapproval feedback) by the supervised worker.

Whether supervisors are influenced by negative incidental moods or the anticipation of regret in their decision to intervene is an empirical question. In this paper, we examine this question with an experiment. Our observations show that those who receive a mood-depressing treatment intervene up to 13% less than those who receive a mood-elevating treatment. Further, we observe that receiving efficiency feedback reduces the intervention rate by 15.87%. Likewise, expecting disapproval

¹In this paper, the terms moods, affects and emotions are used interchangeably, not minding the distinction made in psychological literature.

feedback from the worker leads to a reduction in intervention by 16.11%. While the observed reductions are sizable and in line with the idea that incidental moods and regret anticipation affect the supervisor's decision, however, they fall short of statistical significance.

Whether moods are truly incidental and not integral-thus not arising from the intervention decision itself²—is difficult to control and thus prohibits the use of observational data. Therefore, we designed an experiment where moods are clearly incidental, because they are constructed to be unrelated to the decision to intervene. In order to manipulate the affective state of our supervisors, they were given either negatively framed (Bad News treatment) or positively framed (Good News treatment) information about their career prospect. As subjects tend to draw spurious conclusions from information that is actually thought to be irrelevant, we control for the incidentalness of the mood manipulation: we test whether the framing of career prospects influences the supervisor's expectation that their matched worker performs successfully in the subsequent decision process where he has to solve a logical puzzle. If the expectations turn out to be independent of the mood manipulation, the information about their career prospects can be considered as incidental. Following the mood manipulation, the supervisor has to decide whether to intervene or not: through intervention, she replaces her worker's solution by the correct solution.

Not knowing whether the worker's solution is right or wrong, her intervention decision is under uncertainty. This may induce the supervisor to regret her decision when receiving feedback. Two types of feedback were given: whether the worker has been able to solve the puzzle correctly- in order words whether intervention was needed—(efficiency feedback) and whether the worker has approved of the intervention (disapproval feedback). According to Zeelenberg (1999), the reception of feedback is an antecedent of regret anticipation. By varying the feedback, we isolate the effects of anticipated regret seeking efficiency and anticipated regret arising from social disapproval. In the **Full Feedback treatment**, both type of feedback are given. In the **Rejection Feedback treatment**, supervisors receive only the disapproval feedback. In the **Efficiency Feedback treatment**, only the efficiency feedback is transferred to the supervisors.

Our study contributes to the literature on incidental affects. Recent studies provide evidence for an influence of incidental affects on risk perception (e.g. Johnson

²Contrarily to incidental affects, integral affects are usually in line with to the profitability of intervention. A supervisor who is deciding whether to intervene also uses her feelings for the intervention as a valuable source of information.

and Tversky, 1983; Lerner and Keltner, 2000, 2001) or on risk taking behavior (e.g. Yuen and Lee, 2003; Hockey et al., 2000; Mittal and Ross, 1998; Nygren et al., 1996; Leith and Baumeister, 1996; Isen, 1993; Mano, 1992). However, little attention has been granted to studying the effect of incidental affects on the willingness to intervene in teams. A recent study of Neff et al. (2014) examines the influences of positive state affects on the willingness to seek and share information in teams. Allowing for a mix of incidental - ergo team work unrelated - and integral - team work related - affects in their study, the authors find that communication, both information seeking and sharing, is positively correlated with individual positive affects. In contrast, our design clearly identifies the effect of incidental state affects on intervention activity in teams. Pfaff (2012) provide causal evidence for the influence of incidental affects and stress on team communication. He finds that being in a positive mood increases team awareness: subjects in the happy mood condition inform others about their team contribution more frequently than subjects in the negative mood condition. Rather than looking at the quantity of the exchanged information without accounting for its relevance, the study presented here examines whether negative incidental affects reduce supervisors' willing to intervene in the decision of a team member in order to eliminate uncertainty and increase the expected productivity of the team.

Additionally, this paper contributes to the literature on regret theory by studying the influence of post-decisional feedback on intervention behavior. Various experimental studies have shown that the anticipation of regret affects individual behavior. With respect to regret arising from having taken inefficient decisions, studies have been conducted in the context of risk-taking (Zeelenberg et al., 1996), consumer behavior (Inman and Zeelenberg, 2002), bidding behavior in auctions (Filiz and Ozbay, 2007), and in ultimatum games (Zeelenberg and Beattie, 1997). Other studies provide evidence that anticipating social disapproval alters behaviors in cooperative games: in voluntarily contribution games (Rege and Telle, 2004; Masclet et al., 2003; Gächter and Fehr, 1999), coordination games (Dugar, 2010), and trust games (Lumeau et al., 2015). To the best of our knowledge, the study presented in this paper is the first of examine the role of regret anticipation in the context of supervision where supervisors can intervene with a worker's decision at some costs to eliminate risks. This study provides evidence on regret arising from efficiency feedback as well as on regret arising from disapproval feedback.

The remaining of the paper is organized as follows: Section 2 describes the experimental design. In Section 3, hypotheses are derived. Section 4 describes the procedure of the experiment and descriptive statistics. The results are presented in Section 5 followed by a discussion in Section 6. Section 7 concludes.

2 Experimental Design

This paper studies whether the willingness to intervene in a worker's decision is influenced by psychological factors which are unrelated to the profitability of intervention. Two psychological factors are discussed here. First, we test whether negative incidental moods affect supervisors' inclination to intervene. To test this hypothesis, we introduced mood-manipulation treatments (called the News treatments) prior to the intervention decision. Second, this paper studies whether the anticipation of regret has influence on the intervention decision. This research question requires that supervisors are exposed to regret triggers. Ergo, subsequent to the supervisor's decision on intervention, we integrate feedback treatments where supervisors receive feedback on either the efficiency of intervention or the disapproval by the worker.

In the experiment,³ participants are matched in pairs: a supervisor and a worker engage in a decision process where the worker is in charge of finding and submitting the correct solution to a logical puzzle. First, we give a brief description of this decision process. Then, we describe the treatments in detail. Final, we emphasize some important features of the decision process that are required to elicit reliable observations on the intervention decision and to successfully implement the Feedback treatments.

i. The decision process

The experiment was conducted with students from the Paderborn University. Participants are matched in pairs with each consisting of a supervisor and a worker. Roles are assigned randomly. In the decision process the worker is given the task to solve a logical puzzle, while the supervisor has the option to intervene in the worker's solution at some costs and replace it by the correct solution. Both players have a monetary interest in submitting the correct solution: every player receives a role-specific bonus payment if the correct solution is submitted at the end of the decision process. The bonus payment for the supervisor was $2.50 \in^4$ and for the worker $4 \in$. For a wrong solution, no player receives a bonus payment. Since decisions made during the decision process generate costs, each player receives a role-specific initial endowment which strictly exceeds those costs in order to avoid

³The experiment described in this paper is embedded into another experiment which studies the behavior of the worker. For further information, see Lübbecke and Schnedler (2018).

⁴Payments derived during the experiment are denoted in ECU, and later converted in euros (1 ECU = 5 cents).

bankruptcy. The supervisor receives $5 \in$ as an initial endowment; the worker starts with an endowment of $2 \in$.

The decision process consists of three stages (see Figure 2.1). In the first stage, the worker is given a logical puzzle to solve. Her solution is sent off as a submission for both players. In the second stage, the supervisor reads the puzzle and is shown its correct solution. Not knowing the solution of the worker, the supervisor decides whether to intervene. Through intervention, the supervisor automatically replaces the worker's solution, be it right or wrong, by the correct solution. Ergo, intervention assures that the highest payoff is obtained. Intervention is priced at $0.50 \in$ for the supervisor, but imposes no costs on the worker. In the third stage, the worker fixes a reaction to the intervention of the supervisor: she can either accept or reject it. Rejection generates costs of $0.10 \in$ only for the worker. If she rejects the intervention, she disagrees with the replacement of her solution. In this case, two solutions are submitted: the correct solution for the supervisor and the worker's solution. If she accepts, she agrees to the replacement of her solution.

ii. The News treatments

The news treatment aims at manipulating the supervisors' mood before entering the decision process. Previous studies (e.g. Pfaff, 2012; Kirchsteiger et al., 2006) induce a mood manipulation by showing funny, respectively sad, videos to the subjects. In contrast, our subjects receive either depressing or encouraging information on their expected ability to integrate into the labor market after their graduation. In our eyes, this mood-manipulation procedure is more suitable to activate personal distress that triggers the desire for corrective action than confronting subjects with impersonal, though emotional, videos or movies.

Being students from the Paderborn University, the supervisors receive information about their career prospects extracted from the latest alumni study of the Paderborn University 2014. Given the importance of information for their future, we expect these information to be influential. Supervisors in the **Good News treatment** receive positively framed information, heading "Congratulations. Students with your characteristics have good chances at the labor market" enriched with concrete, positively phrased information about employ-ability and high job satisfaction. Supervisors in the **Bad News Treatment** are given negatively framed information, heading "For students with your characteristics the prospects at the labor market do not look that good" with negatively phrased information about challenges to find jobs and lower job satisfaction added below. In both treatments, the given information is correct but oppositely framed. Before receiving the information,



Figure 2.1: Game Tree with payoffs illustrating the decision process

subjects answer some demographic and study-related questions. However, in order to avoid multiple treatments, the provided information from the alumni study are extracted at the university level and do not regard other personal characteristics as the gender, the faculty or study degree. Hence, information on career prospect are uniform within the treatments. Participants who indicate not to be a student at the Paderborn University are not given any information and are excluded from the data set, since they receive no treatment. Screen shots of the News Treatments are printed in the appendix.

iii. The Feedback treatments

By implementing the Feedback treatments, we vary the scope for anticipating regret. Exposing supervisors to different combinations of regret triggering feedback allows us to test whether intervention reduces when anticipating certain type of regret.

A supervisor who has invested in intervention may later regret to have done so for various reasons. First, she may regret if she learns ex-post that her intervention was not needed and her effort spent in vain. This regret is motivated by efficiency-seeking behavior. Second, she may feel regret if she learns ex-post that others do not appreciate her intervention. This regret is triggered by experiencing social disapproval. For subjects to feel regret, they need to get respective feedback. Here, we design three feedback treatments where supervisors receive efficiency feedback, whether her worker has solved the puzzle correctly, (Efficiency Feedback (EF) treatment) or disapproval feedback, whether the worker has rejected her intervention (Rejection Feedback (RF) treatment) or both efficiency and disapproval feedback (FF) treatment). Expecting such feedback allows supervisors to anticipate potential regret in advance which may ex-ante alter their willingness to intervene.

The treatments share the same sequences (see Table 2.7) and differ only with respect to the feedback that is given to the supervisor. Table 2.1 visualizes the treatment variations. These three Feedback treatments described above were run additional to the Good News treatment, respectively Bad News treatment, which resulted in a 2x3 factorial design. Meaning, every participant received either the Good News treatment or the Bad News treatment and one of the three additional Feedback treatments.

Supervisor receives feedback about		worker's rejection decision		
		YES	NO	
correctness of worker's solution	YES	Full Feedback treatment: anticipate regret due to inefficiency social disapproval	Efficiency Feedback treatment: anticipate regret due to inefficiency	
	NO	Rejection Feedback treatment: anticipate regret due to (inefficiency*) social disapproval		

* The supervisor receives only a noisy signal on the efficiency of her decision.

Table 2.1: Anticipation of potential regret in the Feedback Treatments

iv. Important features of the decision process

In economic experiments the salience of decisions is one of the key feature. For subjects to state decisions that represent their true preferences, they have to face monetary incentives. However, subjects can only form reliable preferences if they are able to oversee the consequences of their strategies. To avoid ambiguity about the consequences of intervention, a supervisor who intervenes automatically induces that the worker's solution, be it correct or incorrect, is replaced by the correct solution. This automatic replacement is irreversible for the supervisor: A worker who later rejects the intervention can only restore her own solution for herself but not for the supervisor. Consequently, through intervention the supervisor is perfectly sure of receiving her bonus payment of $2.50 \in$. If she does not intervene, she chooses to rely on the worker's solution which can be correct or incorrect. In this case, she only receives her bonus payment if the worker's solution is correct. Given the costs for intervention, a supervisor should only intervene if she sufficiently mistrusts the worker's ability to solve the puzzle correctly.

A feedback only triggers regret if it provides unpleasant information that is not available otherwise. Without knowing the worker's solution, the decision to intervene becomes a decision under uncertainty. Ergo, the supervisor does not know ex ante whether intervention is efficient. She only learns about its efficiency when she receive the feedback on whether the worker's solution is correct or not. Likewise, supervisors can only learn whether the worker has rejected the intervention, when receiving the disapproval feedback. A supervisor cannot deduce any of these information from her payoff. Thus, the feedback enables her to anticipate regret which she would not anticipate otherwise.

For the feedback on social disapproval to be credible and to be sufficient to trigger regret, supervisors have to anticipate that rejection is likely occur. If rejection by the worker is unlikely, supervisors would also not expect any rejection and, thus, would not anticipate any regret, even though she receives respective feedback. A risk averse worker would only reject intervention if she is aware of its consequences. Otherwise, risk aversion of workers may lead to a reduction in the rejection rate. This would make the occurrence of rejection unlikely. Therefore, we informed the worker about the consequence of rejection. A worker who rejects the intervention receives the bonus of $4 \in$ only if her solution is correct. If she accepts the intervention, she always gets the bonus payment, irrespective of the correctness of her own solution. To be able to fully assess the consequences of rejection, the workers are informed whether their solution has been correct or not before taking the decision.

Last, for the rejection to trigger the feeling that the intervention is not approved, confounding motives have to be ruled out. In our experiment, a worker cannot use intervention as a mean of punishing the supervisor for mistrusting her ability to solve the puzzle, because rejection has no monetary consequences for the supervisor. Neither is the disapproval confounded with other negative consequences beyond the experiment (e.g. restraining an intervening supervisor from acquiring future gains), since decisions are anonymous and no further interaction takes place. Second, as participants are matched anonymously, rejection cannot be motivated by feelings of antipathy. Admittedly, under this procedure the consequences of social disapproval are weak. But it provides a clean measure of the worker's dislike for the supervisor's intervention, as it is not confounded with alternative motives like taking revenge or expressing antipathy.

3 Hypotheses

3.1 The influence of incidental affects on the intervention rate

Most information, we receive in daily life is not purely informative, but also triggers emotions. The information whether one faces challenges when entering the labor market is a good example. Given the importance of this information for our lives, receiving negative career prospects can trigger negative feelings such as fear, anxiety or even anger. Psychological and economic literature provides support for the influence of affects on decisions (e.g. Loewenstein and Lerner, 2003). Especially, when people use heuristics to take decisions, the way they feel about the situation seems to influence their decisions (Forgas, 1995). Following the rules of economic decision-making, a rational individual should only use information which are relevant for their decision. This rule of relevance also applies to affects as far as they arise directly from the decision to be made. These so-called integral affects serve as relevant information when using heuristics (Slovic et al, 2002). In other words, people should only consult their affects as a source of information if and only if they are integral. In opposite, affects that are induced by factors which are unrelated to the decision do not generate any useful information for the decision. Thus, these so-called incidental affects should be ignored by a rational individual.

In this experiment, the affect arising from receiving negative career prospects is incidental for the intervention decision in the subsequent decision process: it is entirely uninformative about the worker's ability of solving the puzzle correctly and, thus, should not alter the supervisor's likelihood of intervention. However, supervisors may draw spurious inference from their career prospects. A supervisor who has got discouraged about her own success in the labor market, may generally doubt her own abilities and also alter her guess about others performing successfully in a given case. For example, if her own ability is used as a reference point, a discouraged supervisor may also be more pessimistic about the worker's ability to solve the given puzzle. As a consequence, the affect arising from the negative career prospects becomes integral and may influence the perceived value of intervention. To address this point, one has to test whether supervisors' beliefs about the worker's success probability are altered by the News treatments. Our supervisors are asked to assess the probability that the worker has solved the puzzle correctly. If there is no treatment effect on the supervisors' assessment, the affect arising from the reception of negatively framed career prospects can be considered as truly incidental and, therefore, should not affect the supervisor's decision. Accordingly, we postulate the hypothesis ruling under full rationality:

Null Hypothesis (H0: Incidental affects). Assuming full rationality, the intervention rate is independent of the Good News and the Bad News treatment.

However, vast evidence is provided that incidental affects seem to influence people's decision and judgment (e.g. Bodenhausen, 1993; Forgas, 1995; Harlé and Sanfey, 2007; Lerner and Keltner, 2000; Loewenstein and Lerner, 2003; Schwarz, 1990; Schwarz and Clore, 1996). Psychologists like Payne et al. (2010), argue that individuals are unable to sufficiently differentiate between incidental and integral affects which can lead to a misattribution of the affect. Psychology provides several models that describe how incidental affects change judgment and decision-making. One acknowledged model is the *feeling-as-information theory* of Schwarz and Clore (1983, 1988). This model allows for the incorporation of integral as well as incidental affects. In essence, it postulates that being in a positive affective state, irrespectively of its source, reduces a person's motivation to process information, since it tells the individual that an environment is benign and secure. Most people are predominantly in a positive mood. Therefore, as an (induced) positive affect matches the prevailing positive mood of most people, it does not call for updating information or any corrective measures. Ergo, supervisors in a positive mood are prone to stick to the default option (Shevchenko et al., 2014; Yen and Chuang, 2008). Contrarily, negative affects make people more alert and, thus, increases the motivation to process information more systematically. Consequently, negative moods increase the inclination to use modes of rational reasoning and, thus, the likelihood to deviate from the default option.

In order to derive a concrete prediction, we have to define the default option in our experiment. Consider how the supervisor perceives her role and her options in

the decision process. In this experiment, the worker is put in charge of finding the correct solution and submitting it. The supervisor has a rather passive, observing role with the opportunity—but not the obligation—to intervene.⁵ Being assigned to a passive role, it is most natural to consider the non-intervention option as the default. Other factors promote this perception of non-intervention being the default. Remaining passive is for free, while intervention generates costs of 0.50€. Further, when the supervisor has to take her decision, she is asked whether she wants to intervene, instead of asking whether she wants to rely on the worker's solution. In our eyes, this question implies that choosing intervention is the deviation from the default of inactivity. Last, in the instructions and the graphic of the decision process handed out to the participants (see appendix), the non-intervention option is presented as the first and straight-forward option, while the option of intervention is presented as one that interrupts the initial process. In a nutshell, supervisors may see themselves in a passive role with the default behavior of being inactive.⁶ In economic literature, there is an abundance of evidence that subjects have a preference of staying with the default option (e.g. Beshears et al., 2009; Haward et al., 2012; Johnson et al., 2002; Johnson and Goldstein, 2003; Johnson et al., 1993; Madrian and Shea, 2001).

While non-intervention is considered as the default option, it is also the riskier option in this experiment: a non-intervening supervisor decides to play the lottery of relying on the unknown worker's solution instead of taking the intervention option which guarantees the bonus payment. This is in contrast to most of the existing literature, where the default option of staying in the current state or being reluctant to take action is associated with lower risk. Nevertheless, we argue that the general preference for the default is not entirely removed from the decision process, but only mitigated under a reversed risk structure. Choosing to get active in a decision process is usually a deliberate decision to join responsibility for the outcome, as it is in the study here. Since people shy away from taking responsibility, they prefer to stay passive, even though their participation could potentially reduce risk for all parties.

Considering the non-intervention option as the default, the feeling-as-information theory predicts more intervention in the Bad News treatment where the supervisors

⁵These functions are underlined by assigning significant names to each role: Indeed, our participants only know the roles as the decider (worker) and the observer (supervisor).

