COOPERATION, DILIGENCE, AND ETHICAL BEHAVIOR:

FOUR ESSAYS IN EXPERIMENTAL ECONOMICS

Inauguraldissertation zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Universität zu Köln

2019

vorgelegt von
Anna Untertrifaller
aus Bozen, Italien

Referent:	Prof. Dr. Matthias Sutter
Korreferent:	Prof. Dr. Bettina Rockenbach
Tag der Promotion:	15. Januar 2019

To my family.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Matthias Sutter for his outstanding and continuous support. I am deeply convinced that I could not have found a better supervisor. Working on joint projects was the best way for me to learn how to conduct research. His encouragement, advice, and understanding made me succeed in handing in this dissertation.

I would also like to thank the other members of my thesis committee. I am very grateful that Bettina Rockenbach accepted my request to become my second supervisor. I deeply appreciate her support especially at the beginning of my dissertation. Her lectures and talks served as a great inspiration to me. My sincere thanks also go to Oliver Gürtler for being the head of my thesis committee.

My gratitude goes to my co-authors and friends Caroline Stein and Claudia Zoller for the wonderful time we had working together. I take away a lot from our collaborations and found admirable friends. Thanks also to Thomas Lauer for his contribution to our project.

I am also thankful for my dear colleagues at the University of Cologne and the Max Planck Institute for Research on Collective Goods. Thanks to Zvonimir Bašić, Stefania Bortolotti, Justin Buffatt, Zita Coppinger, Rebekka Cordes, Lisa Dickmanns, Zwetelina Iliewa, Felix Kölle, Helene Mass, Sofia Monteiro, Matthias Praxmarer, Shambhavi Priyam, Angelo Romano, Daniel Salicath, Sebastian Schneider, Kristina Steinbrecher, and Jarid Zimmermann for the friendly working atmosphere and their valuable feedback.

I gratefully acknowledge financial support from the Cologne Graduate School (CGS), the Center for Social and Economic Behavior (C-SEB), and the Research Unit Design and Behavior (FOR 1371).

Last, my special appreciation goes to my family whose love and support I owe so much to. I dedicate this thesis to you.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

The field of behavioral economics enriches our understanding of economic phenomena by focusing on humans' preferences and behavior (Camerer and Loewenstein, 2011). This thesis aims to contribute to this field of research by studying three aspects of humans' behavior, namely cooperation, diligence, and ethical behavior. All three aspects are key factors for societal and economic development. Without people cooperating with each other, men would have hardly succeeded in building up a solid social and economic system. Future challenges, such as tracking climate change or keeping away from the usage of nuclear weapons, will increasingly depend on this human ability to cooperate. Being diligent, meaning being able to work hard, is another crucial aspect for economic and social progress. At an individual level it is an important requisite in the labor market and a positive predictor of individuals' success, such as educational outcomes (Duckworth et al., 2007; Eskreis-Winkler et al., 2014; Alaoui and Fons-Rosen, 2016). Ultimately, people's concern for behaving ethically, i.e., adhering to stated rules, is a necessary condition for facilitating humans' co-existence and the pursuit of society's overall well-being.

Each chapter of this thesis aims to closely analyze one of these kinds of behavioral patterns by shedding light on either its origins, development, requirement, or repercussions. For this purpose, all chapters are based on evidence gathered in experiments. The use of experiments - both in the laboratory and the field - is a well-established tool to generate insights regarding individuals' behavior and economic preferences. It allows to study individual decision making in isolated situations and under controlled conditions (Friedman and Sunder, 1994).

Chapter one and two of this thesis present the results of lab-in-the-field experiments conducted with children and in kindergarten. Studying

children's behavior has received growing interest in economics over the last years. First, studies with children contribute to a better understanding regarding the source of heterogeneity in humans' preferences and outcomes we observe later on in life (Dohmen et al., 2011; Bauer et al., 2014; Deckers et al., 2015; Almås et al., 2016). This is also the aim of chapter one as it proposes family background as one explanatory variable for differing levels of cooperation. Second, studies with children shed light on the development of people's preferences and skills as they generate insights on whether economic preferences and behavioral patterns remain stable from early on in life or are shaped with time (Fehr et al., 2008; Sutter et al., 2013; Lergetporer et al., 2014; Sutter et al., 2018). Here chapter two makes a contribution by studying diligence and its development during early childhood.

In contrast, chapter three and four of this thesis present the results of traditional laboratory experiments conducted with a standard subject pool of students. Both chapters investigate people's behavior in situations where people face a trade-off between following the principle of adhering to stated rules or circumventing them for the maximization of one's own payoff. Experimental evidence shows that - contrary to standard economic predictions - people do not only maximize their own earnings but are sensitive to the way this maximization is achieved. In this regard, people, for instance, decide not to lie even if lying in monetary terms would be beneficial to them (Gneezy, 2005; Sutter, 2009; Gneezy et al., 2013; Fischbacher and Föllmi-Heusi, 2013; Abeler et al., 2018). On the other hand, fraudulent employees' behavior - including cases in which employees inflate their expenses, working hours, or efforts - costs industries and countries all over the world billions of dollars (ACFE, 2018) and shows that unethical behavior is widespread in business interactions. This thesis enriches the study of (un)ethical behavior by shedding light on (i) how other group members' dishonesty affects individual lying behavior, and (ii) how assuming responsibility for an ethical or unethical work environment affects subsequent performance. In this regard, chapter three of this thesis focuses on conditional dishonesty, namely on people's tendency to be dishonest if others are so too. Finally, chapter four shows that assuming responsibility for one's own work environment

to be ethical might serve as a non-pecuniary incentive to increase workers' motivation and performance.

In the following I will provide an overview of the main findings of each chapter of this thesis.[1]

OVERVIEW OF MAIN FINDINGS

Chapter 1. *Parental background and children's heterogeneity in cooperation* is joint work with Matthias Sutter. In this chapter we study the development of humans' willingness to cooperate. Social interaction relies crucially on cooperation as it generates synergies and increases the well-being of society. We focus on the heterogeneity in children's cooperation and study how children's cooperation behavior is related to parental socio-economic background, parents' own cooperation behavior, and their subjective perception of their child. Contrary to previous research, we do not only consider the case where mutual cooperation creates symmetric but also asymmetric outcomes. For this purpose, we let four- and five-year-old children and their parents play an experimental prisoner's dilemma game with either symmetric or asymmetric payoffs in case of mutual cooperation. While we find that asymmetries do not decrease cooperation rates, we find that parents with higher education have children who are more likely to cooperate. Parents' and children's likelihood of cooperation is positively, albeit insignificantly, aligned. We also find that parents' subjective perception of their child to be superior to peers is related to higher cooperation rates in children. In a follow-up study, we find that a significant fraction of parents perceives children as superior if these children are particularly popular among their mates, which might explain the positive correlation we observe.

Chapter 2. *Busy little bees - An experiment on diligence and endogenous time scheduling in early childhood* is joint work with Matthias Sutter and Claudia Zoller. In an experimental setting we analyze the development and determinants of diligence in early childhood. Diligence, a non-cognitive skill capturing the individual ability of being hard working, has been shown to

[1] In each chapter of this thesis all authors contributed equally to the work.

be decisive for educational attainment and, thus, later success in life. Since non-cognitive skills have also been proven to be malleable during childhood, it is important to study these skills at an early point in life. We do this by letting children aged three to six work on a real effort task for as long as they want and measure the output. We classify this as the level of diligence they display. We find that giving children the possibility to decide when to work on the task - either today or tomorrow - shows that those who procrastinate the task to the following day provide significantly less effort. This result holds when we control for age, where we find that younger children do not only display lower levels of diligence but are also more likely to procrastinate. In addition, we find that children's diligence relates to their willingness to challenge themselves in an unrelated task. When we introduce family background as an explanatory variable for differing levels of diligence, we see that parents with higher education have children who display higher levels of diligence. We do not find that children's (experimentally elicited) and parents' (self-reported) diligence are intercorrelated, while we find a positive correlation between children's (again experimentally elicitated) and parents' (self-reported) tendency of procrastination.