⁶We purposely have programmed that the non-intervention option is not marked as the default on the computer screens. In the progress of designing a comprehensive, and clean screen layout, we have used a coding that does not allow to mark a single option as a default. Adding to this, using another coding where setting a default is possible brings about the risk that a supervisor accidentally jumps the intervention decision by prematurely clicking on the forward button.

are exposed to the mood-deteriorating information of having negative future career prospects.

According to a different model, incidental affects may also influence decision making through means of *mood-repair strategies*. People usually aim at maintaining their positive mood. When experiencing a deterioration of their mood through e.g. frustration, fear or anger, people usually take mood-repair strategies to restore a positive mood. For example, frustrated people use aggression to improve their mood (Bushman et al., 2001). People in a negative affective state have need for mood repair and, thus, are more likely to opt for a change of their situation instead of sticking with the default. According to Yen and Chuang (2008), mood-deteriorated people focus intensively on the gains of the alternative option. Vice versa, they are also more sensitive towards the losses which are associated with the omission of getting active. Therefore, supervisors should be more inclined to intervene when experiencing a mood deterioration, e.g. by receiving negatively-framed career prospects. This behavior should be even more pronounced when passiveness results in higher risks and activeness promises a significant gain with certainty.

The hypothesis derived from the approach of mood-repairing is identical to the former hypothesis under the feeling-as-information theory. Consequently, we state the following alternative hypothesis:

Alternative Hypothesis (H1). *Given that incidental affects matter, the intervention rate is higher in the Bad News treatment than in the Good News treatment.*

Other psychological approaches, postulating the no-action option as the default, suggest a hypothesis of opposite direction. First, according to the *mood-maintenance hypothesis*, people in a positive mood behave more risk averse in order to not endanger their positive mood. This makes intervention more attractive than remaining passive, as the intervention option is the safe option in our experiment, while remaining passive exhibits the risk of losing the bonus payment. Therefore, intervention also becomes interesting for supervisors in the Good News treatment.⁷

Second, supervisors who have received the negatively framed job prospects may suffer from *ego depletion*. Negative information induces the need to rethink a

⁷In contrast to other studies, the default option in our experiment is the risky option. Note that reversing the risk structure leads to an oppositely directed hypothesis. If doing nothing is safer, supervisors in the Good News treatment would not have an incentive to deviate from the default option. The resulting hypothesis would thus correspond to our alternative hypothesis (H1) and hypotheses derived from the mood-maintenance approach in other studies (e.g. Yen and Chuang, 2008; Shevchenko et al., 2014).

situation and to evaluate the information, while positive information may not be seen as a deviation from what is expected and, thus, a person can continue with her routine. Ergo, negative information may require more cognitive effort to evaluate the information and self-control effort to deal with negative affects that are attached to this information. Following the ego depletion approach, subjects are less willing or able to exert further energy to engage in cognitive reasoning in the subsequent decision process and, thus, ego depleted people may be more inclined to rely on the default option of inactivity.

Deriving a hypothesis from these two approaches, we state the following conflicting alternative hypothesis:

Alternative Hypothesis (H1'). *Given that incidental affects matter, the intervention rate is lower in the Bad News treatment than in the Good News treatment.*

3.2 The role of anticipated regret on the intervention rate

Before deciding to intervene in a worker's decision, a supervisor may ask herself whether her participation is needed or even welcome by others. If this is not the case, a supervisor may ex-post regret her intervention. Anticipating such feelings of regret in advance alter the expected utility of an option and, thus, may induce a supervisor to change decisions. The economists Loomes and Sugden (1982) and Bell (1982) simultaneously developed theoretical decision models where subjects also take anticipated feelings of regret and rejoicing into account when maximizing their expected utility. Various studies provide evidence that the anticipation of regret significantly influences subjects' decision making (for a detailed review see Zeelenberg, 1999).

In this paper, we study two sources of regret that a supervisor may anticipate before getting active. The first source is the psychological impact of forgoing economic payoffs. A rational subject is eager to make an efficient decision. Having taken an inefficient solution burns resources and usually triggers feelings of regret. Thus, she may regret having become active if intervention is not needed and, therefore, inefficient. In our experiment, it is inefficient for supervisors to intervene when the solution of the worker has been correct, because the involved costs are spent in vain. Vice versa, there is also scope for anticipating regret when remaining passive: not intervening is inefficient when the worker's solution is wrong, since the loss of the bonus payment is larger than the costs of intervention. Second, a supervisor may regret her intervention if she faces social disapproval. When getting active, she steps out of her initial role, joins responsibility and becomes accountable for the outcome. Such an interference with the initial distribution of responsibility is not always welcome by the person(s) in charge, in this case the worker, even though the intervention may be efficient. Therefore, an intervening supervisor exposes herself to the risk that her intervention gets rejected. As people strive for social approval, seeing their intervention being rejected by others is psychologically costly. These costs are amplified if intervention is relatively more helpful to others than to oneself, as it is the case in our experiment.⁸ As a result, experiencing opposition against her intervention - especially if it is also seen as an act of help - can cause the supervisor to regret having become active. In this experiment, the worker has the chance to show her disapproval by ex-post rejecting the intervention. Even though, the rejection has no monetary consequences for the supervisor in our experiment, she may incur psychological costs. Anticipating these psychological costs of being rejected reduced the supervisor's propensity to intervene.

A precondition for the anticipation of regret is feedback: people can only feel regret if they learn about the inefficiency or the social disapproval of their decision. As Zeelenberg (1999) points out, "post-decisional feedback is a central determinant of experienced and anticipated regret. When this feedback is present, people anticipate possible regret, but when it is absent regret does not play a significant role in the decision process" (p.103). In this experiment, we designed two types of post-decisional feedback that are given to the supervisors. Under the efficiency feedback, they are informed about the efficiency of their decision by learning whether the worker's solution is correct. This information gives the supervisor a clear signal about whether her decision has been efficient or whether her effort has been spent in vain. Under the disapproval feedback, supervisors are told whether their worker has accepted their intervention. This feedback serves as a direct information on the social disapproval of intervention. Depending on the Feedback treatment, supervisors receive either both or one of these feedback types (see Table 2.1). Supervisors in the FF treatment are exposed to both types of feedback allowing them to feel regret due to inefficiency and due to social disapproval. In the EF treatment, supervisors receive the efficiency feedback but not the disapproval feedback. Therefore, a supervisor in the EF treatment can only anticipate regret due to inefficiency. In opposite, supervisors in the RF treatment do not get the efficiency feedback. They only receive the disapproval feedback. This feedback serves as a direct information on the social disapproval, on the one hand. On the other hand, it is also a noisy signal for the inefficiency of her decision.

⁸The worker receives $4 \in$ for the correct answer, while the supervisor receives only $2.50 \in$.

A worker who rejects the intervention, even though her own solution is incorrect, loses her bonus payment. This makes rejection extremely cost. As a consequence, the supervisor is able to draw inference from the worker's rejection decision on the efficiency of intervention. Nevertheless, in the RF treatment she does not have full assurance, contrarily to the other two treatments: as rejection generates costs, the worker may accept out of profit-maximization, even though her solution is correct.

As described above, regret generates only psychological costs. Whether the supervisor receives any of the above described feedback or not, has de facto no influence on the monetary profitability of intervention: the reception of feedback neither changes the consequences of intervention nor its costs. It may only change the perceived profitability of intervention if regret is triggered. When receiving the feedback that the intervention is inefficient, the supervisor learns that a better outcome could have been achieved if she did not intervene. Feedback allows her to compare the actual payoff to the forgone payoff causing her to be less satisfied with the outcome than she could have been in the absence of feedback. Therefore, the presence of feedback triggers a perceived loss which is exclusively psychological. Likewise, the regret that is triggered by social disapproval is only psychologically grounded. Assuming that regret is only psychologically costly, a rational subject who behaves perfectly profit-maximizing should be immune to the anticipation of regret. Thus, post-decisional reception of any type of feedback should not alter her willingness to intervene. Rather, a profit-maximizing supervisor is expected to choose the option that generated the highest expected payoff for the given payoff structure and her subjective probability that the worker solves the puzzle correctly. Consequently, we state the following null hypothesis ruling under the rationality assumption of profit-maximization:

Null Hypothesis (H0': Feedback). *Given that supervisors behave perfectly profitmaximizing, the intervention rate is identical across all Feedback treatments.*

Observations from behavioral economic studies have proven that subjects do not behave perfectly profit-maximizing, but also account for psychological benefits and costs which causes them to forfeit monetary profits. Zeelenberg et al. (1996) provide evidence that subjects behave regret averse: subjects tend to minimize the psychological costs of regret instead of maximizing profits. For our experiment, the idea of regret aversion suggests that intervention becomes less attractive for supervisors when feedback exposes them to the anticipation of regret. Assuming that subjects are simultaneously receptive to different types of regret, more feedback should increase the scope for anticipating regret and, thus, reduces the willingness to intervene. Consequently, we hypothesize that the intervention rate is lowest in the FF treatment where supervisors anticipate the highest degree of regret, as they are clearly exposed to both types of feedback.

Alternative Hypothesis (H2). *Given that supervisors take post-decisional regret into account, the intervention rate in the FF treatment is lower than in the RF and EF treatments.*

No hypothesis on the difference in intervention rates between the RF and EF treatment is made, since this would require strong assumptions on the quality of different types of regret.⁹

4 Procedure and Descriptive Statistics

The experiment was run in the BaER-Lab laboratory at Paderborn University in March and May 2016. 12 sessions were conducted. Each session was assigned to one of three Feedback treatments. The News treatments were simultaneously run within a session. The numbers of participants were mostly balanced across the treatments and treatment combinations (see Table 2.2) ranging from 18 to 23 observations per treatment combination. A session took between 48-61 minutes excluding the time for paying the participants. The experiment was computerized using the software z-Tree (Fischbacher, 2007). All subjects were recruited from the same subject pool via ORSEE (Greiner, 2015). Every subject could participate in only one session. Those who already participated in March 2016 were not re-invited in May 2016.

In total, 244 participants took part in the experiment. As they were matched in pairs, half of them acted in the role of a supervisor. Therefore, the data set consists of 122 independent observations derived from those participants assigned to the role of the supervisor.¹⁰

⁹Whether regret arising from economic losses is more or less deterrent than regret from social disapproval, should be highly subjective as well as context-bound. Only if economic losses and social disapproval trigger regret of equal size, one could predict that the intervention rate in the RF treatment should be lower than in the EF treatment. Note that the disapproval feedback in the RF treatment also serves as a noisy signal for the inefficiency of intervention. However, given the strong assumption of equal quality of these types of regret, we do not formulate any hypothesis here.

¹⁰There was no supervisor who indicated to be not a student of the Paderborn University. Thus, no observations was deleted, since all supervisors received a News treatment.

	Bad News Treatment	Good News Treatment	Total
RF Treatment			
Observations	23	22	45
Female participants (in %)	34.78	36.36	35.56
Payoff in €	9.26	9.64	9.44
FF Treatment			
Observations	20	19	39
Female participants (in %)	40.00	26.32	33.33
Payoff in €	9.53	9.18	9.36
EF Treatment			
Observations	18	20	38
Female participants (in %)	72.22	30.00	50.00
Payoff in €	9.28	9.40	9.34
Total			
Observations	61	61	122
Female participants (in %)	47.54	31.15	39.34
Payoff in €	9.35	9.42	9.39

Table 2.2: Summary Statistics

The recruited subjects had different study backgrounds: 47.54% of the supervisor intended to get a teaching degree, 31.97% were registered at the faculty of Business Administration and Economics, the remaining participants enrolled in other study fields, such as cultural sciences, natural sciences or engineering, computer sciences and mathematics. In total, 39.34% of the supervisors were female. However, the female participation rate varied strongly across the treatment combinations (see Table 2.2). When analyzing the data, we consider this imbalance.

During the experiment, subjects received their payments in ECU which were converted in euros (1 ECU = 5 cents). Additionally, subjects were paid a fixed participation fee of $2.50 \in$. All payments were handed out in cash at the end of each session. On average, a supervisor received $9.39 \in$. Subjects were provided with printed instructions explaining the experiment. Supervisor and worker received identical instructions. In order to ensure that all subjects had comprehended the instructions, they had to answer control questions before starting the experiment. Subjects could not proceed with the experiment, unless they answered all questions correctly. After the experiment, subjects answered a survey with socio-demographic questions and questions concerning the experiment and their behavior within.

5 Results

This section presents the results. First, we study the effect of receiving negatively framed career prospects on supervisors' willingness to intervene in a worker's decision. Precisely, we ask whether supervisors in the Bad News treatment intervene more often than supervisors in the Good News treatment. Second, we address the question whether anticipating regret deters supervisors from intervention. In other words, we test whether the intervention rate is lower in the FF treatment, where the supervisors receive two types of feedback—on the efficiency and social disapproval of intervention, than in the other treatments (EF and RF treatments) where only one type of feedback—either efficiency or disapproval feedback—is given to the supervisors.

5.1 The influence of incidental affects on the intervention rate

In the Hypotheses section, we postulated that affects arising from negative career prospects are purely incidental and, thus, uninformative. However, it has to be tested whether this is true. If it turns out that supervisors draw spurious inference from their personal career prospects on the worker's ability to perform in the puzzle task, the affect arising from the reception of negatively framed career prospects cannot be considered as incidental. To address this point, our supervisors are asked to assess the probability that the worker has solved the puzzle correctly. The results from a two-sample Wilcoxon rank-sum (Mann-Whitney) test shows that supervisors in the Bad News treatment do not make significant different assessment on the worker's ability than supervisors in the Good News treatment (p=0.3837). Not only is the difference not significant, it also points the opposite direction as expected: supervisors in the Good News treatment are slightly more pessimistic about the worker's ability than supervisors in the Bad News treatment. Supervisors do not seem to link their negative career prospects to the worker's ability to successfully perform in the puzzle task. Ergo, the affect arising from the negative career prospect seems to be irrelevant for the supervisor's intervention decision and can thus be considered as incidental.

Let us first consider the null hypothesis (H0: Incidental affects). Table 2.3 shows that the difference is quite sizable: looking at the entire data set, the intervention rate in the Bad News treatment is almost 12% lower than in the Good News treatment. An illustration of the intervention rates is given in Figure 2.2. This finding is not significant (Fisher's exact test, p=0.264) and is even in the opposite direction of our first alternative hypothesis (H1) which postulates that if incidental affects matter, the intervention rate in the Bad News treatment should be higher than in the Good News treatment. In this experiment, we rather observe that incidental affects seem to work in the opposite direction, as it was postulated by our second alternative hypothesis (H1'). This result is consistent across all Feedback treatment variations: for all feedback treatments, the intervention rate in the Bad News treatment (Fisher's exact tests, the p values are between 0.527 and 0.734; see Figure 2.2). In a nutshell, if incidental negative affects matter at all, they rather may discourage supervisors from intervening in a worker's decision.

A Linear Probability Model (LMP) regression is run (Table 2.4) that does not only account for the size of the effects in the News Treatments and the Feedback treatments but also for the interactions of them, as every supervisor has received a combination of one News treatment and one Feedback treatment. The reference group consists of supervisors who have received the Bad News treatment and the FF treatment. The results of the LPM shows that there is no statistically significant increase in intervention for those supervisors who have participated in the FF treatment and received the Good News instead of the Bad News treatment (see Table 2.4; p=0.413). Likewise, being in the RF treatment, respectively the EF treatment, receiving the Good News instead of the Bad News treatment does not significantly change the intervention rate. Therefore, we cannot reject our null



Figure 2.2: Intervention Rates in the Feedback Treatments, by News treatments

	Bad News Treatment	Good News Treatment	Total
Rejection Feedback Treatment	60.87%	72.73%	66.67%
	(14/23)	(16/22)	(30/45)
Full Feedback Treatment	45.00%	57.89%	51.28%
	(9/20)	(11/19)	(20/39)
Efficiency Feedback Treatment	61.11%	70.00%	65.79%
	(11/18)	(14/20)	(25/38)
Total	55.74%	67.21%	61.48%
	(34/61)	(41/61)	(75/122)

 Table 2.3: Intervention Rates per treatment

hypothesis (H0: incidental affects).

Result 1. The intervention rate in the Bad News treatment is on average 9% to 13% lower than in the Good News treatment. While the difference is sizable, it is not statistically significant.

To account for the large imbalance of female participation across the EF treatment combinations (see Table 2.2), the LPM was extended by adding controls. This specification leads to only minor changes. (The results are available on request.)

5.2 Anticipating regret: the effect of feedback on the intervention rate

In this section, we test whether the anticipation of regret reduces a supervisor's inclination to intervene in a worker's decision. In the Hypotheses section, we postulate that supervisors should be less inclined to intervene, the more they are confronted with feedback that causes them to regret their decision. Assuming that different types of regret do not cancel out each other, we predict that the intervention rate should be lowest in the FF treatment where both type of feedback

Dependent variable	1 if supervisor decided to intervene
Good_News_treat	.1289
	(.1571)
RF_treat	.1587
	(.1499)
EF_treat	.1611
	(.1593)
Good_News*RF_treat	0104
	(.2147)
Good_News*EF_treat	0401
	(.2238)
cons_	.4500***
	(.1097)
N. obs.	122
<i>R</i> ²	0.0343

Table 2.4: Linear Probability Model

Notes: Linear Probability estimates. Robust standard errors are in parentheses. The reference group are supervisors in the Bad News treatment who also receive the FF treatment.

 \ast significant at p<0.10; $\ast\ast$ significant at p<0.05; $\ast\ast\ast\ast$ significant at p<0.01

are shared.

Indeed, looking at Table 2.3 and Figure 2.3 displays that the intervention rate in the FF treatment is lower than in the RF treatment and EF treatment irrespective of the News treatment. In the Bad News treatment, the intervention rate is 15.87% higher in the RF treatment and 16.11% higher in the EF treatment than in the FF treatment (Fisher's exact tests, p=0.366, respectively 0.352; Figure 2.3). Under the Good News treatment, the differences are in the same direction as under the Bad News treatment but less pronounced: 14.84% higher in the RF treatment and 12.11% higher in the EF treatment (Fisher's exact tests, p=0.346, respectively 0.514; Figure 2.3). This finding is in conflict with our null hypothesis (H0': Feedback). Though, the results are statistically not significant.

The LPM (Table 2.4) confirms that the results are statistically not significant for both types of feedback (p=0.292 for the RF treatment; p=0.314 for the EF treatment). Ergo, we cannot rule the possibility out that the observed reduction in intervention is due to random behavior. In other words, we cannot reject our null hypothesis (H0': Feedback).

Result 2. The intervention rate in the FF treatment is on average lower than in the RF treatment and EF treatments. While the differences are sizable (approximately 16% in both cases under the Bad News treatment), they are not statistically significant.

6 Discussion

In this section, our results are discussed. Then, the psychological approaches presented in the Hypotheses part are reviewed with respect to their ability to explain our results on incidental affects.

For the Bad News treatment, we observe an intervention rate which is on average between 9% and 13% lower than for the Good News treatment. While the effect is quite sizable, our finding falls short of statistical significance. Hence, if negative incidental affects matter, they may have a rather hampering than promoting effect on supervisors' willingness to intervene. Among those who would intervene in a worker's decision if being in a positive mood, approximately every eighth (EF treatment) to fourth supervisor (FF treatment) would not do so if being in a negative mood. In the authors' eyes, this magnitude is economically significant. For the role of feedback on intervention, our observations draw a similar picture. If receiving


Figure 2.3: Intervention Rates in News Treatments, by Feedback treatments

feedback that triggers the anticipation of regret influences supervisors in their intervention decision, it may discourage supervisors from intervening. We find that the intervention rate in the FF treatment is approximately 16% lower than in the EF, respectively in RF treatment. Stated differently, among those who would be willing to intervene, round about every fifth (Good News Treatment) to more than every fourth supervisor (Bad News treatment) would not do so if they were exposed to more feedback.