Chapter 3. *Conditional lying* is joint work with Thomas Lauer. Dishonesty is a widespread concern within and across companies. A characteristic feature of companies is that employees who collaborate with each other can observe each others' behavior. Observing that others are dishonest might impact one's own decision whether to be dishonest too. This might be especially the case in group settings where employees bear the consequences of their group members' dishonesty. We study this question in a laboratory experiment and introduce three different group settings where group members' lies have either no or increasingly harmful effects on one another. We find that irrespective of how others' lies affect one's own payoff, one third of all subjects are dishonest if others are too. This result shows that a considerable amount of people do condition their dishonesty on their group members' dishonesty. In addition, the result demonstrates that group members' tendency to condition their own dishonesty on others' dishonesty occurs to a similar extent whether or not group members' dishonesty has

mutual monetary effects on one another. We find that for the vast majority of *conditional liars* having only one dishonest group member is enough to make them switch from being honest to being dishonest. Overall, *conditional liars* tell smaller lies (i.e., inflate their performance to a smaller degree) than *always liars* do. Both types of liars increase the size of their lies the more dishonest group members they face. Taken together, these results suggest that social influence - the notion of other people's behavior and attitudes shaping one's own conduct - contributes strongly to the emergence and magnitude of dishonesty in group settings.

Chapter 4. *The effect of ethical responsibility on performance* is joint work with Caroline Stein. It is common practice that companies offer their employees non-pecuniary incentives in order to increase their motivation and performance. In a laboratory experiment, we study whether being responsible for one's own working environment to be ethical or unethical may serve as such a non-pecuniary incentive to boost workers' performance. Our data shows that workers who prefer to work in an ethical work environment perform better if they are responsible for this ethical work environment compared to a situation where the employer chooses the ethical work environment for them. However, when workers choose to perform in an unethical work environment, we do not find the same positive effect of being responsible. We argue that in the latter case the positive effect of being responsible is counterbalanced by higher ethical costs one bears when being responsible for a violation of rules. We also see that workers care about their work environment even in situations where they are not responsible for this environment. Workers who are forced to work in an environment that violates their own ethical standards perform worse than workers whose own ethical standards are not violated by an imposed environment. Companies might want to take these insights into account when deciding whether and how to shape their employees' work environment.

CHAPTER 1

PARENTAL BACKGROUND AND CHILDREN'S HETEROGENEITY IN COOPERATION

Abstract

Social interaction relies crucially on mutual cooperation to achieve synergies and increase the welfare of human society. We study the development of humans' willingness to cooperate by letting 4- and 5-year-old children and their parents play an experimental prisoner's dilemma game. We focus on how children's cooperation is related to parental socio-economic background, parents' own cooperation behavior, and their subjective perception of their child. We find that parents with higher education have children who are more likely to cooperate. Parents' and children's likelihood of cooperation is positively, albeit insignificantly, aligned. Being perceived to be superior to peers is related to higher cooperation rates. Contrary to previous research, we do not only investigate the case when the benefits from mutual cooperation accrue to the same extent to both parties but also when one party benefits more from mutual cooperation. We find that asymmetric outcomes in case of mutual cooperation do not decrease cooperation rates in children. Overall, these findings take us one step further towards understanding the development and heterogeneity in humans' willingness to cooperate.

This chapter is joint work with Matthias Sutter.

1.1 INTRODUCTION

The ability to cooperate has generated huge efficiency gains for human society and is a key for societal and economic development. Without people cooperating with each other, men would have hardly succeeded in building up a solid social and economic system. Today's economy and multilateral institutions function properly because of people working together. Future challenges, such as tracking climate change or keeping away from the usage of nuclear weapons, will increasingly depend on this human ability to cooperate. In experimental settings, several factors like direct and indirect reciprocity or punishment of norm violations have been shown to contribute to high levels of human cooperation (Fehr and Gächter, 2002; Fehr and Fischbacher, 2004; Nowak, 2006; Herrmann et al., 2008). Yet, even in one-shot interactions and in the absence of punishment opportunities, humans show, in the aggregate, a remarkable ability to cooperate. In prisoner's dilemma games, for instance, children and adults deviate from their dominant strategy of always defecting and cooperate to a considerable extent (Sally, 1995; Lergetporer et al., 2014; Charness et al., 2016).

Hidden behind a substantial degree of cooperation, however, emerges a significant individual heterogeneity in the propensity to cooperate. While some people only cooperate if others do so too, others are by principle either cooperative or free-riders (Fischbacher et al., 2001). This heterogeneity in individual willingness to cooperate potentially develops early on in life and presumably persists over the whole lifespan. It is, therefore, important to understand where the heterogeneity comes from. In this paper, we aim at improving our understanding of individual heterogeneity in cooperation by examining the correlation between children's cooperation and several factors related to their parents, more precisely parental socio-economic status, their own cooperation behavior, and their perception of their child.

We present the results of an experimental prisoner's dilemma game that was run both with 4- to 5-year-old children and with their parents. The choice of 4- to 5-year-old children is motivated by the observation that economic preferences develop early on in life (Fehr et al., 2008; Almås

et al., 2010; Sutter et al., 2015; Alan et al., 2017; Alan and Ertac, 2018), for which reason we want to identify the factors that shape behavior in a strategic game at such young age. Already at the age of 4 and 5 children get involved in activities that are related to cooperation problems (e.g., when playing and afterwards restoring order together, when dividing little duties in kindergarten, etc.). However, this age span has not been considered in the literature on cooperation yet.

First, we analyze the correlation between parental education (as a proxy for socio-economic status) and a child's behavior in a prisoner's dilemma game. Recent work on social preferences has shown that higher socio-economic status of parents is related to more prosocial behavior of their children, so we expect also a positive relationship here. Bauer et al. (2014), for instance, find that family background plays a role in shaping children's preferences in individual decision making tasks. In a study with 4- to 12-year-old children, they show that children of lower educated parents are less altruistic, more selfish, and more likely to be weakly spiteful. Deckers et al. (2015) reinforce these findings by demonstrating that children from families where mothers and fathers have a higher average year of education and where the family disposes of a higher household income are more patient, more altruistic, less likely to be risk seeking, and score higher on IQ tests.

Note, however, that a novel feature of our study is that we consider a strategic setting in game theoretic terms, while the previous studies have exclusively studied non-strategic decisions, such as simple allocation tasks or risk and time preference experiments. Considering behavior in a strategic setting is important because most real world decisions are rarely made in isolation but predominately in interaction with others. Whenever the decision of someone else matters for one's own payoff, beliefs regarding the opponent's behavior play a fundamental role in predicting the own behavior. Previous literature has found that children and adolescents at the age of 7 to 17 make their decisions in strategic settings depend on their expectations regarding their counterpart's behavior (Lergetporer et al., 2014; Czermak et al., 2016). Therefore, considering a strategic game might change or even disband the relationship between parental socio-economic status

and children's behavior because the beliefs about the interaction partner's behavior might influence a child's behavior to an extent that dominates the potential impact of parents' socio-economic status. In order to control for this possibility, we elicit beliefs of children about their interaction partner's behavior.

Second, we let both children and their parents play an experimental prisoner's dilemma game in order to measure whether parental behavior and children's behavior are well aligned. So far, the existing evidence about parental choices and preferences to be correlated with the choices and preferences of their children is mixed. Dohmen et al. (2011), for instance, show that trust and risk preferences of parents and grown-up children are positively associated. However, there seems to be no clear correlation between the behavior of children and their parents in strategic and not strategic games which reveal their social preferences. Cipriani et al. (2013), for instance, find no correlation between prosociality of elementary school children and their parents in a standard public good game. Also, Ben-Ner et al. (2017) find no significant correlation between the fraction transferred to a passive receiver in a dictator game by 3- to 5-year-old children and their parents. It is, therefore, a priori unclear how the correlation between parents' and children's willingness to cooperate should be. Our hypothesis is that cooperation of parents in the prisoner's dilemma game might be positively related to cooperation of their children because we expect social norms to be transmitted from parents to children.

Third, compared to the relatively "hard" factors like socio-economic status of parents and parents' behavior in a prisoner's dilemma game, we also consider a comparatively "soft" factor, which is how parents see and judge their child and how they perceive their child in relation to the child's peers. Recent work in psychology (Brummelman et al., 2015a) has developed two scales on which we draw upon here: one with regard to parents holding an overly optimistic view of their child (parental overvaluation scale) and one with respect to whether parents perceive their child to be superior to their peers (perceived superiority scale). It has been shown that an overvaluation of the own child is particularly strong in narcistic parents and often causes

children to develop a narcistic self-view themselves (Brummelman et al., 2015b). Yet, it has remained unexplored whether the degree of overvaluation and perceived superiority also affects specific aspects of a child's behavior, such as their willingness to cooperate. For this reason, we test whether the parental overvaluation and perceived superiority of a child is related to the level of a child's cooperation. Parents that highly overvalue their child may instill in this child the belief that he or she deserves special and benevolent treatment by others but is less obliged to reciprocate such behavior. Hence, children who are more overvalued and perceived to be superior to others are expected to cooperate less.

Lastly, we do not only concentrate on situations where mutual cooperation benefits all involved parties equally but extend the standard prisoner's dilemma game to asymmetric conditions in case of mutual cooperation. As cooperative behavior in everyday life situation hardly never benefits parties equally, we want to gain a deeper understanding about how cooperation could evolve in mankind despite the fact that it is often more profitable for some than others. It has been shown that for coordination games even small asymmetries decrease people's ability to coordinate (Crawford et al., 2008). Our hypothesis, therefore, is that also cooperation declines when asymmetries in case of mutual cooperation are introduced.

Our experiment yields the following results: 1) Families where at least one parent holds a high-school or university degree have children who are more likely to cooperate compared to children whose parents obtained less than 13 years of schooling. 2) Parents' and children's willingness to cooperate is positively correlated, however this correlation lacks of statistical significance at conventional levels. 3) Parents who perceive their children to be superior to others have children who are more likely to cooperate. 4) Asymmetries in case of mutual cooperation do not affect cooperation rates in children.

Given our results 1) and 3), we conducted a short follow-up study in March 2017 in order to improve our understanding of why parental education and perception might correlate with a child's cooperation. We find that: A) Parents with a higher education want their children more often to cooperate in a prisoner's dilemma game than parents with a lower education,

suggesting that normative expectations might be one important channel through which parental education is related to children's cooperation. B) A significant fraction of parents perceives children as superior if these children are particularly popular among their mates. In this sense, a child's popularity might explain the positive correlation we observe between perceived superiority and higher cooperation rates.

The rest of the paper is organized as follows. In section 1.2 we present the procedures and experimental design of our study. In section 1.3 we analyze our data, illustrate our results, and in addition present and highlight results from the follow-up study. Section 1.4 concludes.

1.2 EXPERIMENT

1.2.1 PROCEDURES

We ran our study with 328 children from nine kindergartens all over the province of South Tyrol, Italy. The project was approved by the internal review board of the University of Innsbruck, Austria and the state board of kindergarten education of South Tyrol. Parents were informed about our visit beforehand and were given the opportunity to opt out their child from participation. Participation of children was, of course, voluntary.

In Italy, kindergarten includes three cohorts of children, aged 3-4, 4-5, and 5-6 years. We ran the experiment only with the oldest two cohorts because when running the experiment (in November 2015) the youngest cohort had only entered kindergarten two months prior to our experiment and teachers still worked on integrating the youngest children into kindergarten. The experiment was conducted during regular kindergarten opening hours and carried out in a separate room where either four or six children were present at the same time such that pairs of children could make simultaneous decisions. Female student helpers, all of whom were future kindergarten or elementary school teachers in their third, ultimate, or penultimate year before concluding their education, explained the rules of the experiment individually and face to face to one child each.

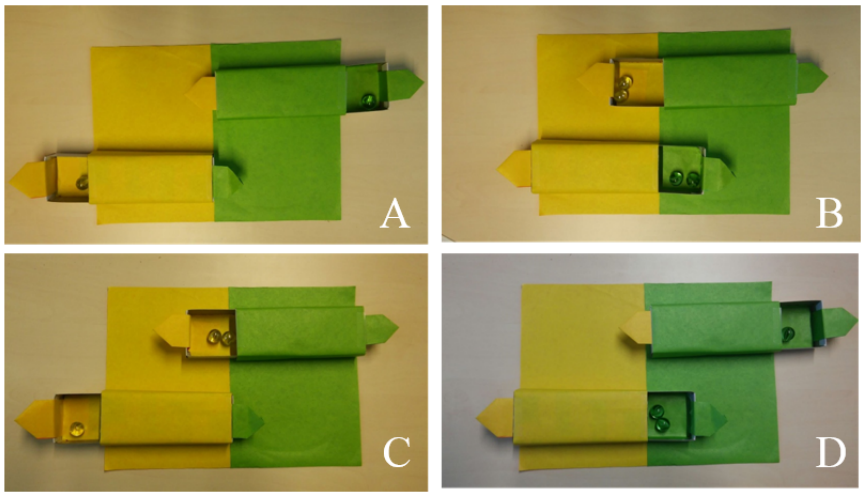
All decisions children made were incentivized by giving children colored marbles which they could at the end of the experiment exchange into a variety of different goods (pencils, balloons, hair clips, finger skateboards, etc.). Before taking part in the experiment, children had to answer control questions in order to check for their understanding. If a child was not able to answer correctly, the student helpers repeated the instructions. This process continued up to three times. If, after the third repetition, a child did not answer correctly the student helper took note of it and proceeded in conducting the experiment.

1.2.2 EXPERIMENTAL DESIGN

In order to make the prisoner's dilemma game and its payoff structure understandable for children, we used a physical game with different colors and a pulling-boxes task to reflect the different decision situations and outcomes (as illustrated in Figure 1.1). Basically, we showed each child a two-colored game. Then we assigned children either the color yellow or green and told them that they were paired with another child in the room of the opposite color. It is worth noting that children did not know with whom of the present children they were paired with, neither before taking part in the experiment nor after having done so.

Each child made the decision whether to cooperate or defect on a separate two-colored game. For this purpose, children could pull one out of two small boxes that were placed in a large box with the child's own color. Pulling out the small box with the own color yielded one marble for this child but nothing for the other child. This stood for the non-cooperative action. Pulling out the small box with the color of the other child yielded two marbles for the other child and nothing for oneself. This represented the cooperative action. Of course, mutual cooperation by pulling the box with the other child's color is the socially optimal outcome here (illustrated in Panel B of Figure 1.1). Yet, defecting by pulling the box with the own color (shown in Panel A of Figure 1.1) is a dominant strategy and leads to

Figure 1.1: Experimental design



Notes: Panel A: Both children picked the small box of their own color and thus defected. Mutual defection made them gain one marble each. Panel B: Both children picked the small box of their partner's color and thus cooperated. Mutual cooperation made them gain two marbles each. Panel C and Panel D: one child cooperated (green in C, yellow in D) and the other defected, yielding three marbles for the defecting child and zero for the cooperating. The illustration refers to payoffs in the symmetric treatment condition.

the only Nash-equilibrium of the game because it maximizes own earnings irrespective of the other child's action.

After both children had independently (and simultaneously) made their decision whether to cooperate or defect, we elicited the belief of each child about the other child's decision. If the belief was correct, the child earned an additional marble, making the belief elicitation incentivized. After the elicitation of the beliefs, student helpers matched the boxes of the paired children and revealed the matched partner's decision.

1.2.3 TREATMENT VARIATION

We ran three different treatments of the described prisoner's dilemma game, where we varied the payoff structure in case of mutual cooperation. 106 children took part in treatment SYM, 112 in treatment ASY, and 110 in treatment MERIT. The exact payoff structures in the different treatments are highlighted in Figure 1.2. While in treatment SYM (see the left hand side of

Figure 1.2) both children gained equally from mutual cooperation (2,2), being cooperative was more profitable for one child than the other in the treatments ASY and MERIT (2,3; see the right hand side of Figure 1.2). In treatment ASY advantageous and disadvantageous asymmetries in case of mutual cooperation were randomly allocated to children. Whereas in treatment MERIT asymmetries were justified based on children's performance in a previous task.

Figure 1.2: Treatment variation

	Child 2 does not cooperate	Child 2 cooperates
Child 1 does not cooperate	1,1	3,0
Child 1 cooperates	0,3	2,2

	Child 2 does not cooperate	Child 2 cooperates
Child 1 does not cooperate	1,1	3,0
Child 1 cooperates	0,4	2,3

Notes: Payoff matrix in treatment SYM (left) and in treatments ASY and MERIT (right).