While the effects are large in magnitude, we cannot exclude that our observations are due to random behavior given the lack of significance. Nevertheless, the relative small sample size resulting from comparing subgroups (n=23 for the largest subgroup, see Table 2.2) may contribute to the low level of significance. Unfortunately, the available subject pool prohibits to extend the data set. Indeed, to obtain the number of observations used in this present study, we exhausted the sizable subject-pool of the BaER-Lab. Given the response rate at the time of conducting the experiment, we were able to recruit 244 students with half of them acting in the role of a supervisor.

The results on incidental affects indicate that if negative incidental affects matter at all, they may induce supervisors to intervene less. This is in opposite to the hypothesis (H1) derived from the psychological models of the *feeling-as-information theory* and the *mood-repair hypothesis*. Therefore, these approaches alone do not seem to explain the observed behavior. If the behavioral pattern described by these approaches is counteracted by the *mood-maintenance strategy* of mood-elevated supervisors in the Good News treatment, it would not be surprising that the treatment effect turns out to be statistically not significant, since these approaches predict opposing behavioral patterns. Ergo, supervisors in both treatments could have an incentive to deviate from the default. Further research is needed to test whether mood-maintenance by mood-elevated supervisors and mood-repair by depressed supervisors are simultaneously at work.

A more promising explanation for our results may be given by the *ego depletion* approach. One indicator for the existence of ego depletion could be the variance in the supervisors' assessments on the worker's ability to successfully perform in the puzzle task. Supervisors are asked to rate the probability that the worker's solution is correct at a scale from 0 (100% incorrect) to 10 (100% correct). If supervisors in the Bad News treatment suffer from ego depletion, the assessments in this treatment are expected to show a greater variance than the assessments made by supervisors in the Good News treatment, since ego depleted subjects are usually less able to make correct estimations. A visual test of Figure 2.4 suggests that the variances of assessments differ across the News treatments. In a nutshell, there is



Figure 2.4: Superviors' assessments about the correctness of the worker's solution

vague indication that supervisors in the Bad News treatment could suffer from ego depletion. Though, further research is needed to confirm this idea.

7 Conclusion

This paper explores two channels which may cause supervisors not to intervene in a worker's decision. First and primary, it is studied whether negative incidental affects have an effect on the willingness to intervene. Our results suggest that if negative incidental emotions matter at all, they have a rather hampering effect on a supervisor's inclination to intervene: in the mood-depressing Bad News treatment supervisors intervene 9% to 13% less than supervisors in the mood-enhancing Good News treatment. This proportion is sizable and economically relevant. Though, it is not statistically significant.

We argue that the psychological theory of *feeling-as-information* of Schwarz and Clore (1983, 1988) and the *mood-repair hypothesis* do not sufficiently predict the influence of incidental affects on intervention behavior. Instead, *ego depletion* might be a possible explanation why supervisors decide to remain passive after suffering from mood-depressing but unrelated incidences. Additionally, attempts of *mood-maintenance* by supervisors in a positive mood may increase the rate of

risk-reducing intervention. This would confound with mood-repairing behavior by depressed supervisors in our experiment who are expected to opt for intervention as a corrective action. Further research is needed to study whether these confounding forces are simultaneously at work. If so, this may contribute the low significance of our results.

Second, this paper studies the role of anticipating regret on a supervisor's willingness to intervene. Our results suggest that if feedback has an influence on supervisors at all, it rather reduces their willingness to intervene: we observe that supervisors intervene less (approximately 16% for depressed supervisors) when they are exposed to regret triggering feedback—be it on the efficiency or the social (dis)approval of intervention. However, the effects are not statistically significant. Considering the disapproval feedback, the null result might be the consequence of the relative weak implementation of social disapproval in our design. As described in the Experimental Design part, we purposely decided on a procedure that measures social disapproval in a very clean way: by excluding other influences that usually come along with social disapproval, we are able to measure the pure effect of fearing others' rejection. Though, the cleanness of this procedure is bought at the price of a weak implementation. It should be clear to the reader that in real situations where people cannot hide behind the veil of anonymity and probably face future interactions, the effect is expected to be much stronger. In this sense, the results derived from this experiment could be regarded as a the lower bound.

Given that the observed effects are statistically weak, we cannot reliably recommend concrete policy measures. Further research is needed to provide more insight into the role of incidental emotions and anticipation of regret for intervention.

References

- Bell, D. (1982). "Regret in Decision Making under Uncertainty." Operations Research, 30(5): 961-981.
- Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2009). "The importance of default options for retirement saving outcomes: Evidence from the United States." In: Social security policy in a changing environment (pp. 167-195). University of Chicago Press.
- Bodenhausen, G. V. (1993). "Emotions, arousal, and stereotypic judgments: A heuristic model of affect and stereotyping." In: Mackie, D. M., & Hamilton, D. L. (Eds.). Affect, cognition, and stereotyping: Interactive processes in

group perception (pp. 13-37). San Diego, CA: Academic Press.

- Bushman, B. J., Baumeister, R. F., & Phillips, C. M. (2001). "Do people aggress to improve their mood? Catharsis beliefs, affect regulation opportunity, and aggressive responding." Journal of Personality and Social Psychology, 81(1): 17-32.
- Dugar, S. (2010). "Nonmonetary sanctions and rewards in an experimental coordination game." Journal of Economic Behavior & Organization, 73(3): 377-386.
- Filiz, E., & Ozbay, E. Y. (2007). "Auctions with anticipated regret: Theory and experiment." The American Economic Review, 97(4): 1407-1418.
- Fischbacher, U. (2007). "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." Experimental Economics, 10(2): 171-178.
- Forgas, J. P. (1995). "Mood and judgment: the affect infusion model (AIM)." Psychological Bulletin, 117(1): 39-66.
- Gächter, S., & Fehr, E. (1999). "Collective action as a social exchange." Journal of Economic Behavior & Organization, 39(4): 341-369.
- Greiner, B. (2015). "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE." Journal of the Economic Science Association, 1(1): 114-125.
- Harlé K. M., & Sanfey, A. G. (2007). "Incidental sadness biases social economic decisions in the Ultimatum Game." Emotion, 7(4): 876-881.
- Haward, M. F., Murphy, R. O., & Lorenz, J. M. (2012). "Default options and neonatal resuscitation decisions." Journal of Medical Ethics, 38(12): 713-718.
- Hockey, G. R. J., John Maule, A., Clough, P. J., & Bdzola, L. (2000). "Effects of negative mood states on risk in everyday decision making." Cognition & Emotion, 14(6): 823-855.
- Inman, J. J., & Zeelenberg, M. (2002). "Regret in repeat purchase versus switching decisions: The attenuating role of decision justifiability." Journal of Consumer Research, 29(1): 116-128.
- Isen, A. M. (1993). "Positive affect and decision making." In: Lewis, M., & Haviland, J. M. (Eds.). Handbook of emotions (pp. 261-277). New York: Guilford Press.
- Johnson, E. J., Bellman, S., & Lohse, G. L. (2002). "Defaults, framing and privacy: Why opting in-opting out." Marketing Letters, 13(1): 5-15.

- Johnson, E. J., & Goldstein, D. (2003). "MEDICINE: Do Defaults Save Lives?" Science, 302(5649): 1338-1339.
- Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). "Framing, probability distortions, and insurance decisions." Journal of Risk and Uncertainty, 7(1): 35-51.
- Johnson, E. J., & Tversky, A. (1983). "Affect, generalization, and the perception of risk." Journal of Personality and Social Psychology, 45(1): 20-31.
- Kirchsteiger, G., Rigotti, L., & Rustichini, A. (2006). "Your morals might be your moods." Journal of Economic Behavior & Organization, 59(2): 155-172.
- Lerner, J. S., & Keltner, D. (2000). "Beyond valence: Toward a model of emotion-specific influences on judgement and choice." Cognition & Emotion, 14(4): 473-493.
- Lerner, J. S., & Keltner, D. (2001). "Fear, anger, and risk." Journal of Personality and Social Psychology, 81(1): 146-159.
- Leith, K. P., & Baumeister, R. F. (1996). "Why do bad moods increase self-defeating behavior? Emotion, risk tasking, and self-regulation." Journal of Personality and Social Psychology, 71(6): 1250-1267.
- Loewenstein, G., & Lerner, J. S. (2003). "The role of affect in decision making." Handbook of Affective Science, 3: 619-642.
- Loomes, G., & Sugden, R. (1982). "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty." The Economic Journal, 92(368): 805-824.
- Lübbecke, S. & Schnedler, W. (2018). "Don't patronize me! An experiment on the Motives behind Rejecting Paternalistic Help." Working Paper Series, No. 2018-06, Paderborn University, Faculty of Business Administration and Economics.
- Lumeau, M., Masclet, D., & Pénard, T. (2015). "Reputation and social (dis)approval in feedback mechanisms: An experimental study." Journal of Economic Behavior & Organization, 112: 127-140.
- Madrian, B. C., & Shea, D. F. (2001). "The power of suggestion: Inertia in 401 (k) participation and savings behavior." The Quarterly Journal of Economics, 116(4): 1149-1187.
- Mano, H. (1992). "Judgments under distress: Assessing the role of unpleasantness and arousal in judgment formation." Organizational Behavior and Human Decision Processes, 52(2): 216-245.

- Masclet, D., Noussair, C., Tucker, S., & Villeval, M. (2003). "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism." The American Economic Review, 93(1): 366-380.
- Mittal, V., & Ross, Jr. W. T. (1998). "The impact of positive and negative affect and issue framing on issue interpretation and risk taking." Organizational Behavior and Human Decision Processes, 76: 298-324.
- Neff, J. J., Fulk, J., & Yuan, Y. C. (2014). "Not in the mood? Affective state and transactive communication." Journal of Communication, 64(5): 785-805.
- Nygren, T. E., Isen, A. M., Taylor, P. J., & Dulin, J. (1996). "The influence of positive affect on the decision rule in risk situations: Focus on outcome (and especially avoidance of loss) rather than probability." Organizational Behavior and Human Decision Processes, 66: 59-72.
- Payne, B. K., Hall, D. L., Cameron, C. D., & Bishara, A. J. (2010). "A process model of affect misattribution." Personality and Social Psychology Bulletin, 36(10): 1397-1408.
- Pfaff, M. S. (2012). "Negative affect reduces team awareness: The effects of mood and stress on computer-mediated team communication." Human Factors, 54(4): 560-571.
- Rege, M., & Telle, K. (2004). "The impact of social approval and framing on cooperation in public good situations." Journal of Public Economics, 88(7): 1625-1644.
- Schwarz, N. (1990). "Feelings as information: Informational and motivational functions of affective states." In: Higgins, E. T., & Sorrentino, R. M. (Eds.). Handbook of motivation and cognition: Foundations of social behavior, 2 (pp. 527-561). New York: Guilford Press.
- Schwarz, N., & Clore, G. L. (1983). "Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states." Journal of Personality and Social Psychology, 45(3): 513-523.
- Schwarz, N., & Clore, G. (1988). "How do I feel about it? Informative functions of affective states." In: Fiedler, K., & Forgas, J. (Eds.). Affect, cognition and social behavior (pp. 44-62). Toronto, Ontario, Canada: Hogrefe International.
- Schwarz, N., & Clore, G. L. (1996). "Feelings and phenomenal experiences." Social psychology: Handbook of Basic Principles, 2: 385-407.
- Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). "Rational actors or rational fools: Implications of the affect heuristic for behavioral

economics." The Journal of Socio-Economics, 31(4): 329-342.

- Shevchenko, Y., Von Helversen, B. & Scheibehenne, B. (2014). "Change and status quo in decisions with defaults: The effect of incidental emotions depends on the type of default." Judgment and Decision Making, 9(3): 287-296.
- Yen, H. R., & Chuang, S. C. (2008). "The effect of incidental affect on preference for the status quo." Journal of the Academy of Marketing Science, 36(4): 522-537.
- Yuen, K. S., & Lee, T. M. (2003). "Could mood state affect risk-taking decisions?" Journal of Affective Disorders, 75(1): 11-18.
- Zeelenberg, M. (1999). "Anticipated Regret, Expected Feedback and Behavioral Decision Making." Journal Of Behavioral Decision Making, 12(2): 93-106.
- Zeelenberg, M., & Beattie, J. (1997). "Consequences of regret aversion 2: Additional evidence for effects of feedback on decision making." Organizational Behavior and Human Decision Processes, 72(1): 63-78.
- Zeelenberg, M., Beattie, J., Van der Pligt, J., & De Vries, N. K. (1996).
 "Consequences of regret aversion: Effects of expected feedback on risky decision making." Organizational Behavior and Human Decision Processes, 65(2): 148-158.

Appendix

Original screen shots in German with translation

The Bad News treatment

Für Studierende mit Ihren Charakteristika sieht es laut Absolventenbefragung nicht so gut auf dem Arbeitsmarkt aus.

Rund jeder 3. ist nach dem Studium längerfristig (min. 4 Monate) auf der Suche nach einer Beschäftigung. Jeder 8. muss mehr als ein Jahr überbrücken, bis er seine erste Stelle antreten kann. Fast jeder 5. muss mehr als 10 Bewerbungen schreiben, um eine Beschäftigung zu finden. Anderthalb Jahre nach dem Studium gehen immer noch mehr als die Hälfte keiner regulären Beschäftigung nach. Mehr als die Hälfte aller Absolventen hat anderthalb Jahre nach dem Studium nur eine befristete Beschäftigung. Und rund die Hälfte ist weniger zufrieden mit Ihrem Beruf.

Figure 2.5: Original screen shot: Bad News treatment (German)

English translation:

"For students with your characteristic it does not look that good at the labor market according to the alumni study.

Round about every 3rd person has been searching long-term (at least 4 months) for an employment after studying.

Every 8th *person has to bridge a year before starting the first job.*

Almost every 5th has to write more than 10 applications to find an employment. One year and a half after graduation more than half do not have a regular employment.

More than half of the alumni only have a limited contract one year and a half after graduation.

And round about half are less satisfied with their job."

Glückwunsch! Studierende mit Ihren Charakteristika haben laut Absolventenbefragung gute Chancen auf dem Arbeitsmarkt.

25% brauchen weniger als 1 Monat, um Ihre erste Stelle nach dem Studium zu finden. Nach 3 Monaten hat rund die Hälfte eine Beschäftigung gefunden.

Fast 40% müssen nur eine oder sogar keine Bewerbung schreiben, um eine Stelle zu finden.

Anderthalb Jahre nach dem Studium suchen nur noch 3% eine Beschäftigung. 3/4 derjenigen, die ausschließlich erwerbstätig sind, haben anderthalb Jahre nach dem Studium eine unbefristete Stelle.

Und mehr als die Hälfte sind mit Ihrem Beruf hoch zufrieden.

Figure 2.6: Original screen shot: Good News treatment (German)

English translation:

"Congratulation! Students with your characteristics have good chances at the labor market according to the alumni study.

25% need less than a month to find the first job after the studies.
After 3 months round about half have found an employment.
Almost 40% have to write only one or even no application to find an employment.
One year and a half after graduation only 3% are looking for an employment.
3/4 of those, who are exclusively gainfully employed, have an unlimited contract one year and a half after graduation.
And more than half are highly satisfied with their job."

Graphic illustrating the decision process—translated into English



Figure 2.7: Graphic of the decision process, handed out in the FF treatment

Supervisors in the RF treatment and EF treatment received a graphic which is identical except for the necessary treatment variations. Note that our participants know the roles in the decision process as the observer (supervisor) and decider (worker).

As the experiment described in this paper is embedded into the experiment presented in Lübbecke and Schnedler (2018), we used the same instructions. Please, see the instructions in the appendix of Lübbecke and Schnedler (2018). Note that the Feedback treatments have been renamed for the purpose of this paper (RF treatment = RI treatment; FF treatment = FI treatment; EF treatment = CI treatment).

Chapter 3

Believing in Others' Dishonesty: An Experimental Study on Beliefs about Lying

Silvia Lübbecke University of Paderborn

Abstract

Several experiments provide evidence for discriminating behavior towards the out-group even in settings where group division is arbitrary. This paper studies whether discriminatory behavior can be traced back to subjects holding discriminating beliefs. An experiment is presented where subjects are randomly assigned to minimal groups. First, subjects are asked to draw a marble in private and report whether it is white or speckled. Second, their beliefs are elicited about how many of the others in the respective group have reported the payoff-maximizing speckled marble. Data show that subjects expect others to behave dishonestly in general, but do not differ in their beliefs about the behavior of in- and outgroup members. Further, the results indicate that subjects' beliefs about others' honesty are positively correlated with the individual lying behavior. Subjects who report the profitmaximizing type also believe in significantly more payoff-maximizing reports by others compared to those subjects who report the unfavorable outcome.

Keywords: Group identity, Minimal groups, Intergroup discrimination, Incentivized belief elicitation, Experimental economics

1 Introduction

Western societies raise the claim of being tolerant, liberal and guaranteeing equality for all people independent of their sex, origin, race, language, religion, or political opinion, as written in the Universal Declaration of Human Rights. Despite this claim, right-winged political parties have recently gained votes with prejudicial statements against certain groups-be it Latin and Black Americans in the USA or Muslim and African immigrants in Europe. All too often, minorities fall victim to discrimination—even in open-minded societies. An easily accepted explanation is that social distance nourishes discriminating behavior, as xenophobia leads people to reject the unfamiliar. Evidence from experiments shows that social distance alone can motivate subjects to discriminate against an out-group. Tajfel et al. (1971) find that inter-group discrimination even occurs in settings of minimal social distance where group identity is artificially induced and, thus, arbitrary. In the sight of arbitrary group formation, this discrimination seems irrational. However, many studies have confirmed this so-called minimal group paradigm (e.g. Brewer, 1979; Chen and Li, 2009). The puzzling results from minimal groups raises the question whether this irrational discrimination behavior arises from preconceived beliefs (akin to statistical discrimination¹) or from preferences (taste-based discrimination²).

In the standard economic theory, preferences are assumed to be stable and, thus, should not be subject to irrelevant information. Mapping this to minimal group manipulations, the information that a person belongs to a certain minimal group should not change a subject's preferences to favor that person. Following this assumption of stable preferences, any discrimination in minimal groups should result from a shift in beliefs. In other words, if preferences did not trigger discrimination, for discrimination to occur, minimal group manipulations should be at least sufficient to trigger unfavorable beliefs about the characteristics of a certain minimal group. But one may wonder whether this is true. Since an arbitrary group distinction does not provide any information about the group's characteristics, a rational subject should not alter her beliefs about a person either when learning that

¹In theoretical economics, the model of statistical discrimination (Arrow, 1973; Phelps, 1972) describes the idea that people discriminate based on their beliefs about the average characteristics of a person's group. This form of discrimination cannot occur in minimal groups under full rationality. However, actors may behave irrationally in the formation of their beliefs and mistakenly perceive differences in group averages. In so far, the role of beliefs in minimal groups would follow the intuition of statistical discrimination, but is not perfectly captured by the rational models of statistical discrimination.

²The model of taste-based discrimination (Becker, 1957) says that people simply prefer their in-group over their out-group and, thus, discriminate against out-groups.

he or she belongs to an arbitrarily formed (out-)group. However, it is known that people do not behave perfectly rational. If people have often enough—and/or in formative periods of life—observed their out-group to be less reliable and honest, they may form a general mistrust against out-groups believing that any out-group is on average less trustworthy than an in-group. Then, the mere confrontation with an out-group may suffice to trigger discriminating beliefs. Consequently, this paper studies whether discriminating beliefs serve as the motivation behind the puzzling observation of minimal group discrimination.

Discrimination based on mistakenly perceived differences in group averages (akin to statistical discrimination) is fundamentally different from discrimination arising from a conscious dislike for a certain group described as taste-based discrimination. In the case of the minimal group paradigm, the existence of discriminating beliefs would imply that marginal degrees of social distance per se are sufficient to trigger prejudices. To be precise, this paper studies the existence of discriminating beliefs asking whether subjects believe that the out-group is on average more dishonest than the in-group. Notwithstanding, believing in others' dishonesty is only one facet of prejudice that may be held against a group. However, dishonesty seems to be one of the major channels that lead to discrimination, as it fundamentally undermines subject's trust in people. If subjects in minimal groups expect more dishonesty from their out-group than from their in-group, it would suggest that in-group favoritism, respectively out-group discrimination, is already so profoundly embedded in people's minds, that they may consider their in-group as more trustworthy irrespective of whether the group distinction is informative about people's honesty behavior. Our results, however, do show that subjects do not discriminate against an out-group when stating their beliefs regarding others' honesty. Since discriminating beliefs seem to play no role, we could conjecture that discrimination observed in minimal groups should be rather taste-based motivated, meaning that subjects would simply prefer one minimal group over the other.

Since purely minimal groups are not existent in real-life situations, one has to study this question with a laboratory experiment. A design is applied where subjects were randomly assigned into either of two arbitrary groups with different labels (blue or yellow).³ In both groups, subjects drew in private from an urn filled with two types of marble (white and speckled) and report the outcome. As only the report of the speckled marble is rewarded, subjects faced monetary incentives to behave dishonestly in their report. Second, subjects stated their beliefs about how many subjects report the profit-maximizing speckled marble type. The treatments only

³The way of inducing group identity here is different from Tajfel et al. (1971), but it still fulfills their criteria of a minimal group.

differed in the belief elicitation: subjects were asked to assess the behavior of either their in-group (baseline) or of their out-group (treatment). Monetary incentives for stating beliefs accurately were given. Further, it is tested how many subjects behave dishonestly. As individuals cannot be observed lying in experiments without lifting their privacy, a design idea by Djawadi and Fahr (2015) is applied to detect dishonesty at group level. In both treatments, the vast majority (65.38% and 69.23%) of the participants dishonestly reported the profit-maximizing speckled marble.

To the best of my knowledge, the motivational role of beliefs in minimal-group discrimination has hitherto not been investigated in an experiment. Lane (2015) makes an approach to provide evidence for discriminating beliefs in his metaanalysis. In this analysis, he compares experiments where discrimination can be motivated by both beliefs and preferences to those where discrimination can only be explained by preferences⁴. Ergo, the role of discriminating beliefs is measured as the difference in discrimination rate between these types of experiments. Further, his meta-analysis involves experiments which differ in design with respect to the chosen games and the group identity level (natural and (near-) minimal group). In comparison, this experiment directly measures the existence of discriminating beliefs in minimal groups and, thus, provides unconfounded evidence on their influence on minimal group behavior.

Cohn et al. (2014) study dishonesty among bank employees. They investigate whether subjects hold discriminating beliefs about the honesty of bankers and other certain professional/social groups among all investigated groups. They found that subjects from general population believe bankers to be the most dishonest group. However, the authors do not elicit bankers' beliefs about certain professional/social groups. Instead, they study whether bankers' beliefs about other banker's honesty is altered when increasing the salience of their professional identity. They do not find any effect. This finding somehow points to the idea of this paper. However, given the strong professional reference to bankers, any conclusion on that matter would be rather vague and can never be generalized. Therefore, this paper studies whether inter-group discrimination occurs irrespective of any professional/social group identity.

Further, this paper adds to the literature studying the relationship between individual lying behavior and beliefs about others' dishonesty. Our findings suggest a positive

⁴According to Lane (2015), discrimination which occurs despite the fact that the opponent's behavior has no impact on the final outcome, as for dictators and trust game returners, can only be motivated by preferences, since any belief about the opponent's type is irrelevant.

correlation between individual lying behavior and the expectation about others' lying behavior. Dishonest reporters expect significantly more profit-maximizing report than honest reporters. This is in line with previous studies (Abeler et al., 2014; Conrads, 2014). To the best of my knowledge, this paper is the first that studies the relationship between individual lying behavior and beliefs about others' dishonesty under an incentivized belief elicitation procedure which enhances the predictive power of our study. Previous studies (Abeler et al., 2014; Conrads, 2014) do not provide any incentives for stating beliefs accurately.

Studies which introduce an incentivized belief elicitation procedure do not test whether the individual behavior of a person is in line with his or her beliefs about others. In Rauhut (2013), the control treatment is comparable to this experiment when disregarding the group identity. He studies whether an over- or underestimation of others' dishonesty induces people to adjust their behavior. Whether the individual behavior and beliefs are aligned was not tested. Fischbacher and Föllmi-Heusi (2013) also provide incentives for stating accurate beliefs. In their main experiment, subjects roll a die and report the outcome. They added a followup study asking subjects to state their beliefs about the reporting behavior of the participants in the previous experiment. That procedure does not allow testing whether subjects' beliefs are in line with their own behavior.

Further studies which employ incentivized belief elicitation procedures let subjects engage in strategic games, such like trust game, after having stated their beliefs about their opponent's behavior in the following game (e.g. López-Pérez and Spiegelman, 2013; Butler et al., 2012). According to Blanco et al. (2010), such designs may cause risk-adverse subjects to state inaccurate beliefs in order to hedge against low payoffs in the game part. The experiment presented in this paper rules out such bias: when reporting the marble outcome, no assessment of others' behavior is required, as the subjects are paid off according to their individual behavior. The behavior of others does not affect subjects' payoff. This rules out incentives to strategically misstate beliefs about the others' behavior later as a payoff insuring mechanism for previous wrong assessments.

The remaining paper is organized as follows: Section 2 presents the experimental design. The hypotheses are derived in Section 3. Section 4 describes the experimental procedure and descriptive statistics. Section 5 presents the results followed by the discussion in Section 6. Section 7 concludes.

2 Experimental Design

This paper studies whether people hold discriminating beliefs against their outgroup under a minimal group setting. Further, we test whether subjects' individual dishonesty behavior is correlated with their beliefs about others' behavior. Testing the later hypothesis demands an approximate identification of who is a potential liar. Unfortunately, allowing subjects to act in private prohibits to detect lying at the individual level. At best, the individual reporting behavior can be used as an approximate identifier whether an individual behaved honest or dishonest. First of all, a clear definition for who is a liar is needed. This can be achieved best with a dichotomous payoff structure where subjects receive a payoff for reporting the success outcome or receive nothing for reporting the loss outcome. Under the assumption that dishonest subjects are more likely to be among the successful reporters, subjects who report the loss outcome can be considered to be honest. In this study, subjects are asked to privately draw a marble out of an urn filled with white (loss) and speckled (success) marbles and report the outcome. Most experiments studying dishonesty in the laboratory are dice experiments (e.g. Fischbacher and Föllmi-Heusi, 2013; Rauhut, 2013). These experiments allow for different degrees of dishonesty where one may not necessarily report the best outcome, but still behave dishonestly by reporting better paid outcomes than their actual one, even though they do not go for the profit-maximizing outcome (partial lying). Of course, under a die task, an approximation is still possible by assuming that liars are more likely to be among the "five" (best paid) or "four" (second best paid) reporters, while "six" reporters (worst paid) are very likely to behave honestly using the payment scheme given in Fischbacher and Föllmi-Heusi (2013). However, for moderate reports such statements are rather difficult and rather arbitrary, as the incentives for lying are still present but less clear: some subjects may tell the truth, while some reporting moderate numbers may be partially lying, as their true outcome may have been even lower. Consequently, using this identification strategy would lead to a loss in observations, as the moderate reports cannot be reliably used for an approximate identification of liars. For this reason, the author purposely decided against a die task. Other papers (e.g. Abeler et al., 2014) use coin flip tasks which would also satisfy our demand for a binary device. Nevertheless, the idea of drawing a marble has a useful feature. Gathering the drawn marbles allows observing lying at group level by comparing the number of returned speckled and white marbles to the reported numbers in a group. This idea was borrowed from Djawadi and Fahr (2015).⁵

⁵They run an experiment where subjects draw marbles in private facing incentives to claim the draw of a red marble which entitles them to participate in a raffle. In order to measure the rate of lying in the experiment, they collect the drawn marbles, count the total amount of red marbles

Our experimental process followed four steps: the group allocation, tournament part, individual part and a survey.

i. Group allocation

Subjects are randomly assigned to one of two groups: either to the blue or to the yellow group, each group consisting of 13 participants. Randomization is assured by telling subjects to draw a number which directs them to a certain seat. The seats are prepared in such a way that the room is divided into two symmetrical groups. Subjects who draw a number directing them to a seat on the left side of the room are assigned to a group called the blue group. Subjects with a seat located at the right side belong to the group called the yellow group. This way, subjects of the same group sit together. This makes the group distinction visible for the subjects. Corresponding to their group's name, the subjects receive a wristband in either blue or yellow. The group color also appears constantly at the top of the subjects' computer screens throughout the experiment to strengthen the group identification.

ii. The Experiment: Tournament part

The main experiment consists of two parts: a tournament part and an individual part including the drawing of marbles and the belief elicitation. Our data set is based exclusively on observations gathered in the individual part. The tournament part has been introduced only for technical reasons. First, it helps to create a competitive atmosphere which serves as an additional in-group-out-group distinction. Findings from previous experiments suggest that the salience of identity plays a great role in the effectiveness of induced identity. For example, Eckel and Grossman (2005) study the effect of identity on cooperative behavior varying the strength of group identity. They find that "just being identified with a team is, alone, insufficient to overcome self-interest. [...] that actions designed to enhance team identification contribute to higher levels of team cooperation." Among others, they set up a treatment where a tournament is implemented to create an in-group-out-group conflict: group effects on cooperation are among the largest in the tournament treatment. Similar findings on increased group effects under enhanced group identification are reported in Charness et al. (2007) and Chen and Chen (2011). Based on these evidence, beliefs about honesty of the in- and

and compare them to the amount of handed-in tickets to the raffle which claim the drawn of a red marble.

out-group are expected to differ more when using this enhanced design instead of a pure minimal group setting alone due to stronger group identity. Adding to this, the introduction of the tournament part serves a second purpose. It prevents a demand effect, since the author is concerned that the treatment would be quite obvious when running the individual part alone. As the tournament part is also used as a preparation for the individual part, it takes place before the individual part.

As a first task of the tournament, subjects are given a mathematical text problem to solve.⁶ The solution reveals the amount of speckled marbles inside a hose that contains the marbles for the individual part. Throughout the tournament part, the hose is physically placed but covered on a table in front of the participants. Subjects submit their answers to the text problem individually, communication between the group members is not possible. Competition between the groups is induced by rewarding only the winner group which is the group with most correct answers. In the winner group, every member receives a payment of $1.50 \in$ —independent of the individual performance. If there is a draw in correct answers, no group is paid for simplicity reasons. After the subjects have typed in their answers at the computer screens, the hose is uncovered, lifted in the air and the speckled and white marbles are counted publicly. The correct composition is 80 white marbles and 40 speckled marbles.



Figure 3.1: Hose with 80 white and 40 speckled marbles

The next task in the tournament is to estimate how many out of 12 subjects have solved the story problem correctly and, thus, guessed the correct amount of speckled marbles inside the hose. In the out-group treatment, estimation is made on the other group's members; in the in-group treatment, it is elicited on the own group's members. The payment scheme of this task is identical to the one of the

⁶The text problem and some design challenges are described in the appendix.

former task: only the winner group with most correct guesses, receives a payment of $1.50 \in$ for each group member. In case of a draw there is no payment.

Subjects receives their payoff and the information on who is the winner of the tournament tasks at the end of the experiment. This way, we can rule out wealth effects driven behavior in the individual part. Second, confounding in the beliefs due to different strengths of group identification within winner and loser groups is ruled out. Nevertheless, the pure experience of working in teams causes the group identity to be more salient, as the tournament makes subjects more conscious of their team identity. This procedure minimizes a potential over-time deterioration of group identification under arbitrary group allocation as observed in Chen and Li (2009). A comparable procedure is used in Chen and Chen (2011) where subjects also receive information on their payoffs generated during the group enhancing part at the end of the experiment.

iii. The Experiment: Individual part

The first task of the individual part is the marble drawing task. As a preparation, subjects are shown the transparent hose which is filled with 80 white marbles and 40 speckled marbles. Counting the marbles publicly ensures that this distribution is known to the subjects. Next, the marbles from the hose are poured into an empty urn which is non-transparent and had its opening covered by a cloth to prevent subjects from viewing inside while drawing. After preparing the urn, all subjects draw from this same urn. As the groups are spatially separated, the urn is taken through the seat rows in such a way that no group is served first. Subjects are told to draw only one marble and to be careful to hold the marble in the palm of their hand, so that no one but them can ever see the marble. These instructions are repeated in order to make subjects aware that the actual type of their marble is only known to themselves. After all subjects have drawn, they mark the type of their marble at the computer screen. For reporting a speckled marble, the individual receives $6 \in .^7$ The report of a white marble generates no money. Since the actual marble type is private information to the subjects, they have a strong incentive to misreport their type in order to generate a higher payoff. Next, the marbles are recollected. Djawadi and Fahr (2015) have developed a mechanism to observe lying at group level. Their idea is integrated into the design of this experiment: the drawn marbles are recollected in empty urns separately for each group. This allows us to check the actual number of white and speckled draws in each group and to

⁷Payoffs generated during the experiments were expressed in the artificial currency ECU and converted into euros (10 ECU = $0.30 \in$) at the end of the experiment.

compare them to the reports. Lying is thus observed at group level as the deviation of reported number of speckled marbles from the actually returned number.

The last task of the individual part is the belief elicitation task: subjects estimate how many out of 12 people have reported a speckled marble. The answers are used as a measure of subjects' belief about the number of liars. This interpretation can only be made if all participants are able to compute the statistically expected number of "speckled" reports under complete honesty. After reading the instructions, subjects answer some comprehension questions to ensure that all participants have properly understood the instructions.⁸ Among others, they are given an example for which they calculate the expected number of "marbled" reports under the assumption that all subjects report honestly. The treatment variation was induced by either assessing the out- or in-group's reporting behavior. As the answers are checked for correctness, all subjects are expected to know the statistically expected number of "speckled" reports under complete honesty. Regarding the belief elicitation task, in the treatment (out-group treatment), subjects make a guess on the members of the other group. In the baseline treatment (in-group treatment), subjects guess the "speckled" reports of the fellow members in their own group. Since everybody knows their own report, there are only 12 subjects left to assess in the in-group treatment. To avoid confounding, subjects in the out-group treatment are told to make an estimate based on 12 randomly chosen people of the out-group.⁹ Beliefs are purposely elicited in absolute terms, as stating beliefs in relative shares is less intuitive and may induce subjects to misstate beliefs due to wrong calculation or understanding. For guessing correctly, the subject receives an individual payment of $6 \in$. The reader may ask whether subjects state their correct beliefs, as subjects may have an incentive to deviate in order to justify their own reporting behavior. However, in an anonymous setting justifying oneself can only be motivated by self-image concerns and only if subjects can sufficiently suppress their true beliefs through their inaccurate report. The author doubts whether this is possible at all. Moreover, self-image concerns that lead to a distorted guess should not play a role in this setting. The relative high incentives of $6 \in$ for a correct guess makes the

⁸Subjects are unable to proceed unless all comprehension questions are answered correctly. Clarification questions are answered in private.

⁹The subjects are purposely not asked to guess the number of false reports, ergo liars. As counting the number of returned marbles allows identifying the number of false speckled reports, the experimenter would basically be able to correctly incentivize the subjects for guessing the number of false reports. However, this is not done due to administrative and more importantly confounding reasons: Given that every subject knows whether she or he has been honest, the subjects in the in-group treatment would possess more information than those in the out-group treatment. To avoid confounding, the returned marble of a respective subject has to be singled out in the in-group treatment which is impossible without lifting the privacy, as doing so requires to observe lying on the individual level.

pursuit of such self-image concerns prohibitive. Therefore, the incentive scheme is expected to elicit true beliefs.

iv. Survey

At the end of the experiment, subjects are asked to answer a survey which includes demographic questions and questions concerning the experiment. Among other questions, they are asked to give a truthful report about the marble that they have drawn. They are reminded that any ex post deviation from their initial report made in the experiment does not affect their payment. From their answers, the experimenter can draw some inference about the individual honesty behavior.

v. Linking the reporting behavior to individual dishonesty

It remains a challenge in an experiment with private information to figure out who behaves honestly and who did not. Since observing an individual lying is impossible without sacrificing privacy, economic experimenters can only detect lying at the aggregated level by either using the statistically distribution as a benchmark or-as in our case-by observing lying at group level through counting the number of returned speckled and white marbles. Despite these limitations, inference on a person's honesty behavior can be drawn from her reporting behavior. By assumption, subjects are expected to lie only if it is advantageous to them. Meaning, subjects will rather deviate to the payoff-maximizing option (report "speckled") than to the zero payoff (report "white"). Disadvantageous lying by giving a false "white" report is only rational if a subject wants to signal an honest image to others. Such considerations are irrational under anonymity. Another motive is preserving a positive self-image if subjects follow a behavioral norm, like being modest. Fischbacher and Utikal (2013) run an experiment with nuns and female students. They find that even in absence of any social interactions people may lie to their own disadvantage. In their experiment, evidence falls only on the side of the nuns from a Franciscan community who took a poverty vow. For female students, however, they observe profit-increasing lying instead. Consequently, I argue that disadvantageous lying is rather unlike to happen in our experiment. Assuming full rationality, subjects reporting the minimal payoff option "white" are considered to be honest. Accordingly, liars are believed to be exclusively present among the "speckled" reporters. It is clear to the reader that not all subjects reporting the payoff-maximizing option behave dishonestly. Therefore, "speckled" reports do not perfectly identify liars. Another proxy can be derived from the survey where subjects are asked to give an honest ex-post statement about the marble type they

have actually drawn. Further, the ex-post reports are used to verify our assumption of full rationality, namely that subjects do not report "white" while actually having drawn a speckled marble.

vi. Additional design feature

In each session, 27 participants are recruited. Beside the 26 subjects playing in groups, there is the function of a so-called *independent verifier*¹⁰ who is not allocated to any group and does not take part in the experiment itself. Instead, she is assigned the task of assuring that the experiment was executed as described in the instructions. One may argue whether such an independent verifier is actually needed. The desired implementation is one that assures complete anonymity, especially in the drawing task. Therefore, an additional subject advocating on behalf of the other participants, comparable to a lottery queen or a guarantor who verifies that the rules (e.g. anonymity)¹¹ are kept at best has been assigned. Further, the independent verifier takes over the task of handing the subjects the urns for drawing and recollecting the marbles. This prevents the experimenter from entering the subjects' area during the experiment and preserves a stronger perception of anonymity and group integrity in the author's sight.

3 Hypotheses

3.1 Theoretical Considerations

This paper studies whether subjects hold discriminating beliefs against their outgroup regarding their dishonesty. If so, this may serve as a motivation for subjects to disadvantage outsiders, respectively to advantage their peers. Especially, the existence of discriminating beliefs under a minimal group setting is studied.

Tajfel et al. (1971, p.154) postulate several minimal criteria for the existence of inter-group bias:

¹⁰The independent verifier is randomly chosen from the pool of participants by directing her to the seat number 27. For her participation she receives a fixed amount of $6 \in$. The participants receive printed instructions of the experiment which also explain the function of the independent verifier.