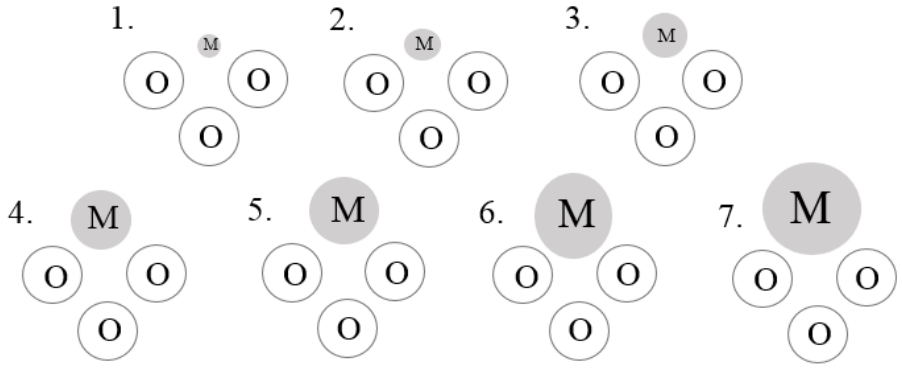
In particular, in order to justify asymmetries in the MERIT treatment, we made all children at the very beginning of the experiment participate in a searching task, where they were asked to collect as fast as possible all corks from a basket containing different items. The faster child of two matched children gained more in case of mutual cooperation in the MERIT treatment than her counterpart (this is child 2 on the right hand side in Figure 1.2). Each child always got one marble for this first searching task in order to ensure that no child left the experimental room with no rewards.

1.2.4 PARENTAL QUESTIONNAIRE

At the end of the day of the experiment, we distributed a questionnaire to the children's parents, asking either the mother or the father to fill it in at home and return it to us within two weeks. The questionnaire had three parts: In Part (i) we asked for background information about the family (e.g., education of parents, working status, number of children, single household). Part (ii) contained what we consider the "soft" factors. More

specifically, we elicited the parental perception of their child by using the scales developed to measure the overvaluation and perceived superiority of one's child (Brummelman et al., 2015a). In Figure 1.3 we present the question for superiority and in the appendix we include the questions for the overvaluation. Finally, Part (iii) of the questionnaire was, in fact, a prisoner's dilemma game in a pen-and-paper version. Parents could choose between keeping €50 or passing them on to another, unknown parent. In the latter case, the amount was doubled or tripled, making cooperation by passing on the money the socially most efficient choice but keeping defection still as the dominant strategy for each person. Note that parents took part in the same treatment as their children did, with the only exception of parents whose children participated in the MERIT treatment. Those parents took part in the ASY treatment instead. After we collected all questionnaires, one pair of parents was randomly selected for payment, and parents could earn up to €200.

Figure 1.3: Perceived superiority scale



Notes: Question about perceived superiority of own child (taken from Brummelman et al., 2015a). Parents were asked to answer the following question: "Please write the number of the diagram (1-7) that best represents how you see your child "M" compared to other children "O"?"

In March 2017 we conducted a short follow-up study with parents in order to deepen our understanding regarding our previous findings. For this purpose, we approached parents of kindergarten children from the same villages where we had conducted the first study in 2015. In this second

questionnaire, we presented parents with a prisoner's dilemma game and asked them how they would want their child to act in such a game (see appendix for the questionnaire). In addition, we also asked parents for the reasons why parents in the 2015-study might have perceived their child to be superior to others.

1.3 RESULTS

In Table 1.1 we report the summary statistics of the children who participated in our study. Overall, 328 children took part in our study. Children were either 4 or 5 years old and about half of the participating children were girls (47%). We find that a considerable amount of children, namely 24%, cooperated. This fraction of cooperating children is not different between 4- and 5-year-olds (25% vs. 24%, p=.963, Fisher-exact test). While in our study 27% of the females cooperated, 22% of the males opted for cooperation. This difference in cooperation rates, however, is not significant (p=.440, Fisher-exact test). Most importantly, the fraction of cooperating children does not statistically differ across treatments (see Figure 1.4, pairwise comparisons of treatment differences according to Fisher-exact tests yield p-values>.532[1]). In the following analysis, we will aggregate the data collected in the different treatments and control for treatment condition in the regressions. The dummy variables indicating the different treatment conditions turn out not to be significant (see Table A-1 in the appendix). This suggests that the probability to cooperate is not influenced by the payoff structure in case of mutual cooperation. Expected cooperation (35%) is slightly higher than actual cooperation indicating that children in general expect other to cooperate more often than they themselves are willing to. 85% of the participating children were able to answer all control questions correctly. We include all children in our analysis and control for understanding in our

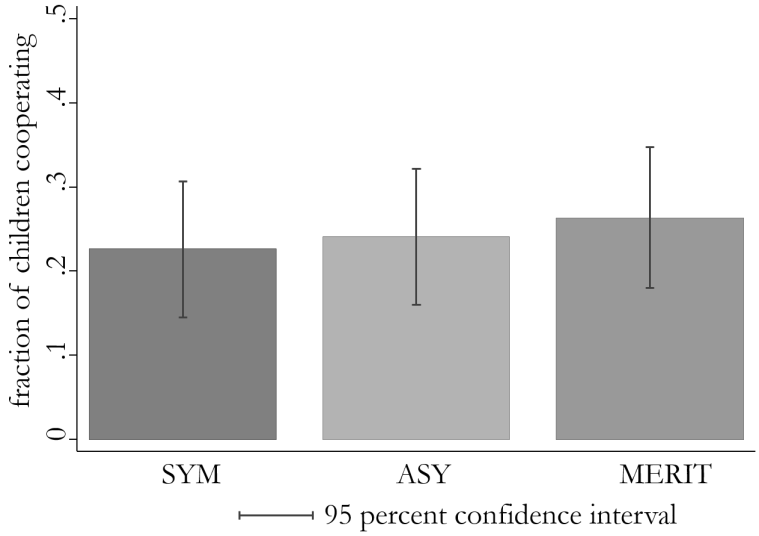
[1]See Figure A-1 in the appendix for the comparison of the 5 possible conditions: symmetric, advantageous, disadvantageous, justified advantageous, and justified disadvantageous asymmetric conditions. Pairwise comparisons with Fisher exact tests yield p-values>.078.

regressions. The dummy variable for understanding turns out not to be significant (see Table A-1 in the appendix).

Table 1.1: Summary statistics

Variable	**Mean**	**Std.Dev**	**N**
Age (in years)	4.491	0.536	328
Female (relative frequency)	0.470	0.500	328
Cooperation (relative frequency)	0.244	0.430	328
Expected cooperation (relative frequency)	0.352	0.478	327
Understood instructions (relative frequency)	0.845	0.363	328

Figure 1.4: Cooperation rates of children by treatment



Considering the parents who participated in our study, we see that in total 156 parents returned the questionnaire completely filled in. This allows us to match the data of 156 parents with their child's behavior in the prisoner's dilemma experiment. In addition, we have 20 parents who returned the questionnaire incompletely, and we use their data when possible.

In Table 1.2 we show that the children whose parents participated completely in our survey do not differ with respect to willingness to cooperate, beliefs, age, gender, or understanding from the children whose parents did not participate (also see appendix, Table A-1 for further comparison of children whose parents (did not) participate). In the following, we will report data of all children and when evaluating the impact of family background focus on those children whose parents participated in the survey.

Table 1.2: Comparison of children's behavior conditional on parents participating in the survey (column [1]) or parents not participating in the survey (column [2]).

Variable	**[1] Parents participated**	**[2] Parents didn't participate**	**p-value**
Cooperation	0.26	0.23	0.616
Expected coop.	0.32	0.38	0.260
Age (in years)	4.47	4.51	0.452
Female	0.47	0.47	0.957
Understood inst.	0.87	0.82	0.195
No. of observations	156	172	

Notes: The p-value refers to two-sided Mann-Whitney-U-tests comparing columns [1] and [2].