¹¹For example, the independent verifier is in charge of controlling the amount of the white and speckled marbles inside the hose from where the marble are poured into the urn.

- 1. Subjects do not interact face-to-face, be it within their own group or between the groups.
- 2. Group membership is completely anonymous.
- 3. Subjects cannot rationally or instrumentally link their group discriminating behavior to the criteria of group categorization.
- 4. Subjects do not gain any utilitarian value from their choices.
- 5. For every subject there exist other strategies which are more rational and useful than the group discriminating strategy.
- 6. The choices are "made as important as possible to the subjects'."

The design of this experiment fulfills all criteria. Since the experiment focuses on studying beliefs, no interaction between subjects is involved. Group membership is largely anonymous. The identity of group members is not revealed. Nevertheless, anonymity is only imperfectly assured: groups are spatially separated in the room, allowing participants to know that their direct seat neighbors belong to their group. However, the laboratory is equipped with blinds which prevent face-to face recognition of other group members. The third criteria can be fulfilled by a random allocation rule: in our experiment subjects are randomly assigned to one of two groups by drawing a seat number. The formulation of the fourth criteria seems rather ambiguous. One could conclude that subjects' choice should not affect their own payoff (e.g. Chen and Chen, 2011; Chen and Li, 2009; Ostrom and Sedikides, 1992). More light on this criterion is shed in Tajfel and Turner (1986): "there [is not] any rational link between economic self-interest and the strategy of in-group favorism" (p.14). In this sense, our experiment meets this criterion, as there are no economic incentives provided for discriminating any group. If this were the case, we would face problems of confounding: observing any group effect does not allow concluding that group distinction per se is sufficient to trigger discriminating belief, as it is hypothesized in this paper. Such an effect would be rather driven by other economic interests. Further, under the random group classification which is entirely uninformative about people's honesty, discrimination is not even economic. There exists a more rational and also useful strategy which is to not discriminate. This is in conflict with the group discrimination strategy. Thus, the fifth condition is fulfilled as well. When formulating the sixth condition, the authors have aimed for salience of decisions: their idea is to introduce real decisions on the payments of others instead of asking for unincentivized subjective evaluation of others which is likely to be arbitrary. In this experiment, salience is ensured, because subjects' choices are relevant for their own payoff but not for the others' payoff as suggested by Tajfel and Turner (1971). In the author's eyes, referring incentives to one's own

payoff does even increase salience.

Since all criteria for the existence of inter-group bias are fulfilled in our design, we can confidently expect to find discriminating beliefs against the out-group in this experiment if such an inter-group bias exists.

An abundance of literature provides evidence for the minimal group paradigm stating that even in minimal groups subjects tend to discriminate against their out-group, respectively to favor their in-group (e.g. Tajfel et al., 1971; Brewer, 1979; Chen and Li, 2009). Still, it remains unclear whether subjects in minimal groups discriminate because they hold prejudicial beliefs against the out-group and, hence, are convinced that the other group is on average different in some characteristics that matter for honesty, or because they have a preference for out-group discrimination. If prejudicial beliefs serve as a motivation, we should expect that subjects hold beliefs about their out-group's behavior which are more unfavorable than those about their in-group's behavior. There are many facets of prejudice that a person can hold against a group. However, the belief that an out-group is less honest than one's own in-group is one major channel that could cause subjects to discriminate, since dishonesty undermines subjects' trust in people and may lead them to draw inference to other unethical behavior. For example, a sender in a trust game may easily fear to be exploited if she or he expects the receiver to be a rather dishonest person. Therefore, if discriminating beliefs serves as a motivation, it should be reflected in increased mistrust against the out-group: subjects should expect more dishonesty from their out-group than from their in-group.

3.2 Hypotheses

According to this paper's interest of research and in line with the theoretical considerations above, the central hypothesis is the following:

Hypothesis 1. Subjects expect more dishonesty in their out-group than in their *in-group*.

Findings from other experimental studies (e.g. Abeler et al., 2014; Conrads; 2014) have suggested a positive relationship between beliefs and individual lying behavior. This relationship also seems relevant for inter-group discriminatory behavior in trust, especially if subjects tend to project their own behavior more strongly onto those subjects who they feel more familiar with, namely their in-group. As a result, honest people may discriminate differently than dishonest individuals. Then, our central hypothesis may not hold whenever subjects' beliefs about others' honesty

is interacting with their own honesty behavior.

Therefore, the following hypotheses need to be added:

Hypothesis 2. Given that the subject is honest, she expects more dishonesty in her out-group than in her in-group.

Hypothesis 3. *Given that the subject is dishonest, she expects less dishonesty in her out-group than in her in-group.*

Due to the counteracting character of these hypotheses, it is important to distinguish between different honesty behaviors. If honest subjects assume more honesty inside their group than outside, while dishonest subject expect more lying within their group, those effects may outweigh each other and, thus, a discriminating belief cannot be detected at the overall level.

A necessary assumption for the experiment is that subjects misreport the outcome from the marble draw.

Assumption 1. A substantial share of subjects in both treatments is expected to behave dishonestly.

Further, subjects are expected to anticipate that others behave dishonestly. As the same monetary incentives are applied to subjects in both treatments, it is assumed that the belief in dishonesty is omnipresent in both treatments.

Assumption 2. A substantial share of subjects is predicted to expect dishonesty in the in- and out-group.

4 Procedure and Descriptive Statistics

The experiment was run at the BaER-Lab of the Paderborn University in July 2015. Apart from the independent verifiers, 52 students were recruited, subdivided into two consecutive sessions of 26 participants.¹² All participants were recruited via ORSEE (Greiner, 2015). The sessions differed with respect to the treatment and were run the same morning, the in-group treatment directly followed by the out-group treatment. 46.15% of the participants were female. Approximately 42% were enrolled in the field of business administration or economics. The remaining

¹²Due to her special role, no data points were collected for the independent verifier, and she was not counted among the subjects.

students studied in different fields, such as engineering, computer science and education science. Table 3.1 provides the summary statistics. Randomization between the in- and out-group treatment has been successful. Female participation rate and the average age were identical in both treatments.

The experiment was programmed using the software z-Tree (Fischbacher, 2007). Subjects received printed instructions, but instructions were also recalled on the computer screens during the experiment. Payments during the experiment were expressed in ECU and converted into euros at the end of the experiment: 10 ECU accounted for $0.30 \in$. Additionally, the participants received a fixed participation fee of $2.50 \in$. On average, subjects received total payment of $9.51 \in$ which was handed out immediately after the sessions. Finally, subjects filled in a survey eliciting socio-demographic questions and questions concerning the experiment. Each session took between 61-65 minutes excluding the time for paying participants.

	Out-group Treatment	In-group Treatment	Total
Participants	26	26	52
"Speckled" reports	23	23	46
Counted speckled marbles	6	5	11
False "speckled" reports	17	18	35
Recalled "speckled" reports (survey)	13	16	29
Female participants (in%)	46.15	46.15	46.15
Final Payoff in €	9.77	9.25	9.51
Average age	23.58	23.58	23.58
Std. dev. (age)	3.18	4.92	4.10

Table 3.1: Summary statistics

5 Results

In this section, the results are presented. First, we compare the actual distribution of speckled marbles that have been drawn with the reports of speckled marbles to find out whether subjects behave dishonestly in the experiment. Second, we study whether subjects differ in their beliefs about others' lying behavior across the treatments in general and conditional on their own behavior in the experiment. Then, we test whether subjects' beliefs are generally in line with their own behavior as the existing literature suggests. At last, given strong evidence that men behave more dishonestly than women (e.g. Friesen and Gangadharan, 2012; Abeler et al., 2014), some results on gender effects are presented.

5.1 Belief Elicitation in In- and Out-groups

Lying is very attractive in this experiment, since the participants face strong incentives for reporting the draw of a speckled marble. Adding to this, lying does not involve any externalities in our setting. Costs of lying occur solely to the liar herself in non-monetary form, e.g. bad consciousness. However, the extent of lying in this experiment is surprising. We detect lying at group level, since the returned marbles are counted separately in the blue and yellow groups. In all groups, at least 11 speckled marbles (84.62%) were reported, while at most 3 speckled marbles (23.08%) are counted. See Table 3.4 for a detailed comparison of the groups. The behavior at group levels is merely of interest and, thus, data from yellow and blue groups is pooled at treatment level since the yellow and blue groups do not differ with respect to important features (the empirical distribution of speckled marbles, the distributions of the reported outcomes, of the stated beliefs and of the revocations made in the survey). The respective test results are presented in Table 3.4 of the appendix.

In both treatments, 23 out of 26 participants (88.46%) report the payoff-maximizing speckled outcome as shown in Table 3.1, while only 6, respectively 5, speckled marbled are counted. This shows that a substantial fraction of subjects has lied in both treatments. Indeed, if we assume that subjects do not lie at their own disadvantage by falsely reporting a "white" marble, 17 out of 26 subjects (65.38%) in the out-group treatment and 18 out of 26 subjects (69.23%) in the in-group treatment have been dishonest about their draw.¹³

In order to elicit beliefs on others' lying behavior, subjects are asked to assess the reporting behavior of either the 12 other members of their group (in-group treatment) or 12 randomly chosen subjects from the other group (out-group treatment). In the absence of lying, the expected number of reported speckled marbles can be computed using the statistical distribution. Here, an individual believing that all 12 members are completely honest should expect 4 reports of speckled marbles.

¹³Djawadi and Fahr (2015) find that approximately 32% of the subjects do not tell the truth about their outcome.

Note that participants have been informed about the amount of speckled and white marbles in the urn and, thus, are able to compute the statistically expected number under complete honesty.¹⁴ Therefore, any deviating belief from the statistically expected number signifies the expectation of lying. The larger the stated belief deviates from the statistically expected number, the more liars the subject expects.

As shown in Table 3.2, the expectation of lying is prominent in both treatments. Using a Wilcoxon signed-rank test, the median expectations of lying are significantly different from 4 (p=0.000 for both treatments).¹⁵ This also holds true when controlling for the individual reporting behavior as seen in Table 3.2. In both treatments, subjects believe that others behave dishonestly.



Figure 3.2: Aggregate beliefs in the Out-group and In-group treatment

The research interest in this paper is to show whether the beliefs of subjects about others' honesty behavior differ substantially when assessing the behavior of an out-group instead of an in-group. Figure 3.2 reports the aggregate beliefs of sub-

¹⁴To facilitate the subjects' understanding, some respective comprehension questions are asked prior to the experiment which help to calculate the statistically expected number under complete honesty.

¹⁵In the comprehension questions, subjects are given an example where 6 "speckled" reports are expected under complete honesty. In case that this example serves as an anchor, the median expectations are also tested against the number of 6. For all subgroups but the honest, the null hypothesis of the Wilcoxon signed-rank test is significantly rejected.

jects in the out- and in-group treatment. In both treatments, most participants (12 of 26 subjects in the in-group treatment and 11 of 26 subjects in the out-group treatment) expect that all subjects report the profit-maximizing option. However, a first rough visual analysis of Figure 3.2 suggests some small differences in the distribution across the treatments. Additionally, we may also suspect the variance of the distributions to differ, since the minimal quantity of "speckled" reports that is expected is 3 in the in-group treatment (1 subject), while it is 6 in the out-group treatment (2 subjects).

Despite the evidence that lying is broadly expected, tests for equality in distributions indicate that there is no significant difference in beliefs between in- and out-group treatments. The results from the two-sample Wilcoxon rank-sum (Mann-Whitney) test (MWW) suggests that the treatments do not differ with respect to the median beliefs about the number of profit-maximizing reports (p>0.95; see Table 3.2). Following the visual analysis of Figure 3.2, a two-sample Kolmogorov-Smirnov (KS) test for equality in variance is added. In none of the cases, the null hypothesis can be rejected. The beliefs do not seem to differ across the treatments.

As already mentioned an excess proportion of subjects in both treatments (46.15%, respectively 42.31%) expect all 12 subjects to report "speckled". Those people may differ from people who indicate a moderate number for various reasons. For example, they may be more mistrusting and pessimistic than others, believing that they live in a world where a person has to make use for any chance. Thus, they may presume that all subjects behave profit-maximizing, even if this requires dishonesty. I argue that such attitude is rather independent of the treatment: these people state beliefs of complete dishonesty in both treatments, even though they may hold slightly discriminating beliefs. If so, it is not surprising that no treatment effect is observed given the dominant proportion of people expecting complete dishonesty. Consequently, the MWW test is rerun without those subjects who have indicated to expect 12 "speckled" reports. The result is in line with the former one.¹⁶ Therefore, Hypothesis 1 cannot be supported. Subjects in the out-group treatment.

Result 1. The belief about the in-group's lying behavior is not different from the belief about the out-group's lying behavior.

From the results so far, we could argue that people do not hold discriminating beliefs against their out-group nor their in-group in terms of honesty. However, it

¹⁶The analysis is based on 29 subjects, 14 in the in-group treatment and 15 in the out-group treatment.

	Out-group Treatment		In-group Treatment		Tests for equality of distribu- tion: Prob > z		
					Kolmogorov- Smirnov: equal variances	Wilcoxon Ranksum: equal medians	
	Median ¹⁾	Obs.	Median ¹⁾	Obs.			
all	11***	26	11***	26	1.000	0.954	
white (honest)	6	3	6	3	1.000^{2}	0.637 ²⁾	
speckled	11***	23	12***	23	1.000	0.906	
dishonest	12***	13	11.5***	16	1.000	0.454	
male	11***	14	12***	14	0.635	0.295	
female	11^{**}	12	10^{**}	12	0.536	0.194	

Table 3.2: Median values of expected "speckled" reports in treatments

Note: ¹⁾ Medians differ from the statistically expected number of 4 in most subgroups. Significance levels of the Wilcoxon signed-rank tests are reported behind the respective medians: * significant at p<0.10; ** significant at p<0.05; *** significant at p<0.01²⁾ There are not enough observations (n=6) for any reliable analysis.

may be the case that honest subjects assume more honesty inside their group than outside, while dishonest subject expect more lying within their group. Assuming these counteracting beliefs, it is not surprising to see no treatment effect at the overall level. Consequently, we next test whether the individual lying behavior affects people's beliefs as stated in Hypothesis 2 and 3. For this purpose, the data set is split into subsets which are generated based on subjects' reporting behavior. For example, subjects who have reported "white" are believed to have made an honest report. Thus, the subset consisting only of "white" reporters is called the subset of honest subjects.

Following Hypothesis 2, it is tested whether honest people expect less lying in their in-group than in their out-group. Looking at the mean expectations, honest subjects reporting a white marble in the out-group treatment seem to believe in more payoff-maximizing reports (7 reports expected on average) than the honest reporters of the in-group treatment (6 reports expected). However, the median values (6 reports) do not differ across the treatments, as shown in Table 3.2. Unfortunately, lying is so prevalent in our experiment that only a total of 6 people reporting a white marble, 3 subjects in the control group and 3 subjects in the treatment group, is observed. The scarcity of observations does not allow drawing final conclusion. Therefore, Hypothesis 2 cannot be tested in this study. However, if one could rely on these observations, they would suggest that subjects behaving honestly do not differ in their beliefs across the treatments.

Next, we test whether dishonest people have different beliefs about others' honesty across the treatments. Since dishonesty of subjects cannot be observed directly, the reporting behavior of subjects is used as a first approximation to dishonesty. The vast majority of subjects reporting "speckled" have behaved dishonestly, as the number of "speckled" reports exceeds the number of returned speckled marbles by approximate 4 times in both treatments. Thus, the subset consisting only of "speckled" reporters is used as an approximation for the subgroup of dishonest subjects in this experiment. For this subset of "speckled" reporters, no treatment effect is observed (MWW: p>0.90; KS test: p=1.000; see Table 3.2). Subjects of out-group treatment who have reported "speckled" do not seem to indicate higher expectations of payoff-maximizing reports than the "speckled" reporters of the in-group treatment.¹⁷

Result 2. The belief of those subjects who have reported the payoff-maximizing option is independent of the treatment.

¹⁷The histograms of the subset of "speckled" reporters are presented in the appendix (Figure 3.5).

Since there are also honest "speckled" reporters, an effect may not be captured sufficiently by applying the report type as a proxy for dishonesty. Therefore, the identification of dishonest people must be tracked down further. In the survey following the experiment, the participants are asked to give a truthful ex-post report about their marble. 13 subjects of the out-group treatment and 16 subjects of the in-group treatment deviated from their former payoff-maximizing report made in the experiment. These numbers closely correspond to the numbers of excess "speckled" reports (see Table 3.1). There has been no revocation of "white" reports. This underlines the author's assumption that lying solely occurs among the "speckled" reporters. The author has no reason to doubt the truthfulness of any revoked statement. A deviation from the initial "speckled" report is only rational if subjects want to appear smart by having explored their opportunities. But these image-concerns are not applicable in an anonymous setting, since the experimenter cannot identify a certain subject as a deviator and, thus, creating an image of being smart is not possible.

Based on their revocation, a new more accurate subset of dishonest subjects is generated which includes only those subjects who reveal their dishonesty by later stating in the survey that they have actually drawn a white marble. Subjects who persist on their "speckled" report cannot be classified reliably given that they could be honest or dishonest, as the number of revocations do not perfectly match the excess quantity of "speckled" reports meaning that not all have admitted their dishonesty. The reader should be aware that the identification of dishonesty used here relies on softer data than the report type which is incentivized. As mentioned above, a revocation of a "speckled" report is assumed to be reliable, but we cannot tell with certainty.

The results from the MWW test shows no evidence for a treatment effect among the subset of dishonest subjects (MWW test: p>0.45; KS test: p=1.000; see Table 3.2). Dishonest subjects of the out-group treatment do not make significantly different assessments than do the dishonest ones of the in-group treatment.¹⁸ Therefore, Hypothesis 3 cannot be supported.

Result 3. The belief of those subjects who have revealed their dishonesty is independent of the treatment.

Summarizing, there is no evidence that subjects hold discriminating beliefs against their out-group when assessing their lying behavior. This seems true irrespectively of the individual honesty behavior.

¹⁸The histograms of the subset of dishonest reporters are presented in the appendix (Figure 3.6).



Figure 3.3: Aggregate beliefs of "white" and "speckled" reporters

5.2 Belief Elicitation based on Individual Lying Behavior

Evidence from previous studies suggests that people's belief about others' honesty is line with their own lying behavior. Since there is no evidence for any treatment effect, we can confidently pool data from the treatments. This allows us to test whether subjects' belief about others' lying is related to their own behavior, as suggested in the literature.

Figure 3.3 suggests that subjects of different reporting behavior differ in their beliefs about others lying. A Kolmogorov-Smirnov (KS) test reveals that the distribution of the beliefs stated by the "speckled" reporters is significantly different from the "white" reporters' distribution (p=0.001, Table 3.3). Using a MWW test, we can assume a difference in median expectations of lying (p=0.0002). Note that the MWW test is actually not an appropriate test in this case, since the MWW test relies on the assumption of equal variance of the samples which is rejected by Kolmogorov-Smirnov test. Nevertheless, the results of the MWW are listed, since it is the most commonly used test for equality in distribution for small sample size. Running a right-censored Tobit model confirms our results (Table 3.5 estimation (5), appendix). The reader should take the results of the Tobit regression with caution, because the small sample size may not meet the necessary assumption of normal distribution.¹⁹

¹⁹A censored least absolute deviations estimation may serve as a more appropriate regression

characteristics	equal 1 ¹⁾		equal 0		Tests for equality of distribu- tion: Prob > z	
					Kolmogorov- Smirnov: equal variances	Wilcoxon Ranksum: equal medians
	Median ²⁾	Obs.	Median ²⁾	Obs.		
speckled (=0: white)	11.5***	46	6**	6	0.001***	0.0002***
dishonest	12***	29	6**	6	0.000***	0.0002***
male	12***	28	10^{***}	24	0.216	0.148
out-group: male	11^{***}	14	11^{**}	12	0.853	0.8717
in-group: male	12***	14	10^{**}	12	0.102	0.0366**

Table 3.3: Median values of expected "speckled" reports by characteristics

Note:

* significant at p<0.10; ** significant at p<0.05; *** significant at p<0.01

¹⁾ "equal 1" denotes that the subgroup is formed based on the respective characteristic given in the first column; "equal 0" is the comparison group formed by those subjects who do not possess the respective characteristic.