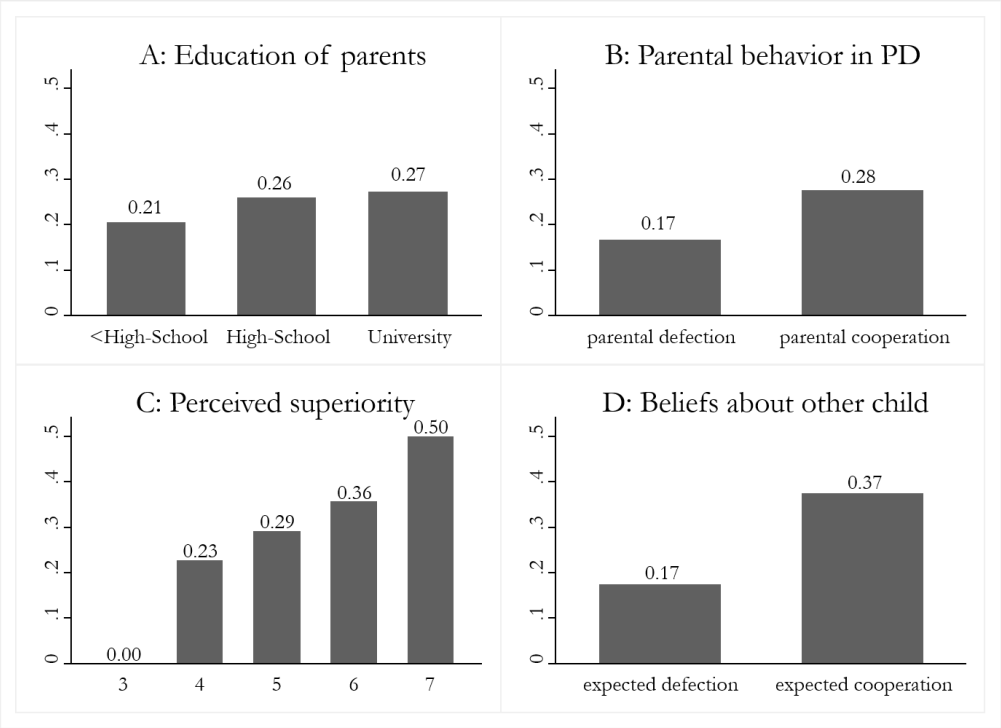
To increase our understanding regarding the determinants of heterogeneity in children's willingness to cooperate, we concentrate on factors which potentially correlate with a child's cooperation behavior. For this purpose, Figure 1.5 illustrates children's cooperation rates conditional on parental education (Panel A, Figure 1.5), parental cooperation (Panel B), perceived superiority (Panel C), and children's beliefs (Panel D). Panel A shows that children from families with higher education (measured here as the highest degree attained by either father or mother) are more likely to cooperate. While 21% of children from families where both parents obtained less than

13 years of schooling cooperated, this fraction increases monotonically as parents have higher education. The rate is 26% for children from families with a high-school degree of one or both parents and 27% for children from families where at least one parent holds a university degree. The correlation between parents' highest degree and children's cooperation becomes statistically significant in our regression in Table 1.3 when we control for further children and family specific characteristics. Panel B illustrates that children whose parents cooperated display an average cooperation rate of 28%, while the rate for children whose parents defected is only 17%. Although quite sizeable, this correlation lacks of statistical significance (Fisher-exact test, p=.216).

Contrary to our expectations, and quite surprisingly, parents' perceived superiority of their child has a significantly positive effect on the level of cooperation (see Panel C). Specifically, there is a monotonic increase of the likelihood to cooperate with the superiority score (p =.042, Jonckheere-Terpstra test, ordered alternatives in either direction). Note that the score of superiority ranged from 1 to 7 with scores of 1 or 2 never being assigned by parents. A score of 4 indicates that the own child is judged as equal to peers; higher numbers indicate that the own child is perceived as superior to others. In this sense, the more superior a child was ranked by his or her parents, the higher was his or her likelihood to cooperate. Finally, Panel D highlights that children who believed the other child cooperated were cooperative in 37% of cases, while only 17% of children cooperated when they believed that the other child defected (Fisher-exact test, p<.001). We interpret this as evidence for conditional cooperation potentially emerging at an early age already.

When matching children's average cooperation rates with their beliefs about their interaction partner's decision, Figure 1.6 reveals an interesting aspect. This figure shows the average cooperation rates and beliefs of children conditional on the parents' highest degree of education. For children from families where neither the father nor the mother have a high-school degree (less than thirteen years of schooling), the difference between average beliefs and cooperation rates is 20 percentage points large (and statistically

Figure 1.5: Cooperation rates of children contional on parents' education, parents' cooperation, perceived superiority, and children's beliefs

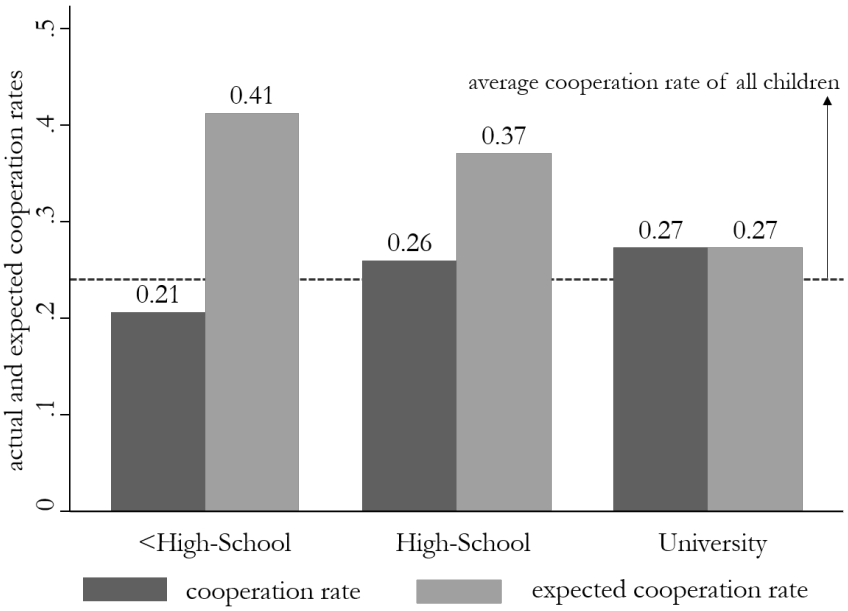


Notes: Panel A: Cooperation rates conditional on highest education of parents. **Panel B**: Cooperation rates of children conditional on the parents' own behavior in the prisoner's dilemma game. **Panel C**: Cooperation rates conditional on perceived superiority of the child by their parents' answer to the superiority scale. **Panel D**: Cooperation rates conditional on a child's belief about the choice (defection or cooperation) of the other child.

significant: p=.0348; two-sided Wilcoxon signed rank test). In the aggregate, these children expect others to cooperate much more frequently than these children on aggregate are willing to cooperate themselves. The middle column for children from families with a high-school degree of one or both parents shows that the difference between average actual and expected cooperation rates shrinks to 11 percentage points, and this difference vanishes completely (to zero percentage points) for children from families where at least one parent has a university degree. This indicates that children from higher educated families are better calibrated in how accurately their

aggregate beliefs of others' behavior match their own aggregate behavior than children from lower educated families. In addition, Figure 1.6 reveals that the group of children from families where at least one parent holds a university degree, in the aggregate, does well in assessing the actual cooperation rate of all children (dashed line in Figure 1.6).

Figure 1.6: Cooperation rates and beliefs of children by parental education



We run a PROBIT regression in order to deepen our understanding regarding cooperation in childhood. Table 1.3 reports the results of the regression displaying average marginal effects and standard errors clustered at the level of present children in the room. As a dependent variable, we introduce children's likelihood to cooperate. The regression confirms what Figure 1.5 has displayed: i) children's beliefs are significant. If a child believes that the other child cooperates, he or she is much more likely to cooperate him- or herself. ii) The relationship between a parent's cooperation and the child's cooperation is quite sizeable and always positive in the estimated coefficient, but it fails significance at conventional levels. iii) Having at least one parent holding a high-school or university degree increases children's likelihood to cooperate. Here it is interesting to notice that the coefficients

of children's beliefs and of parental education are about the same size, meaning that parental education matters as much as children's beliefs in predicting their cooperation rates. iv) Being perceived to be superior by one's own parents goes hand in hand with higher cooperation rates, while the score obtained in the overvaluation scale (see the POS-questionnaire in the appendix) is not significantly correlated with children's cooperation. In addition to these results, the regressions shows that having more siblings is positively related to the likelihood of cooperation. Also, having a mother working part-time has a positive - albeit only weakly statistically significant - effect on children's willingness to cooperate.

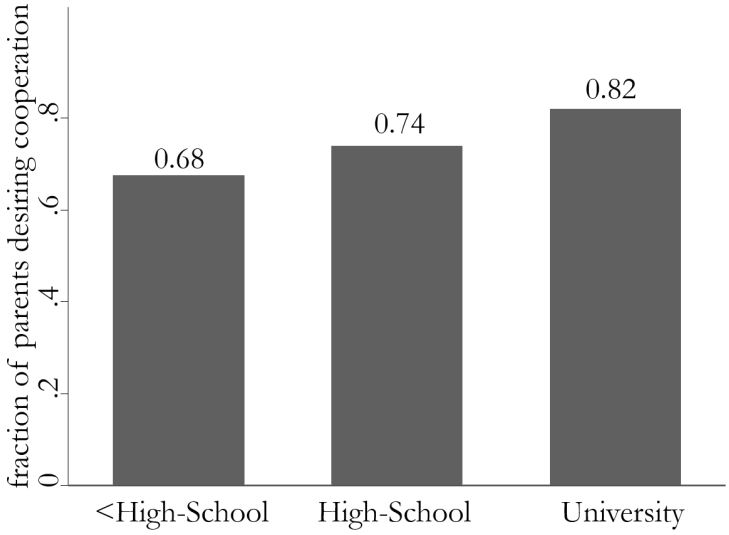
Table 1.3: Probit regression analysis of children's likelihood to cooperate

Dep. var.: cooperation		
Age	-0.015	(0.057)
Female	0.100	(0.062)
Expected cooperation of other child	0.241***	(0.062)
Parental cooperation in PD	0.134	(0.090)
High-School as parents' highest degree	0.231**	(0.109)
University as parents' highest degree	0.234**	(0.111)
Mother staying at home	0.006	(0.096)
Mother part-time working	0.110*	(0.067)
Father part-time working	0.063	(0.179)
Single parent	-0.038	(0.189)
Number of siblings	0.098**	(0.043)
Parental overvaluation (POS)	-0.008	(0.012)
Perceived superiority of child	0.078***	(0.030)
Controls	Yes	
(Number of observations)	156	

Notes: We report average marginal effects. Standard errors are in parenthesis. ***, **, * indicate significance at the 1%, 5%, and 10% level, respectively. Further controls included (all insignificant at the 5% level): dummies for treatment conditions, dummy for not understanding the instructions properly, dummy for presence of a friend in the room, time used to finish the sorting task in stage one of the experiment, dummy for running the experiment in Italian. We clustered the standard errors at the level of present children in the room.

Given our results on the influence of parents' education and their perception of their child, we conducted a short follow-up study. Out of approximately 300 parents that we contacted, a total of 185 parents took part in this study, where we, among other things, asked parents how they think their child should behave when asked to play a prisoner's dilemma game. Figure 1.7 presents the fraction of parents wishing their child to cooperate in a prisoner's dilemma game conditional on parents' education. While 68% of the low educated parents would like their child to cooperate, this fraction increases monotonically as parents have higher education. The rate is 74% for parents with a high-school degree and 82% for parents with a university degree (p =.0697, Jonckheere-Terpstra test, ordered alternatives in either direction). We interpret this as suggestive evidence for education being related to normative expectations of parents about their child's behavior. These normative expectations might be one important channel through which parental education correlates with children's willingness to cooperate.

Figure 1.7: Parents wishing their child to cooperate by parental education



Notes: Percentage of parents who want their child to cooperate when playing the prisoner's dilemma game conditional on highest education of parents (results from follow-up study).

In the follow-up study, we also asked parents for the reasons why parents in the 2015-study might have perceived their child to be superior to others, thus, indicated scores above 4 (i.e., scores 5, 6, or 7 in Figure 1.3) in the perceived superiority scale. The most frequent interpretation of respondents (29%) was that a score above 4 indicates that one's child is particularly popular among peers. This matches the findings of our 2015-study, where parents who rated their child higher in the perceived superiority scale were also more likely to agree with the following statements (taken from the parental overvaluation scale): "Without my child his or her class would be less fun", and "My child is a great example for others to follow" ($p<.001$, $\rho=.357$ and $\rho=.301$, respectively, Spearman correlation). In this sense, a child's popularity might explain the positive correlation between perceived superiority and higher cooperation rates.

1.4 CONCLUSION

Cooperation is important for the welfare of human society. In experimental settings, it has been shown that people cooperate to a substantial degree. All these studies, however, show that people are heterogeneous in their willingness to cooperate. This heterogeneity potentially develops early in childhood. Our study addresses the relationship between children's cooperation and parental background (such as parental education, behavior, and perception) as one possible source for the heterogeneity in humans' willingness to cooperate. For this, we played a prisoner's dilemma game with 328 children and their parents.

First, our results show that children's likelihood to cooperate in a prisoner's dilemma game is associated with parents' education controlling for other children and family relevant factors. Specifically, children whose parents hold a high-school or university degree are more likely to cooperate compared to children whose parents obtained less than 13 years of schooling. In an extension of our study, we tried to understand the channels which explain why parental education and children's cooperation are related. We find that the higher the parental education, the more often parents think

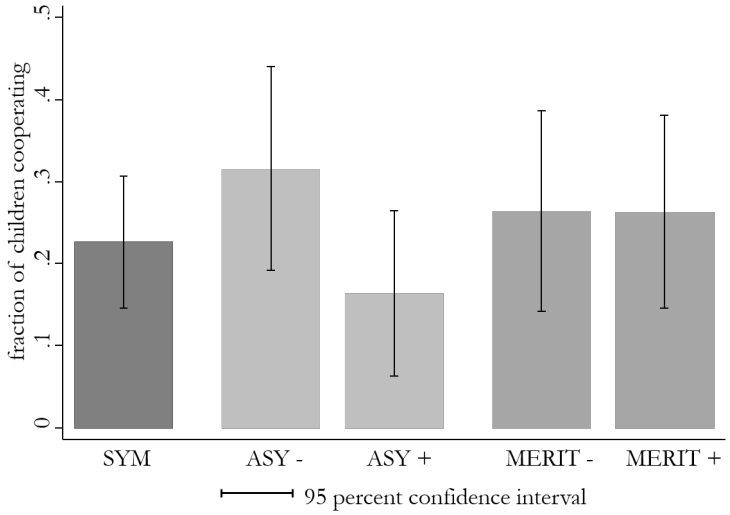
their child should cooperate in a prisoner's dilemma game. From this, we conclude that different educational levels might foster different normative views, which again correlate with children's willingness to cooperate. Interesting to mention is that parental education plays an additional role in shaping children's behavior. While children from lower educated parents expects their counterparts to cooperate more often than they on average are willing to do, this gap in average beliefs and average cooperation rate decreases when parents hold a high-school or university degree.

Second, our results show that although parents and children's choices in a prisoner's dilemma game are well aligned, this association lacks statistical significance. More than parents' behavior, we find that parents' perception of their children correlates with children's cooperation. This is our third result. In particular, we find that the more parents perceive their children to be superior to other children, the more likely it is that children cooperate. Our follow-up study was aimed at shedding light on this (to us) surprising phenomenon. We found that parents often perceive their child to be superior to others if children are particularly appreciated by their peers. This could actually explain the positive correlation with cooperative behavior. Popularity has, indeed, been shown to be related to cooperative behavior in social networks (Currarini et al., 2009; Branas-Garza et al., 2010).

Lastly, we find that children who expect others to cooperate are more likely cooperative too. We do not find asymmetries in case of mutual cooperation to affect children's willingness to cooperate. Taken together, all these findings take us one step further towards understanding what shapes human cooperation. Of course, besides parental influence, other factors like direct and indirect reciprocity or punishment of norm violations play a role for cooperation. It is beyond the scope of our present study to investigate the relative importance of the different factors, e.g., what is the relative impact of parental education compared to the effect of direct reciprocity in peers. Future research might want to address such questions in order to assess the relative importance of the different factors that drive human cooperation.