"Out-group: male" equal 0 denotes females in the out-group treatment.

²⁾ Medians differ from the statistically expected number of 4 in most subgroups. Significance levels of the Wilcoxon signed-rank tests are reported behind the respective medians.

Next, we narrow the subset of "speckled" reporters by using dishonest reporters (those who have admitted their dishonesty in the survey). Then, the subset of the dishonest reporters is tested against the subset of honest ("white") reporters. The aggregate beliefs are displayed in Figure 3.4. Using a KS test reveals that the distribution of the subset of dishonest reporters is significantly different from the distribution of the "white" reporters' subset (p=0.000, Table 3.3).²⁰ The result from the Tobit regressions (Table 3.5 estimation (6), appendix) confirms this finding.

Result 4. The belief about others' lying behavior is in line with the individual behavior: "speckled" reporters expect significantly more subjects to report the profit-maximizing option than "white" reporters. The difference is even more pronounced when controlling for the individual honesty behavior instead of the reporting behavior.

model (Powell, 1984). However, in our case, this model leads to convergence problems which may be due to the relatively heavy censoring in our data set.

²⁰The result from the MWW test points in the same direction (p=0.0002): "white" reporters expect a significantly lower median number of subjects reporting the payoff-maximizing option (6 reports) than the dishonest reporters (12 reports).


Figure 3.4: Aggregate beliefs by subjects' honesty behavior

5.3 Further Indications

In Friesen and Gangadharan (2012) and Abeler et al. (2014), it is shown that females tend to lie less than males. Given that people's beliefs seem to be line with their own behaviors, females may also expect less lying than males.

Our results do not seem in line with the above. In our experiment, females lie as much as males; they do not significantly report the payoff-maximizing option less often than men²¹. Further, there is no evidence for discriminating beliefs about others' honesty when controlling for gender (Table 3.2). Nevertheless, the null hypothesis of equal medians is significantly rejected when testing for gender differences in the in-group treatment (MWW test, p=0.0366 in Table 3.3) indicating that, in the in-group treatment, females expect significantly less "speckled" reports than males. This points to the idea that females may trust more in their peer's honesty than males. Note that the KS test does not suggest inequality in distributions (p=0.102). Therefore, another explanation could be that females may not per se differ in their beliefs, but may be more conservative when stating their beliefs about peer's honesty behavior and, thus, indicate lower numbers of excepted "speckled" reports in the in-group treatment.

 $^{^{21}}$ Almost 86% of the males and almost 92% of the females report a speckled draw (MWW test: p=0.5071). Neither is there a gender difference in revoking reports. The results are available on request.

Based on these results, one could argue that it might be less socially acceptable for females to mistrust their peers than for males. Gender specific characteristics seem less pronounced in the case of mistrust against the out-group. Nevertheless, given the relatively small sample in this study, these findings on gender cannot be generalized. Further research based on larger, more representative data is needed to confirm whether females exhibit more trust towards their peers than males.

Result 5. In the in-group treatment, females expect fewer people to lie than males.

6 Discussion

6.1 The absence of discriminating beliefs

No evidence for the existence of discriminating beliefs in minimal group setting is found in this study. In the author's eyes, the null result does not stem from a lack of group identity. Other studies successfully show that subjects discriminate even in minimal groups. According to Lane (2015), discrimination under artificially induced group identity is relatively strong and not affected by the method of creating group identity (e.g. preference elicitation or labeling). Given that the method of labeling in this experiment is enriched by an inter-group competition which strengthens group identity, evidence for discrimination should have been found if group division triggers discriminating beliefs.

Rather, the null result could be explained by the relative high acceptance of lying in this experiment: control questions from the survey reveal that many subjects do not classify a false payoff-maximizing report as a moral offense, as it is usually the case. Instead, they seem to regard dishonesty as a legitimate way of making money in this experiment.²² This can be explained by inter alia the relative low costs of lying. Other studies, where lying involves higher costs, find lower levels of dishonesty. Djawadi and Fahr (2015) observe that approximately 32% of the subjects lie which is less than half of the dishonesty rate in the experiment presented here (65.38% in the out-group treatment and 69.23% in the in-group treatment). In their study, negative externalities make lying costly²³ and, thus, less attractive. Adding to this, they use a lab-in-the field experiment where participants face lower impulse to behave dishonestly than in a laboratory experiment where

 $^{^{22}40}$ out of 52 participants say that a person who reports falsely the payoff-maximizing option does not need to feel guilty in this experiment.

²³Dishonesty leads to illegitimate participation in a raffle which reduces the chance of winning for all other subjects who play the raffle.

subjects participate with the predominant motivation to make money, as in our experiment. Other studies where less dishonesty is observed introduce their subjects to additional money-generating tasks. This deteriorates the incentives to behave dishonestly. For example, in a laboratory study of Houser et al. (2012), 74.5% of the participants (approximately 14% points less than in our experiment²⁴) report the profit-maximizing option. In contrast, our experiment presents the marble drawing task as the only opportunity to generate profit with certainty. This entices subjects to make money through lying. In the extreme case of Abeler et al. (2014), no evidence of dishonesty is found at all. Only 44.4% of their participants report the profit-maximizing option in their coin-flip study. The absence of profit-maximizers may surprise at first glance. However, their procedure suggests the implementation of an honesty norm. Prior to the cheating task, subjects have answered questions to their socio-demographic background and risk and trust preferences using questions from the GSOEP and World Value Survey. This introductory phase demands honest answers which seem to have reflected on the cheating task as well. Such an honesty norm is not present in our experiment. For the reasons given above, it is not surprising that lying and the expectation of lying is prevalent in this study. As a result, the differences between the treatments with respect to reporting behavior and beliefs are marginal. It is left for further research to test whether discriminating beliefs are an issue when lying is rather perceived as immoral and less pressing. However, in case of cheap lying (such like daily lies) people do not seem to believe that their peers behave differently than outsiders.

6.2 The interrelation of beliefs and behavior

Without drawing conclusion on causality or specific motives, this study suggests three potential motives that are discussed here: Best guess about others' behavior, the justification of the own behavior and its assimilation to first-order beliefs due to social conformity.

In context of no or marginal information, the best guess about others is to believe that others behave just like oneself. Then, the own behavior may affect the person's first-order belief. Second, subjects may also project the own behavior onto others in order to justify their own behavior (e.g. Abeler et al., 2014). This justification is motivated by the wish to preserve a positive self-image and, thus, should hold

²⁴As they apply a fair flip coin cheating task, the statistically expected percentage of dishonest subjects is approximately 25%, while the observed dishonesty percentage in our experiment is approximately 65%, respectively 69%, which is at least 40% points more than in Houser et al. (2012).

independent of the information level. If subjects justify their dishonest behavior by ex-post stating a higher belief in this experiment, it is worth mentioning that preserving a positive self-image through justification is quite costly in this setting, since the belief elicitation was incentivized with participants receiving a payment only if they have guessed the correct number of persons reporting the payoff-maximizing option. Following this motivation, the non-monetary gains from maintaining a positive self-image must be sufficiently large to compensate for the loss in payoff.

On the other hand, the first-order beliefs could determine the own behavior. When subjects believe in others lying, the wish for social conformity may drive them to behave as they expect others to behave (López-Pérez, 2010, 2012; López-Pérez et al., 2013). Especially, when lying is relatively beneficial and rather seen as a sign of smartness, it turns to be acceptable and may even manifest into a moral norm which most people follow. Since the disobedience of a moral norm involves feelings of guilt and shame, subjects decide to conform, since they expect that others follow the moral norm as well. Further, a social norm is a mechanism which can explain an influence of beliefs on behavior. According to Elster (2009), a social norm is different from a moral norm and requires the observability of a violation and the possibility of being sanctioned, while a moral norm is obeyed out of internal motivation and does not depend up on the observation of others. Following this definition, a social norm cannot be enforced in an anonymous setting. Consequently, if a person's first-order belief determines her behavior, it should be due to a moral norm in this experiment.

The above mentioned motives are equally suitable to explain our findings. Further research is needed to disentangle these motives.

7 Conclusion

This paper explores whether people's beliefs about others' honesty are subject to out-group discrimination. It contributes to the minimal group literature by providing evidence on the motivation behind discrimination in minimal groups. A first approach to this question is made by a meta-analysis of Lane (2015) comparing minimal- and natural-group experiments of different designs. In contrast, the experiment presented here eliminates noisiness from the analysis by providing a direct study on the existence of discriminating beliefs. Our results suggest that subjects in minimal groups do not hold discriminating beliefs with respect to honesty expectation. Therefore, it seems unlikely that discriminating behavior observed in minimal group experiments can be explained by the existence of discriminating beliefs. This result is in line with the study of Lane (2015). He confirms that beliefs may have a rather moderate effect on discrimination in minimal group experiments. Ergo, one could argue that marginal degrees of social distance alone do not seem to trigger prejudice. This would contradict the idea that mistrust against out-groups is already so embedded in people's minds that an arbitrary division in "us" and "them" is enough to discriminate. Rather, the absence of discriminating beliefs leads to the conjecture that subjects in minimal groups may have a preference to favor their own group, respectively to discriminate against their out-group. This would contradict the standard economic assumption that preferences are stable and unaffected by irrelevant information—here arbitrary group distinction. But further research is needed to confirm this conjecture.

Further, this paper contributes to the literature studying the relationship between individual lying behavior and beliefs about others' dishonesty. To the best of my knowledge, it is the first that studies this relationship under an incentivized and unbiased belief elicitation procedure. Introducing incentives to the belief elicitation improves the validity of our results compared to the non-incentivized elicitation procedures used in previous studies. Adding to this, our design allows drawing inferences, even though imperfectly, on individual dishonesty and, thus, to derive a more precise estimation on its correlation with corresponding first-order beliefs than comparable studies which use the individual reporting behavior as an approximation for dishonesty. Our results show a strong relationship between the behavior of individuals and their beliefs: subjects who reported the payoff-maximizing option also believe in more subjects reporting the same option. The difference is even more pronounced when controlling for the individual honesty behavior instead of the reporting behavior. This finding is in with the literature broadly suggesting a positive correlation between individual lying behavior and subjects' belief about others' lying behavior (e.g. Abeler et al., 2014; Conrads, 2014).

Even though, the positive correlation between first-order beliefs and individual lying behavior is established in the current literature and replicable, there is still a need for further research on the causality and the motives underlying a certain causality. A first attempted is made by Diekmann et al. (2011) and Rauhut (2013) who provide evidence that others' lying behavior is causal for the own behavior.

References

- Abeler, J., Becker, A., & Falk, A. (2014). "Representative evidence on lying costs." Journal of Public Economics, 113: 96-104.
- Arrow, K. J. (1973). "The Theory of Discrimination." In: Ashenfelter, O., & Rees, A. (eds.). Discrimination in Labor Markets (pp.3-33).Princeton, NJ: Princeton University Press.
- Becker, G.S. (1957 [1971]). "The Economics of Discrimination." Chicago: University of Chicago Press.
- Blanco, M., Engelmann, D., Koch, A. K., & Normann, H. T. (2010). "Belief elicitation in experiments: is there a hedging problem?." Experimental Economics, 13(4): 412-438.
- Brewer, M. B. (1979). "Ingroup bias in the minimal intergroup situation: A cognitive motivational analysis." Psychological Bulletin, 86: 307-324.
- Butler, J. V., Giuliano, P., & Guiso, L. (2012). "Trust and Cheating." Discussion Paper Series No. 6961, National Bureau of Economic Research.
- Charness, G., Rigotti, L., & Rustichini, A. (2007). "Individual Behavior and Group Member ship." The American Economic Review, 97(4): 1340-1352.
- Chen, R., & Chen, Y. (2011). "The Potential of Social Identity for Equilibrium Selection." The American Economic Review, 101(6): 2562-2589.
- Chen, Y., & Li, S. X. (2009). "Group Identity and Social Preferences." The American Economic Review, 99(1): 431-457.
- Cohn, A., Fehr, E., & Maréchal, M. A. (2014). "Business culture and dishonesty in the banking industry." Nature, 516(7529): 86-89.
- Conrads, J. (2014). "The Effect of Communication Channels on Lying." Working Paper No. 05-06, Cologne Graduate School in Management, Economics and Social Sciences.
- Diekmann, A., Przepiorka, W., & Rauhut, H. (2011). "Lifting the veil of ignorance: An experiment on the contagiousness of norm violations." Discussion Papers 4, Nuffield Centre for Experimental Social Sciences.
- Djawadi, B. M., & Fahr, R. (2015). "'... and they are really lying': Clean evidence on the pervasiveness of cheating in professional contexts from a field experiment." Journal of Economic Psychology, 48: 48-59.
- Eckel, C. C., & Grossman, P. J. (2005). "Managing Diversity by Creating Team Identity." Journal of Economic Behavior and Organization, 58(3): 371-392.

- Elster, J. (2009). "Social norms and the explanation of behavior." In: Hedström, P., & Bearman, P. (Eds.). The oxford handbook of analytical sociology (pp. 195-217). Oxford: Oxford University Press.
- Fischbacher, U., & Föllmi-Heusi, F. (2013). "Lies in disguise—an experimental study on cheating." Journal of the European Economic Association, 11(3): 525-547.
- Fischbacher, U. (2007). "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." Experimental Economics, 10(2): 171-178.
- Friesen, L., & Gangadharan, L. (2012). "Individual level evidence of dishonesty and the gender effect." Economics Letters, 117(3): 624-626.
- Greiner, B. (2015). "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE." Journal of the Economic Science Association, 1(1): 114-125.
- Houser, D., Vetter, S., & Winter, J. (2012). "Fairness and cheating." European Economic Review, 56(8): 1645-1655.
- Lane, T. (2015). "Discrimination in the laboratory: a meta-analysis of economics experiments." CeDEx Discussion Paper No. 2015-03, University of Nottingham.
- López-Pérez R. (2010). "Guilt and shame: An axiomatic analysis." Theory and Decision, 69(4): 569-586.
- López-Pérez, R. (2012). "The power of words: A model of honesty and fairness." Journal of Economic Psychology, 33(3): 642-658.
- López-Pérez, R., & Spiegelman, E. (2013). "Why do people tell the truth? Experimental evidence for pure lie aversion." Experimental Economics, 16(3): 233-247.
- Ostrom, T. M., & Sedikides, C. (1992). "Out-group homogeneity effects in natural and minimal groups." Psychological Bulletin, 112(3): 536-552.
- Phelps, E. S. (1972). "The Statistical Theory of Racism and Sexism." The American Economic Review, 62: 659-661.
- Powell, J.L. (1984). "Least Absolute Deviations Estimation for the Censored Regression Model." Journal of Econometrics, 25: 303-325.
- Rauhut, H. (2013). "Beliefs about lying and spreading of dishonesty: Undetected lies and their constructive and destructive social dynamics in dice experiments." PLoS One, 8(11).
- Tajfel, H., Flament, C., Billig, M. G., & Bundy, R. F. (1971). "Social

categorization and intergroup behaviour." European Journal of Social Psychology, 1: 149-177.

- Tajfel, H., & Turner, J. C. (1986). "The social identity theory of intergroup behavior." In: Worchel, S., & Austin, W. G. (Eds.). Psychology of intergroup relations (pp. 7-24). Chicago, IL: Nelson-Hall.
- Utikal, V., & Fischbacher, U. (2013). "Disadvantageous lies in individual decisions." Journal of Economic Behavior & Organization, 85: 108-111.

Appendix

Appendix A: Regression and other tests

			Prob > $ z $ of tests for equality		
	Yellow Group	Blue Group	(1)	(2)	(3)
Out-group treatment:					
counted speckled marbles	3	3	0.000		
reported "speckled" marbles	11	12	0.500		
average belief	9.92	10.84		0.2830	0.898
admitted dishonesty	7	6	0.600		
In-group treatment:					
counted speckled marbles	2	3	0.000		
reported "speckled" marbles	12	11	0.500		
average belief	10.38	10.08		0.4650	0.999
admitted dishonesty	7	6	0.542		

Table 3.4: Comparing Yellow and Blue Groups

Note: "admitted dishonesty" captures those subjects who report to have drawn a speckled marbles in the experiment, but make a opposite ex-post report in the survey.

(1) Fisher's exact test for equality of proportions

(2) Test for equality of medians (Wilcoxon Ranksum)

(3) Test for equality of variances (Kolmogorov-Smirnov)
* significant at p<0.10; ** significant at p<0.05; *** significant at p<0.01

Dependent Variable: Expected percentage share of "speckled" reports						
	(1)	(2)	(3)	(4)	(5)	(6)
treat	0.006	0.083	0.083	0.144		
	(0.082)	(0.164)	(0.159)	(0.109)		
speckled		0.478***			0.436***	
		(0.126)			(0.090)	
treat*speckled		-0.087				
		(0.177)				
dishonest			0.477***			0.467***
			(0.125)			(0.093)
treat*dishonest			-0.008			
			(0.179)			
male				0.265**		
				(0.112)		
treat*male				-0.279*		
				(0.808)		
Observations	52	52	35	52	52	35

Table 3.5: Tobit Regression Analysis of expected "speckled" reports

Note: Standard errors in parentheses. The dummy variable "dishonest" is 1 for those subjects who have admitted their dishonesty by deviating from their initial report, equal 0 denotes those subjects that are assumed to behave honest. We are unable to categorize those subjects that report "speckled" and do not state a deviating outcome in the survey. Thus, we remove them from the data set used in equation (3) and (6). * significant at p<0.10; ** significant at p<0.05; *** significant at p<0.01

Appendix B: Figures



Figure 3.5: Aggregate beliefs of "speckled" reporters in the Out-group and In-group treatment



Figure 3.6: Aggregate beliefs of dishonest reporters in the Out-group and In-group treatment

Appendix C: Further Material

The mathematical text problem used in the tournament

"Mr. Müller has got 5 vessels with speckled and non- speckled marbles. In vessel No. 1 are 60 marbles, half of them are non-speckled. In vessel No. 2 there are 20 marbles inside: three-quarter are non-speckled. In vessel No.3 there are only 15 speckled marbles. In vessel No. 4 are the double amount of marbles as in vessel No. 2. 35 of all marbles in vessel No. 4 are non-speckled. In vessel No. 5 are exclusively non-speckled marbles. There are twice as many marbles in vessel No. 5 as there are in vessel No. 2. Now, the hose gets filled with the contents of vessel No. 1, 2 and 4.

How many speckled marbles are inside the hose?"

Subjects have unlimited time to solve the task. Pen and blank paper are placed at each desk as auxiliary material for calculation.

Challenges in designing the text problem:

In order to make the tournament credible and effective, the text problem has to be of a medium difficulty level in such a way that many but not necessary all students are able to solve the task. First, if a task which all could solve has been chosen, the tournament would become arbitrary, as both groups would perform equally well. Consequently, it may be questionable whether the tournament enhances the group identification, as belonging to a group would seem obsolete. The same applies to a task which is too difficult. Second, seriousness of the following task requires uncertainty about the success rate in the text problem task: the participants are asked to estimate how many subjects of a certain group have answered correctly. Consequently, only a task of a medium difficulty level ensures the required variation of correct answers.