1.A APPENDIX

Figure A-1: Cooperation rates of children by treatment conditions



Notes: The treatment add-ons (+) and (-) refer to advantageous and disadvantageous asymmetries in case of mutual cooperation. In none of the asymmetric conditions (ASY-, ASY+, MERIT-, and MERIT+) we find cooperation rates which are statistically different from cooperation rates in treatment SYM (Fisher exact tests yield p-values $>.260$).

Table A-1: Probit regression analysis of children's likelihood to cooperate for children whose parents did not participate in the survey (column [1]) and children whose parents participated in the survey (column [2])

Dep. var.: cooperation	**[1] Parents didn't participate**		**[2] Parents participated**	
Age	0.029	(0.062)	-0.003	(0.063)
Female	0.013	(0.064)	0.101	(0.065)
Expected cooperation	0.164***	(0.060)	0.242***	(0.060)
ASY (-)	0.038	(0.100)	0.128	(0.082)
ASY (+)	-0.079	(0.104)	-0.071	(0.100)
MERIT (-)	-0.047	(0.099)	0.163	(0.117)
MERIT (+)	-0.071	(0.096)	0.105	(0.093)
Presence friend	-0.034	(0.073)	-0.063	(0.065)
Conducted in Italian	0.106	(0.067)	0.140	(0.086)
Time searching task	0.000	(0.005)	0.007	(0.008)
No Understanding	-0.122	(0.094)	0.123	(0.099)
N	171		156	

Notes: We report average marginal effects. Standard errors are in parenthesis. ***, **, * indicate significance at the 1%, 5%, and 10% level, respectively. The treatment add-ons (+) and (-) refer to advantageous and disadvantageous asymmetries in case of mutual cooperation. We clustered the standard errors at the session level (i.e., at the set of children present in the room).

INSTRUCTIONS

Instructions translated from German. German instructions available upon request.

Hello! My name is . . . [*name of student helper*]. What is your name? Do you want to play a game with me? Yes, then let's go in the room over there to play the game.

[*Near the entrance area of the room the rewards are placed.*] Look, these are some presents you may earn for this game. But first we will play the game.

So look, here is free space. Let's sit down here. I am happy to have you play this game with me. I will now explain the rules of the game and you will listen very closely. After that you will explain the rules back to me. Ok?

Here is a basket with different items. Can you tell me what you see in the basket? That's right, these are stars and cubicles made of wood and figures and corks. Your job is it to collect all corks and to put them into this glass here. You should do this as fast as you can. I will stop the time. For this game I will afterwards give you one marble. You can later exchange this marble into little presents we have seen over there.

We can start soon. Can you beforehand please tell me what you have to do?

Well done, so we can start now. Are you ready? Ready, steady, go! [*Student helper records the time needed.*]

Well done! Here is the marble you have collected for this first game. Because you did such a great job we will now play another game where you can again earn marbles.

Look this is the game we are going to play now. The game has two different colored sides. Do you know why this is the case? This is the case because you will play the game with another child. Your color is yellow and the color of the other child is green. The child with the green color is from your kindergarten and sits with us in this room. Look in here there are other children too. Do you know the other children? Is one of the children your friend?

Ok, you will play this game with one of them. But I am not allowed to tell you with whom you'll play, because this is a secret. Also the other child does not know that he/she is playing with you.

In this game you can collect yellow marbles and the other child can collect green marbles. Do you know what these are? These are two big boxes. One big box is yellow and it's on your side (yellow side) and the other big box is green and it's on the side of the other child (green side). Two small boxes are hidden in your big box. You can open one of these small boxes. The other child can do the same with his or her big box and can also choose which of the small boxes to open from his or her big box. Let's have a look at what is in the little boxes:

- If you open the yellow box then there is one yellow marble inside. This yellow marble is for you and you can keep it.
- If you open the green box then there are two [*in another treatment: three*] green marbles inside. The green marbles are for the other child. I will take the box, bring it to the other child and the other child will take the marbles.

The other child also decides which of the two boxes he/she wants to open.
- If the other child opens the green box then there is one green marble inside. The green marble is for the other child and the child can take it.
- If the other child opens the yellow box then there are two [*in another treatment: three*] yellow marbles inside. The yellow marbles are for you. I will go to get the box from the other child, bring it to you and you can take

the two [*in another treatment: three*] marbles.

These are the rules of the game. Can you explain me how the game works?
- With whom are you playing the game?
- Which color are the marbles you are collecting?
- How many little boxes inside your big box can you open?
- What happens if you open the small yellow box?
- What happens if you open the small green box?
- What happens if you and the other child both open the yellow boxes? How many marbles does each of you get?
- What happens if you and the other child both open the green boxes? How many marbles does each of you get?
- What happens if you open the yellow box and the other child opens the green box? How many marbles does each of you get?
- What happens if you open the green box and the other child opens the yellow box? How many marbles does each of you get?

[*Student helpers repeat instructions if child cannot answer correctly up to two additional times. Control questions asked in randomized order for each student helper.*]

Well done! You can now decide which of the two little boxes you want to open. I will then go the other child, bring him/her the box you have chosen and I will bring you the box the other child has chosen. Please pull one of the boxes!

Ok, you have chosen this box. Can you explain to me why you did so? You can collect one more marble if you can correctly guess what the other child did. Do you think the other child opened the small yellow or green box?

[*Student helpers exchange the chosen boxes of the matched children.*]

Look, the other child chose the green/yellow box. These marbles are for you [*depending upon outcome:* There are no marbles for you]. [*If applicable:* Because you were right in guessing here is one more marble for you.]

Thank you very much for playing this game with me. You did a great job. You can now exchange all your marbles into presents. We will then put your presents into a little bag and store it with your belongings. You won't be allowed to open the bag until you are at home. At home you can play with the things you have chosen. In the bag there is also a letter for your mum and dad which they should fill in and put it in a box near the entrance.

PARENTAL QUESTIONNAIRE - NOVEMBER 2015 (MATCHED WITH CHILDREN'S CHOICES)

Parental questionnaire translated from German. German version available upon request.

Dear parents, we kindly ask you to fill in this anonymous questionnaire. All answers are voluntary. Thank you for your collaboration!

Demographic information about your person

Gender:
O female
O male

Occupation:
O full-time job
O part-time job
O momentarily unemployed

Highest educational degree:
O Mandatory schooling
O Vocational training
O High-School
O University (Bachelor/Master Degree)

Language, primarily spoken at your home (please indicate only one):
O German
O Italian
O Ladin
O other: .....

Number of children:
O 1 child
O 2 children
O 3 children
O 4 children or more

I'm raising my children:
O alone
O with my partner

Partner's occupation:
O full-time job
O part-time job
O momentarily unemployed

Partner's highest educational degree:
O Mandatory schooling
O Vocational training
O High-School
O University (Bachelor/Master Degree)

Personal evaluation

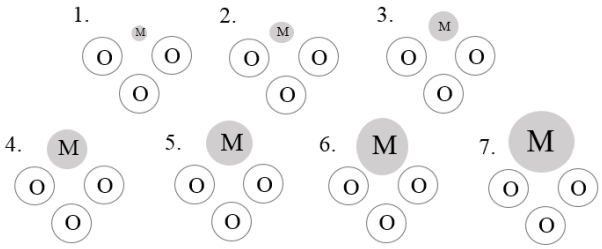
The following paragraph will help us to better understand children's decisions. The following statements refer to how parents see their child. Please indicate for each statement, how well it describes the way you think about your child. If you have more children, please think of the child who is currently enrolled at kindergarten. If you have more children who attend kindergarten, please think of the oldest.

1= not applicable, 2=partially applicable, 3=fully applicable;

Without my child his/her class would be less fun.	1	2	3
My child deserves to get special treatment.	1	2	3
I would not be surprised to learn that my child has extraordinary talents and abilities.	1	2	3
I would find it disappointing to learn that my child was just a "regular" child.	1	2	3
My child is more special than other children.	1	2	3
My child deserves something extra in life.	1	2	3
My child is a great example for other children to follow.	1	2	3

My child compared to others

Please indicate which diagram (1-7) best represents how you see your child (M) compared to other children (O). Please mark the corresponding number.