Allgemeine Erklärungen für die Teilnehmer

Sie nehmen nun an einem wirtschaftswissenschaftlichen Experiment teil. Diese Anleitung ist für alle Teilnehmer gleich, bitte lesen Sie sie genau durch. Es wird Ihnen alles erklärt, was Sie für die Teilnahme am Experiment wissen müssen. Falls Sie Fragen haben, melden Sie sich bitte per Handzeichen. Ihre Frage wird dann an Ihrem Platz beantwortet. Ansonsten gilt während des ganzen Experiments ein absolutes Kommunikationsverbot.

Für die Teilnahme am Experiment erhalten Sie 2,50 €. Im Verlauf des Experiments können Sie zusätzliche Taler dazu verdienen. Dieses Einkommen wird am Ende des Experiments in Euro umgerechnet. Hierbei gilt, dass:

10 Punkt = 30 Cent

Das umgerechnete Einkommen wird am Ende des Experiments zusammen mit dem Startgeld bar ausbezahlt.

Das Experiment

Durch das Ziehen Ihrer Sitzplatznummer sind Sie und die anderen Teilnehmer des Experiments völlig zufällig in zwei Gruppen aufgeteilt worden: die Gelben und die Blauen. Damit später keine Unklarheiten auftreten, legen Sie nun das Armband in der Farbe Ihrer Gruppe an.

Eine besondere Rolle hat die Teilnehmerin oder der Teilnehmer, die der Sitzplatz mit der Nummer 27 zugeteilt wurde. Diese Person ist keiner Gruppe zugeordnet und fungiert als unabhängige Kontrollperson. Sie wird während des Experimentes sicherstellen, dass das Experiment diesen Regeln entsprechend durchgeführt wird. Für ihren Einsatz erhält die unabhängige Kontrollperson $6 \in$ zusätzlich zum Teilnahmeentgelt.

Zu dem Experiment gehören zwei Teile. Im ersten Teil können Sie Geld für Ihre Gruppe verdienen (Turnier). Im zweiten Teil verdienen Sie Geld nur für sich (Einzelaufgaben). Die Details beider Teile werden im Folgenden erklärt.

Turnier: BLAU gegen GELB

In diesem Teil gibt es zwei Aufgaben, mit denen sich Ihre Gruppe eine Bonuszahlung erarbeiten kann. Die Bonus-Zahlung erhalten alle Angehörigen Ihrer Gruppe.

Ihre erste Gruppenaufgabe ist eine **Rechenaufgabe**. Es geht darum, die Anzahl marmorierter Murmeln in einem verborgenen Schlauch zu bestimmen.

Sie sehen vorne im Raum einen Tisch, auf dem ein Tuch liegt. Unter diesem Tuch befindet sich ein Schlauch, in dem 120 Murmeln enthalten sind. Die Anzahl der marmorierten Murmeln entspricht der Lösung einer Rechenaufgabe, die Sie später auf Ihrem Bildschirm sehen werden. Alle Teilnehmer in diesem Experiment erhalten die gleiche Rechenaufgabe. Wenn Sie die Aufgabe richtig gelöst haben, kennen Sie die Anzahl der marmorierten Murmeln. Ihre Antwort geben Sie über ein entsprechendes Feld auf dem Bildschirm ein.

Wenn in Ihrer Gruppe **mehr** Teilnehmer als in der anderen Gruppe die Anzahl der marmorierten Murmeln korrekt bestimmt haben, erhält jeder Teilnehmer aus Ihrer Gruppe einen **Bonus** von 50 Talern. Wenn in Ihrer Gruppe **weniger** Teilnehmer als in der anderen Gruppe korrekt geantwortet haben, erhält Ihre Gruppe **keinen Bonus**. Stattdessen erhält jeder Teilnehmer aus der anderen Gruppe einen **Bonus** von 50 Talern. Wenn in beiden Gruppen gleich viele Teilnehmer die Aufgabe richtig gelöst haben, erhält keine Gruppe einen Bonus.

Wenn alle Teilnehmer Ihre Antwort eingegeben haben, wird die richtige Anzahl an marmorierten Murmeln bekannt gegeben und das Tuch über dem Schlauch entfernt.

Die zweite Gruppenaufgabe ist eine **Schätzaufgabe**. Sie sollen schätzen, wie viele von 12 Teilnehmern der anderen Gruppe [*Control group: wie viele der 12 anderen Teilnehmer Ihrer Gruppe*], die Rechenaufgabe richtig gelöst haben. Ihre Antwort geben Sie über ein entsprechendes Feld auf dem Bildschirm ein. Ihre Einschätzung gilt als richtig, wenn Sie die Anzahl der richtigen Antworten exakt geschätzt haben.

Wenn in Ihrer Gruppe **mehr** Teilnehmer als in der anderen Gruppe bei der Schätzaufgabe richtig liegen, erhält jeder Teilnehmer aus Ihrer Gruppe einen **Bonus** von 50 Talern. Wenn in Ihrer Gruppe **weniger** Teilnehmer als in der anderen Gruppe richtig liegen, erhält Ihre Gruppe **keinen Bonus**. Stattdessen erhält jeder Teilnehmer aus der anderen Gruppe einen **Bonus** von 50 Talern. Wenn in beiden Gruppen gleich viele Teilnehmer bei dieser Schätzaufgabe richtig liegen, erhält keine Gruppe einen Bonus.

Einzelaufgaben

In diesem Teil gibt es erneut zwei Aufgaben. Diesmal beeinflussen Ihre Entscheidungen jedoch nur Ihre eigene Auszahlung und nicht die der Gruppe.

Die unabhängige Kontrollperson wird die 120 Murmeln aus dem Schlauch in eine sichtgeschützte Urne umfüllen. Die Urne ist so konstruiert, dass Sie eine Murmel entnehmen können, ohne dass die Experimentatoren, Teilnehmer oder die Kontrollperson sehen können, ob die entnommene Murmel marmoriert ist oder nicht.

Ihre erste Einzelaufgabe ist ein **Murmelzug**. Sie sollen verborgen eine Murmel aus der Urne ziehen. Die Kontrollperson wird dazu mit der Urne durch die Reihen gehen. Wenn Sie an der Reihe sind, greifen Sie bitte unter das Tuch und ziehen aus der Urne **genau eine Murmel** ohne selbst hinzusehen. Achten Sie dabei darauf, dass weder andere Teilnehmer noch die Kontrollperson oder die Experimentatoren sehen, ob Ihre Murmel marmoriert ist oder nicht. Lassen Sie daher die Hand zunächst geschlossen und schauen Sie selbst erst nach Ihrer Murmel, wenn die Kontrollperson bereits weitergegangen ist.

Merken Sie sich gut, ob Ihre Murmel marmoriert ist. Sie benötigen diese Information später. Die Kontrollperson wird nun erneut durch beide Gruppen mit je einer Urne gehen und die gezogenen Murmeln wieder einsammeln. Achten Sie bitte beim Zurücklegen der Murmel erneut darauf, dass niemand sieht, ob Ihre Murmel marmoriert ist.

Wenn alle Murmeln eingesammelt sind, wird auf dem Bildschirm eine Eingabemaske erscheinen. Geben Sie nun an, ob Ihre Murmel marmoriert war.

Sollte Sie angeben, dass Ihre Murmel **marmoriert** gewesen sei, dann erhalten Sie 200 Taler. Sollte Sie angeben, dass Ihre Murmel **nicht marmoriert** gewesen sei, dann gewinnen Sie nichts.

Weder die unabhängige Kontrollperson noch der Experimentator noch die Teilnehmer wissen, ob Ihre Murmel marmoriert war oder nicht. Daher hängt Ihre Auszahlung allein von den Angaben, die Sie in der Eingabemaske machen, ab und nicht davon, ob Ihre Murmel tatsächlich marmoriert war oder nicht.

Für die zweite Individualaufgabe werden Sie wieder um Ihre **Einschätzung** gebeten: Wie viele von 12 Teilnehmern der anderen Gruppe [*Control group: Wie viele der 12 anderen Teilnehmer Ihrer Gruppe*] haben angegeben, eine marmorierte Murmel gezogen zu haben? Geben Sie Ihre Antwort in die Eingabemaske auf dem Bildschirm ein.

Wenn Sie mit Ihrer Einschätzung exakt richtig liegen, erhalten Sie eine Erfolgsprämie von 200 Talern.

Im Anschluss stellen wir Ihnen noch einige demographische Fragen. Diese sind nicht auszahlungsrelevant und lassen auch keine Rückschlüsse auf Ihre Person oder Entscheidung zu. Bitte bleiben Sie danach sitzen, bis der Experimentator Sie für Ihre Auszahlung aufruft.

Wenn Sie diese Anleitung gelesen und verstanden haben, klicken Sie bitte auf "Weiter".

Kurzbeschreibung des Experiment

Turnier: BLAU gegen GELB

Rechenaufgabe:

- i Sie lösen die Rechenaufgabe zu den marmorierten Murmeln.
- ii Sie geben die berechnete Anzahl der marmorierten Murmeln an.

Wenn in Ihrer Gruppe mehr Teilnehmer als in der anderen Gruppe richtig gerechnet haben, erhält jeder Teilnehmer in Ihrer Gruppe einen Bonus von 50 Talern.

Sie erfahren, wie viele der 120 Murmeln marmoriert sind.

Schätzaufgabe:

Sie geben an, wie viele von 12 Teilnehmern der anderen Gruppe [*Control group: wie viele der 12 anderen Teilnehmer Ihrer Gruppe*], die Rechenaufgabe richtig gelöst haben.

Wenn in Ihrer Gruppe mehr Teilnehmer als in der anderen Gruppe richtig gelegen haben, erhält jeder Teilnehmer in Ihrer Gruppe einen Bonus von 50 Talern.

Teil 2: Einzelaufgaben

Murmelzug:

- i Sie ziehen eine Murmel.
- ii Sie schauen sie sich unbeobachtet an und merken sich, ob die Murmel marmoriert ist.
- iii Sie legen die Murmel unbeobachtet zurück.
- iv Sie warten bis die unabhängige Kontrollperson alle Murmeln eingesammelt hat.
- v Sie geben ein, ob Ihre Murmel marmoriert war.
- vi Sie erhalten 200 Taler, falls Sie angegeben haben, dass Ihre Murmel marmoriert war.

Schätzaufgabe:

Sie geben ein, wie viele von 12 Teilnehmern der anderen Gruppe [*Control group: wie viele der 12 anderen Teilnehmer Ihrer Gruppe*] angegeben haben, eine marmorierte Murmel gezogen zu haben.

Wenn Sie richtig liegen, erhalten Sie 200 Taler.

Organisatorisches nach dem Experiment

- i Sie beantworten demographische Fragen
- ii Sie warten an Ihrem Platz, bis Sie an der Reihe sind.
- iii Sie werden aufgerufen und erhalten Ihre Auszahlung.

Bitte beachten Sie:

- Während des gesamten Experiments ist keine Kommunikation mit anderen Teilnehmern gestattet.
- Alle Handys müssen während der kompletten Experimentdauer ausgeschaltet sein.
- Wenn Sie eine Frage haben, bleiben Sie bitte an Ihrem Platz sitzen und heben die Hand. Stellen Sie bitte Ihre Frage so, dass kein anderer Teilnehmer Ihre Frage mithören kann.
- Sämtliche Entscheidungen erfolgen anonym, d.h. keiner der anderen Teilnehmer erfährt die Identität desjenigen, der eine bestimmte Entscheidung getroffen hat.
- Auch die Auszahlung erfolgt anonym, d.h. kein Teilnehmer erfährt, wie hoch die Auszahlung eines anderen Teilnehmers ist.
- Bitte bleiben Sie bis zum Ende des Experiments an Ihrem Platz sitzen, Sie werden zur Auszahlung mittels der Ihnen zugeordneten Platznummer aufgerufen.

Viel Erfolg und Danke für Ihre Teilnahme an diesem Experiment!

Chapter 4

Cheating for My or for Your Benefit? A Field Experiment with Children

Julia Kramer, Silvia Lübbecke, and Nina Lucia Stephan University of Paderborn

Abstract

"The end justifies the means." Experimental literature has proven that subjects cheat when costs from being detected are zero and the reward sufficiently positive. People cheat if this generates an increased payoff to them, while they have little reason to cheat if thereby they can only reward someone else. While recently, some studies investigate altruistic cheating behavior of adults, little interest has been granted to children. In this paper, we run a field experiment with children to find out whether the cheating behavior is something that lies in human nature and is thus already present in young children. This would speak against theories claiming that social behavior is only developed over time, as people grow up. Our data show that children are more likely to cheat if the cheating benefits themselves instead of someone else. We find a significant difference of 16 percentage points in cheating behavior depending on who would benefit from cheating. However, many children do not seem to cheat at all, and cheating behavior also differs between genders and different age groups. Surprisingly, whether the child can benefit a friend through cheating, rather than a stranger, does only make a marginally significant difference.

Keywords: cheating, dishonesty, altruism, children, field experiment

1 Introduction

According to the saying "the end justifies the means", cheating should be a common behavior among rational agents, assuming that costs from being observed and detected are zero and the reward is sufficiently positive. In such cases, costs of cheating are non-monetary, relatively small, and mainly occurring to the actor himself in form of a bad conscience, especially if no harm is done to third parties. Previous literature confirms that among adults cheating is the predominant strategy, as it simply generates the higher payoff. But under what circumstances does this pattern of behavior hold and where does it come from?

First, this paper considers also the circumstance where a person is given a risk-free opportunity to cheat not for his own benefit but for the benefit of another person. It will be shown whether cheating is only observed when a subject can benefit himself, or whether cheating also occurs when a subject has the opportunity to instead benefit someone else. Some may call the latter altruistic behavior, since the person enhances the payoff of the other while incurring the psychological costs of cheating. Thus, we define cheating for another person's benefit as altruistic cheating. It is known that people differ regarding how altruistic they are, and thus, not everyone behaves altruistically. Moreover, altruistic cheating may simply happen less often than cheating for oneself because the person incurring the costs from cheating is not the one receiving the benefit, and thus the cheater himself can only gain indirect benefits from liking to favor someone else, which may not make up for the full costs of lying. Additionally, one may also argue that individuals refrain from cheating that benefits someone else in order not to reduce their relative standing. However, according to Erat & Gneezy (2012), some people still lie in such "Pareto white lying" situations, when lying makes somebody else better off and at least does not make themselves worse off. People may cheat in these situations for different reasons. One rather less altruistic reason may be that people may draw benefit from a winner-feeling by having tricked the system, as through cheating they have exploited their opportunities. In so far, cheating for others is not necessarily altruistically motivated. Another potential reason could be that an actor may gain benefit from behaving altruistically in form of good conscience. In this case, the net costs of cheating would be reduced and people would no longer feel that honesty is the only right thing to do. Instead, they may even feel less guilty, as they are doing someone else a favor, and cheating would happen more often. In other words, people face a moral conflict. On the one hand, they want to behave honestly. On the other hand, they may feel morally obliged to cheat in favor of the other person. Given that the benefits of behaving altruistically are sufficiently large, one could thus expect that overall there should be at least as much cheating going

on if the recipient of the payoff is another person and no longer the actor himself.

However, Fehr and Schmidt (1999) point out that the relative payoff matters and that people prefer to be relatively better off than worse off. In this sense, subjects may even cheat to reduce another person's payoff. In this paper, we define cheating in order to prevent another person from receiving the entitled payoff as negative cheating. Given that subjects display Fehr-Schmidt-preferences with sufficiently high inequity aversion, one could expect to see either no cheating regardless who is the beneficiary or at least more cheating for their own benefit than for another person's benefit.

The second, and most fundamental question that this paper aims to answer is whether both selfish cheating for one's own benefit and cheating for the benefit of someone else are behavioral patterns which are innate, rather than being learned and adapted to as people grow up. For this purpose, an experiment is run with children of different ages, in order to find out whether cheating for one's own benefit or cheating for someone else's benefit is already present in young, or very young children.

In this paper, we report data from a field experiment with 512 children. We let children play in pairs and vary the recipient of the gift that can be won in a game, either through luck with a probability of 50 percent or through cheating. With the results, we answer whether children are less likely to cheat if through cheating they benefit someone else, instead of themselves. Even more importantly, the results answer whether cheating for someone else's rather than oneself's benefit is already present at young ages, and, thus, lies in the nature of human beings.

Currently, honesty or dishonesty is a broadly discussed topic in the literature. Rosenbaum et al. (2014) review 63 economic and psychological experiments about cheating. While most experimental literature investigates the rationality of adults' lying behavior (e.g. Fischbacher and Föllmi-Heusi, 2013; Abeler et al., 2014; Conrads, 2014; Fosgaard et al., 2013; Houser et al., 2012; Jiang, 2013), little research has been done to investigate the behavior of children cheating for their own benefit (e.g. Bucciol and Piovesan, 2011; Chytilova and Korbel, 2014).

Our work is inspired by Bucciol and Piovesan (2011) who study cheating behavior of children aged between 5 and 15 years in a field experiment. They find that the majority of children does cheat, especially children above the age of eight. An explanation may be that younger children are less experienced in rational thinking than older ones and thus often act emotionally instead. The authors investigate how age development affects honesty, but do not investigate cheating when it benefits someone else. In our data set, the age differences between the children are large, including children of all ages between 3 and 16. We exploit this fact to learn more about how cheating behavior is developing over time as people grow up. Moreover, we introduce a variation in the recipient of the benefit that can be gained through cheating. This allows us to study whether children also cheat out of altruistic reasons, and whether they increasingly or decreasingly do so as they grow older.

Lately, some studies address adults cheating behavior as cheating allows them to increase or decrease others' payoff (e.g. Utikal and Fischbacher, 2013; Houser et al., 2015; Glätzle-Rützler and Lergetporer, 2015; Maggian and Villeval, 2014). Utikal and Fischbacher (2013) run a laboratory experiment in order to answer whether people prefer to appear honest rather than to actually be honest.

The present paper contributes an explanation for where the adults' cheating behavior may originate from, by running an experiment with children. Furthermore, our paper not only controls for age but also for the social relationship between the choice maker and the possibly affected person. It makes intuitive sense that interacting with a stranger has less of an impact on your own utility, as compared to interacting with a person who is known to you. Therefore, we contribute a study that takes into account the effect of social proximity to the choice maker. This is especially important when allowing for altruistic behavior in a field experiment, where participants are very aware of each other.

All together, the present paper aims at confirming the hypothesis that people are less likely to cheat if through cheating they can benefit someone else, instead of themselves. Our data will show that we cannot confirm a prevalence of altruistic cheating among very young children. Moreover, in our sample we find that older children make less of a difference between others receiving a benefit and themselves receiving a benefit. This could potentially be explained by the idea that altruistic cheating behavior develops as children grow up, and is not yet fully developed in very young children. Accordingly, altruistic behavior expressed in the act of cheating for the benefit of others may not be innate but rather learned as children grow up.

The remainder of the paper is organized as follows: Section 2 describes the experiment in more detail. Section 3 presents descriptive statistics. Section 4 continues with a summary of the results and their analysis. Section 5 concludes.

2 Experimental Design

In this section, the design and realization of our experiment, which is based on the experiment of Bucciol & Piovesan (2011), is described. Bucciol & Piovesan investigate whether children at different ages are honest, but do not investigate cheating in a situation where it benefits someone else, as it is done in our experiment.

Our experiment was conducted during three subsequent open days of a large museum in Germany in June 2015. This event was visited by school classes, kindergarten groups and families. To establish a comparable control and treatment group, we conducted two equal experimental sessions next to each other (but separated) at the same time. This setup was chosen in order to avoid having an entire class of children from one specific school or kindergarten in the control group, and a bit later an entire class of children from another school or kindergarten in the treatment group which would have resulted in more heterogeneity between the control and treatment group. In our case, children were thus waiting in a single line in front of the entrance to the experimental area and then pairs of children were randomly distributed into control- and treatment group. From each pair, one child was given the opportunity to cheat, which we call the "active child". The other child is called the "passive child". These roles were assigned randomly by rolling a dice with each of the six sides showing one out of two different colors.

Subsequently, the first experimenter explained the entire set of instructions to both children, carefully ensuring that both understood them fully. In order to avoid asymmetric information amongst the participants joining the experiment at different times of the day, the experimenter also showed the children the gift they could win in the cheating and the extra gadget that all children would receive for their participation. The gift that could be won by the children was a pencil with rainbow colors, which is tangible and whose value is easier to grasp for children, as opposed to that of an amount of money. The extra gadget was a small die.

While the active child was then carrying out the main task, the passive child received a simple entertainment task. This was important in order to ensure that both children took the experiment serious as a game, and would not be bored and find the situation strange. The active child was asked to enter a private corner, where no one could observe her actions. She was given a fair yellow-and-black coin, along with instructions to toss the coin in private, and to report the outcome (yellow or black) on a report sheet in the private corner. The experimenter made sure that she knew she would not be observed by anyone while carrying out these tasks. Hence, the active child had the possibility to cheat and report another color than the coin actually showed. Next, she handed the report sheet to the second experimenter waiting outside at the exit of the private corner. As the passive child's entertainment task was to walk around the private corner and bring a paper to the second experimenter waiting there, both children met again once the active child had finished her tasks. In most cases, the active child's task did not take longer than a minute, such that there was no long waiting time for the passive child.

Next, the second experimenter, who received the report sheet of the active child, evaluated the outcome reported on it in front of the children. In the control group, the active child was then awarded the rainbow colored pencil in case she marked yellow. If she marked black, no one received a gift. In the treatment group, the passive child was awarded the rainbow colored pencil in case the active child marked yellow. Again, in case the active child marked black, no one received a gift. The payoffs are displayed in the game tree below (Figure 4.1). After asking the children a set of control questions, both were rewarded with small dies as extra gadgets.

Participating in the experiment appeared to the children as playing a game of fortune, as it was all about throwing dice or coins and winning something. Moreover, the experiment was just one out of many stands that were especially set up for the occasion of the open day. This ensured that children were behaving rather naturally, and not make choices different from how they would usually make them. The diversity of participants (kindergarten groups, school classes and families) allowed us to obtain data from children of very different ages and relations to each other. We ensured that every child participated only once by using a congratulations stamp to mark their hands after successful participation in the experiment. Overall, 512 children participated in the experiment over the course of the three days. Descriptive statistics of the data that were gathered are presented in the third section.

3 Descriptive Statistics

Table 4.1 reports some descriptive statistics of our samples. A total of 512 children participated in our experiment. Since they were paired, our final sample consists of 256 observations stemming from the active children.

Overall, the samples from the control and treatment groups are approximately equal regarding the sample size and demographic characteristics of the participants. Noticeably, in both samples many children were paired up with a child that is known to them. This is due to the fact that the open days were mainly visited



Figure 4.1: game tree with payouts for active and passive child

	Total	Control Group	Treatment Group
Observations (512 participants)	256	121	135
Average age (min:3; max:16)	7.711	7.277	8.097
Female participants (%)	46.9	51.7	42.5
Knows passive child (%)	71.3	67.5	74.6

 Table 4.1: Summary Statistics of Active Children

by families with many having more than one child, or school classes and kinder garden groups who consisted out of larger groups of children knowing each other. As in any field experiment, the groups are of course not perfectly equal to each other regarding their mean values. Therefore, it will be interesting to not only look at how average behavior differs between the two groups, but also to calculate marginal effects taking into account some of the above characteristics as controls.

If children are truly less likely to cheat when through cheating they benefit someone else—as in the treatment group, it may matter to control for the little age differences between the treatment and control group that we observed. If most of the older children behave altruistically and most children in the treatment group are older, one may otherwise ascribe the effect of cheating to the treatment, rather than to the age of the participant. It may also matter to take the gender into account. If females are more likely than males to cheat in favor of another person, and there are much fewer girls in the treatment group compared to the control group, data may show that the treatment has the effect of reducing the likelihood of cheating. However, the actual explanation may be that there are simply fewer girls in the treatment group, and therefore we see less cheating. In the following section, we will thus also regard the results while holding gender constant, even though the percentages of male and female participants in the two groups are not significantly different from each other. Similarly, it may matter to take into account whether the children know each other. If most children would value a gift for a person known to them as much as a gift to themselves, and most children in the treatment group do know each other, we may find that there is no difference between the treatment and control group. However, whether the gift is rewarded to the active child herself or to someone unknown to them may still make a difference.

In order to avoid drawing biased conclusions by disregarding the above stories,

the following section that presents the results of the experiment will also look at the treatment effect holding many other characteristics constant. This will allow us to learn whether children are actually less likely to cheat if through cheating they benefit someone else instead of themselves, and to learn about which forms of cheating are already present in very young children.

4 Results and Discussion

We now discuss the main empirical outcomes of the study, which can be summarized in five results, where the third one is the most fundamental to learn from.

First of all, it is surprising that, in contrast to existing literature, the rate of yellow reports by all children taken together is not significantly different from 50%. At the aggregated level, children do not seem to cheat at all. Only for some subgroups based on gender and age, a tendency to cheat is observed (see Table 4.4 in the appendix). At the entire sample, 57.4% of the children reported a yellow outcome. As the coin used for the experiment is fair, one would expect 50% of the reported outcomes to be yellow. Thus, it cannot be rejected that there is no difference between the expected and reported outcome.

Result 1. On average, children do not seem to cheat. A tendency to cheat is observed in some subgroups based on age or gender, and cheating is more likely to be observed in the control group. But looking at the entire sample, there is no uniformly observed tendency to cheat.

Only looking at the averages, one may conclude that there is no cheating going on. However, the reported outcomes of the treatment group alone show that 49.2% of these children reported yellow, which is less than 50%. On the other side, 66.7% of the children in the control group reported yellow, which is significantly more than 50%. We find a highly significant difference of 17.5 percentage points in reporting yellow between the control and treatment group, based on a χ^2 -test (Figure 4.2). This suggests that at least in the control group children do cheat to obtain a gift for themselves, by reporting yellow despite the coin flip showing the opposite result. At the same time, part of this difference stems from treatment group children reporting black despite the coin flip showing yellow. By doing so, they stop the other child from receiving a gift, which they would have won through luck. This leads us to the next major result. Given the negligible differences between expected and observed outcome, one may argue that cheating is virtually not an issue in the treatment group. However, children may still cheat when cheating benefits the



Figure 4.2: percentage of children reporting yellow

other child. Assume that some children do cheat negatively, in order to prevent the other child from receiving the gift, while some other children cheat altruistically, in order to grant the gift to the other child. If this is the case, we may not be observing any cheating on the treatment level, since the two forms of cheating outweigh each other in reports of yellow outcomes, but this would not imply that no cheating is going on.

Result 2. Children seem to care significantly more for their own benefit than for that of others. They are likely to cheat for their own favor in the control group to obtain a gift which they would not have obtained through luck. They are also likely to cheat to the disadvantage of other children in the treatment group, to avoid that these other children obtain a gift.

In a next step, we run a Probit regression to confirm the treatment effect and calculate marginal effects for certain types of children. We include the following controls: age > 6, *female*, *monday*, *knows_passive_child*. The resulting overall marginal effect suggests that being in the treatment group significantly decreases the likelihood of an average child marking yellow by 16 percentage points, confirming the result from the χ^2 test. An extended table with the Probit estimates from Table 4.3 is shown in the Appendix.

	Probit marginal effects	Probit marginal effects
Treatment	-0.172*** (0.060)	-0.161*** (0.062)
Controls	NO	YES
Observations	242	241

Table 4.2: Probit regression

Standard errors in parentheses ***p<0.01, **p<0.05,*p<0.10

To find out whether the results are driven by a certain type of child, we calculate marginal effects for children with selected characteristics. We start by looking at a child with median characteristics, which could be regarded as a child best representing the sample. In the following steps, one single characteristic is deviated from at a time, in order to find out how the marginal effect changes with a specific characteristic. First, it will be looked at the effect of age, as we want to answer whether cheating only develops as children grow up. Next, the effect of gender as well as the personal relation between the active and the passive child is looked at.

The median child is a seven-year-old boy who participated on a Sunday or Tuesday and knew the passive child. The marginal effect for a child with these characteristics is 16.7 percentage points. We look at how marginal treatment effects differ between children at school age compared to younger ones. Results show that for the median child up to the age of six the marginal effect of being in the treatment group is larger (16.9 percentage points) than for the median child with an age above six (15.2 percentage points). In other words, older children seem make less of a difference between others receiving a benefit and themselves receiving a benefit. Even though, the difference in the average marginal effects falls short of statistical significance according to the Probit regression (p=0.2364), from an economic perspective the difference may still be significant. Further research is needed to confirm this observation. A possible explanation for a difference in behavior may be that the older the children, the more likely they might be to cheat altruistically when the reward from cheating benefits another child instead of themselves. This is line with the idea that as children grow up, they learn that altruistic behavior does pay off. They may not yet think about any potential long term consequences

of their social behavior. At the same time, older children still cheat in favor of their own benefit but to a lesser extent than young children (see Table 4.4, appendix). We may conclude that Result 2 is rather driven by those children who are not old enough to be at school.

Result 3. While very young children are significantly more likely to promote their own benefit than that of others, children who have at least school age seem to make less of a difference between others receiving a benefit and themselves receiving a benefit. Older children still cheat in favor of their own benefit but to a lesser extent than young children. Adding to this, they may behave more altruistically than younger children. This would match the idea that altruistic behavior rather develops as people grow up, and is not fully developed in very young children.

As the median child is a boy, but a large share of the participants (46%) were female, let us now look at the median child if it was female. The marginal effect for the 7-year-old median child if it was a girl is only 15.4 percentage points. Ergo, the marginal effect for a female median child is smaller (though statistically not significant according to the Probit regression, p=0.2117) than for a male median child. This suggests that girls may not make as much difference as boys do, when the benefit is no longer awarded to themselves but to the passive child instead. If this is confirmed robust by future studies, it would show that girls may behave more altruistically than boys. In so far, this finding would be in line with other findings in the field of literature, where commonly women are observed to be more willing to do good to another person. Notwithstanding, boys may cheat as much as girls. If altruistic cheating among boys is as common as among girls and negative cheating is more frequent among boys, we would also observe boys to cheat less for others' benefit than girls do. Result 2 may thus be driven more by the male participants than by the female ones. To disentangle these explanations, future research is needed.

Result 4. While overall children are significantly more likely to promote their own benefit than that of others, this pattern of behavior may be expressed more strongly among boys than among girls. This may be explained by the idea that altruistic behavior might be rather present in females compared to their male counterparts.

The last characteristic to look at is the personal relation between the active and the passive child. Thus, we check whether marginal effects change if the median child does no longer know the passive child. The difference of the marginal effects in this case is negligible (17.1 percentage points if the passive child is unknown versus 16.7 percentage points if it is known, p=0.5277). Still, the direction of the

difference points in the expected direction, which is that children are even more likely to make a difference between themselves and another child receiving the gift if this other child is a stranger to them. However, according to the nonparametric χ^2 -test, children in the treatment group are weakly significantly more likely to report yellow when the gift is awarded to a friend instead of a stranger: when children can award the prize to a known child (friend or sibling), 53.61% report yellow, while only 35.48% do so when they can award the prize to an unknown child (see Table 4.4, appendix; p=0.079).

Result 5. Children are significantly more likely to promote their own benefit rather than the benefit of others, even if this other person is known to them or a friend of theirs. Though, friends are statistically treated differently from strangers: children seem to report yellow more frequently if the other child is a friend instead of a stranger.

All in all, one can summarize that the results of calculating marginal effects of the treatment for children with selected characteristics confirm the expectations from the descriptive statistics. While on average it is not evident that children cheat, they are significantly more likely to cheat for their own rather than for someone else's benefit. Moreover, our sample shows that children do start to make less of a difference between others receiving a benefit and themselves receiving a benefit, as they grow older. Looking at some additional characteristics reveals, that the treatment effect is rather driven by boys. Surprisingly, whether the children know each other or not seems to matter only marginally for their choice.

5 Conclusion

This paper investigates whether children adapt their cheating behavior if they cannot benefit themselves through cheating but instead benefit someone else. The results suggest that even young children are already more likely to cheat with the intent to promote benefits if they themselves are the beneficiary of cheating instead of others. This speaks for the idea that cheating for one's own rather than someone else's benefit lies in the nature of human beings, and does not only develop as people grow to adults. Some children even already seem to have some basic sense of inequity aversion with respect to their own relative payoff, since they care to not have less than others, as some are believed to cheat in order to avoid that the other child wins a gift through luck. This general pattern of behavior seems to already exist in young children. For young children, our result may be pronounced stronger than for older children. This would suggest that this behavioral pattern might be overthrown by another pattern, as children grow up. However, the difference between young and old children is statistically not significant. Likewise, there is a slight—but not statistically significant—indication that boys may make a larger difference between themselves and others than girls do. This would also speak for the idea of a changing behavioral pattern. As girls are brought up to become women, and thus likely to be confronted with female role models, they may develop a behavior that is increasingly different from that of boys, who are taught to follow the men's example instead. In this case, younger children and boys drive the results.

Further, we find that children make a significant difference between themselves and others, even when the other person is a friend instead of a stranger. However, they seem to be more gift-giving towards their peers than they are towards strangers. This find could be potentially explained by the idea that children behave altruistically towards their friends, or respectively are more envious towards strangers.

Even though, the age range of participants in this experiment is relatively large, it was only designed for rather young children at kinder garden or elementary school age. In order to confirm whether the result continues to grow stronger with age, and whether the difference of behavior between the genders continues to grow, or whether boys catch up over time, additional research will have to be conducted. Moreover, carrying out this project made us wonder under which circumstances subjects would put an effort and cheat for avoiding that others receive benefits through luck. This would be especially interesting in a setup like this, where there is nothing to lose by cheating but no direct gain from cheating for (or against) the benefit of someone else either.

References

- Abeler J., Becker A., & Falk A. (2014). "Representative evidence on lying costs." Journal of Public Economics, 113: 96-104.
- Bucciol, A., & Piovesan, M. (2011). "Luck or cheating? A field experiment on honesty with children." Journal of Economic Psychology, 32 (1): 73-78.
- Chytilova, J., & Korbel, V. (2014). "Individual and group cheating behavior: A field experiment with adolescents." IES Working Paper, 06: http://hdl.handle.net/10419/102591.
- Conrads, J. (2014). "The Effect of Communication Channels on Lying." Cologne Graduate School Working Paper, 5(6): 1-31.

- Erat, S., & Gneezy, U. (2012). "White lies." Management Science, 58(4): 723-733.
- Fehr, E., & Schmidt, K.M. (1999). "A Theory of Fairness, Competition and Cooperation." The Quarterly Journal of Economics, 114 (3): 817-868.
- Fischbacher, U., & Föllmi-Heusi, F. (2013). "Lies in disguise—an experimental study on cheating." Journal of the European Economic Association, 11(3): 525-547.
- Fosgaard T. R., Hansen L. G., & Piovesan M. (2013). "Separating Will from Grace: An experiment on conformity and awareness in cheating." Journal of Economic Behavior & Organization, 93: 279-284.
- Glätzle-Rützler, D., & Lergetporer, P. (2015). "Lying and age: An experimental study." Journal of Economic Psychology, 46: 12-25.
- Houser, D., List, J. A., Piovesan, M., Samek, A. S., & Winter, J. (2015). "On the Origins of Dishonesty: From Parents to Children." National Bureau of Economic Research, w20897.
- Houser D., Vetter S., & Winter J. (2012). "Fairness and cheating." European Economic Review, 56 (8): 1645-1655.
- Jiang T., (2013). "Cheating in mind games: The subtlety of rules matters." Journal of Economic Behavior & Organization, 93: 328-336.
- Maggian, V., & Villeval, M. C. (2014). "Social preferences and lying aversion in children." Social Science Research Network, 2368098, http://dx.doi.org/10.2139/ssrn.2368098
- Rosenbaum, S. M., Billinger, S., & Stieglitz, N. (2014). "Let's be honest: A review of experimental evidence of honesty and truth-telling." Journal of Economic Psychology, 45: 181-196.
- Utikal, V., & Fischbacher, U. (2013). "Disadvantageous lies in individual decisions." Journal of Economic Behavior & Organization, 85: 108-111.

Appendix

	Probit	Probit marginal effects	Probit	Probit marginal effects
treatment	-0.450***	-0.172***	-0.435**	-0.161***
	(0.164)	(0.060)	(0.175)	(0.062)
Age > 6			-0.495***	-0.183***
			(0.190)	(0.067)
Female			0.275	0 101
			(0.173)	(0.063)
Monday			0 420*	0 158**
·			(0.221)	(0.080)
Knows passive child			0 104	0.028
_1 _			(0.192)	(0.071)
Constant	0 /21***		0 6/1***	
	(0.121)		(0.227)	
observations	242	242	241	241

Table 4.3: Probit Regression Full

Standard errors in parentheses ***p<0.01, **p<0.05,*p<0.10

	Control Group ¹⁾	Treatment Group ¹⁾	Pearson's chi-square test of independence
Pairs where			
All pairs	66.67%***	49.22%	0.006***
	(76/114)	(63/128)	
other child unknown	71.79%***	35.48%	0.002***
	(23/32)	(11/31)	
other child known	63.51%**	53.61%	0.194
	(47/74)	(52/97)	
active child aged>7	60.42%	47.76%	0.180
	(29/48)	(32/67)	
active child aged<=7	71.21%***	50.82%	0.0180**
	(47/66)	(31/61)	
active child male	60%	47.37%	0.153
	(33/55)	(36/76)	
active child female	72.41%***	51.92%	0.026**
	(42/58)	(27/52)	

Note:

* significant at p<0.10; ** significant at p<0.05; *** significant at p<0.01
¹⁾ Goodness-of fit tests are applied to test whether the gift reception rates significantly differ from the 50%. Significance levels are reported behind the respective gift reception rates.

Table 4.4: Testing for equality in the rates of yellow reports

Appendix

Erklärung zu Studien in Koautorenschaft

Ich, Silvia Lübbecke, erkläre hiermit, dass keine der Studien, die Teil dieser Dissertation sind, bereits in anderen laufenden oder abgeschlossenen Promotionsverfahren benutzt worden sind.

Paderborn, den 20. März 2018

Silvia Lübbecke

Eidesstattliche Erklärung

Ich, Silvia Lübbecke, erkläre hiermit eidesstattlich, dass ich vorliegende Dissertation selbständig und ohne unerlaubte Hilfe verfasst habe. Ich habe ausschließlich die in der Arbeit angegebenen Literatur und Hilfsmittel verwendet. Alle Stellen, die wörtlich oder inhaltlich den Schriften anderer Autorinnen und Autoren entnommen wurden, habe ich als solche kenntlich gemacht. Zudem erkläre ich, dass ich keine Hilfe von Vermittlungs- und Beratungsdiensten in Anspruch genommen habe.

Ich erkläre ebenfalls, dass ich keine vorherigen Promotionsverfahren beantragt habe. Insbesondere versichere ich, dass die vorliegende Dissertation in gleicher oder anderer Form in keinem anderen Prüfungsverfahren vorgelegt oder veröffentlicht wurde.

Paderborn, den 20. März 2018

Silvia Lübbecke