Game

Lastly, we invite you to participate in a game. If you are lucky you will earn some money. For this game, your decision will be paired with the decision of another parent of the kindergarten of your son/daughter. You and the other parent form a pair. Neither you nor the other parent knows with whom you are partnered. You will both receive one 1 coin (coin= currency in this game) and you have to make a decision:

- You can either keep your 1 coin.
- Or you can pass your 1 coin on to the other parent. In the latter case the coin will be doubled so that the other person will gain 2 coins.

The other person in your pair will make the same decision and will also decide whether to i) keep 1 coin, or ii) pass 1 coin on to you. In the latter case the coin will be doubled so that you will gain 2 coins.

Four possible situations might occur:

1. Both you and your counterpart keep the 1 coin. In this case both of you have 1 coin each.

2. You pass your 1 coin on while your counterpart keeps his/her 1 coin. In this case you will have no coin while your partner will have 3 coins.

3. Both you and your counterpart pass the 1 coin on to each other. In this case both of you end up with 2 coins each.

4. You keep your 1 coin while your counterpart passes his/her coin on to you. In this case you gain 3 coins while your counterpart gets no coin.

Please indicate your choice: ◯ I want to keep my 1 coin. ◯ I want to pass my 1 coin on to the other person.

For the last part of the questionnaire, one pair of parents out of all participating pairs (approx.. 300 parents) will be randomly chosen. The decisions of these two parents will be paid out. For each coin parents have earned in the game they will receive €50. In your envelope you find an orange piece of paper with a code. For your payment it is important that you copy this code which consists of numbers and letters correctly:

My code is the following: .....

Please store the piece of paper with your code in a safe place. The codes of the two winners will be published in the kindergarten. If you are one of the winners, you have to hand in the piece of paper with your code in order to get your money. You will receive your money in a sealed envelope. We will additionally use the codes in order to match the decisions of the children with those of their parents. Since we neither know the complete names of the children nor those of the parents the codes guarantee the full anonymity of all participants.

Please submit the filled in questionnaire in the corresponding box in the kindergarten.

THANK YOU FOR YOUR COLLABORATION!

PARENTAL QUESTIONNAIRE - FOLLOW-UP STUDY (MARCH 2017)

Parental questionnaire translated from German. German version available upon request.

Dear parents, we kindly ask you to fill in this anonymous questionnaire. All answers are voluntary. Thank you for your collaboration!

PART A: DEMOGRAPHIC INFORMATION ABOUT YOUR PERSON

Gender:
O female
O male

Age: .... years

Occupation:
O full-time job
O part-time job
O momentarily unemployed

Highest educational degree:
O Mandatory schooling
O Vocational training
O High-School
O University (Bachelor/Master Degree)

Language, primarily spoken at your home (please indicate only one):
O German
O Italian
O Bilingual: German & Italian
O other: .....

Number and age of children:
O 1 child, age: .... years
O 2 children, age: .... years & .... years
O 3 children, age: .... years & .... years & .... years
O 4 children or more, age: .... years & .... years & .... years & .... years & .... years

I'm raising my children:
O alone
O with my partner

Age of my partner: .... years

Partner's occupation:
O full-time job
O part-time job
O momentarily unemployed

Partner's highest educational degree:
O Mandatory schooling
O Vocational training
O High-School
O University (Bachelor/Master Degree)

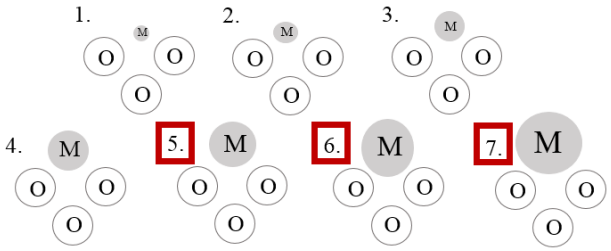
Net monthly income of our family (voluntary disclosure):
O below €1,500
O €1,500- €2,500
O €2,500- €3,500
O above €3,500

We live in a:
O rental flat/house

O own flat/house

PART B: PERSONAL EVALUATION

In our last study, we asked parents to indicate how they see their own child (M) compared to other children (0). The "M" circle illustrates their own child, while the "O" circles illustrate other children. Parents were asked to choose one of the numbers (1-7) which best describes how they see their own child compared to other children. We would like you to interpret the answers given by parents.



What do you think, how did parents interpret scores 5, 6 or 7? In these cases parents attributed a bigger circle to their own child than to other children. In your opinion, what is the reason for parents to choose numbers 5, 6 or 7? (please indicate only one possibility)

Parents, who chose numbers 5, 6 or 7 thought:

◯ that their child is taller than other children.

◯ that their child is in general superior to other children.

◯ that their child is particularly appreciated by his or her peers.

◯ that they can express their strong love for their child.

◯ None of the previously given possibilities describes what I think. I think parents attribute a bigger circle to their child for the following reason:

.....

During our last study, we asked children to participate in a game. The game was structured in a similar way as the following:

Each child got another child assigned as a partner. The identity of the child's partner was kept secret.

Each child got one marble. Then both children were simultaneously asked to either keep the marble or pass it on to the other child without knowing the decision of their counterpart.

◯ KEEP: Whenever a child kept his/her marble, the child gained 1 marble (the marble could later be exchanged into a small present).

◯ PASS: Whenever a child passed his/her marble on to his/her (unknown) counterpart, the marble was doubled so that the other child gained two additional marbles.

The following situations could occur:

1. Both children passed on their marbles. In this case each child gained 2 marbles each.
2. One child passed on the marble and one kept it. The child who passed on the marble had no marble at all, while the child who kept the marble gained 3 marbles.
3. Both children kept their marble. In this case each child gained 1 marble each.

We now kindly ask you to answer two questions. When answering please think of your child who is currently enrolled at kindergarten. If you have more children enrolled at kindergarten please think of the oldest of your children attending kindergarten.

In your opinion, how would your child act in a similar situation?
◯ I think my child would keep his/her marble.
◯ I think my child would pass his/her marble on.

In your opinion, how should your child act in a similar situation?
◯ I think my child should keep his/her marble.
◯ I think my child should pass his/her marble on.

For your information: Our previous study in the kindergarten was completely anonymous. We have no possibility to check how your child - if it participated in that study - behaved in a similar situation.

PART C: LOTTERY

You have now the chance to win €50 if you correctly guess the right statement. In our previous study we have - among other things - analyzed whether a child's willingness to cooperate is correlated with the way parents see their child compared to other children (figure in Part B). Among all parents who correctly anticipate the right statement we will randomly choose 5 parents to receive €50 in cash.

Children who in the eyes of their parents occupy a bigger circle than others
◯ cooperate more.
◯ cooperate less.
◯ do neither cooperate more nor less.

You are now invited to generate a code composed of letters and numbers which you are kindly asked to store in a secure place. Should you correctly guess which of the previous statements is right and your code will be randomly chosen for payment you can pick up your prize at the kindergarten. To claim your prize you have to report your code to the kindergarten teacher. The kindergarten will publish the first four elements of the winning code. It is your responsibility to complete the code correctly in order to claim your prize.

My code is the following: ... ...
1. ... (third letter of the first name of my child)
2. ... (second letter of the surname of my child)
3. ... (months of birth (01 - 12) of my father)
4. ... (the last two digits of the year of birth of my mother)
5. ... (second letter of the first name of my child)
6. ... (third letter of the surname of my child)

â

This code guarantees the full anonymity of all participants.

PART D: FINAL QUESTIONS

We kindly ask you to answer some final questions. In the following we report some traits and abilities which can be fostered by education. Please tell us how important it is for you to achieve the following educational goals.

1= not at all important, 2= rather unimportant, 3=indifferent, 4= rather important, 5= very important

That a child ...					
goes along well with others.	1	2	3	4	5
is interested in how and why certain things happen.	1	2	3	4	5
is honest.	1	2	3	4	5
has good manners.	1	2	3	4	5
possesses self-control.	1	2	3	4	5
is responsible.	1	2	3	4	5
has regards for others.	1	2	3	4	5
conforms to his/her parents.	1	2	3	4	5
possesses the capacity of good judgement.	1	2	3	4	5
is tidy.	1	2	3	4	5
tries to achieve his/her goals.	1	2	3	4	5
fits well in a team.	1	2	3	4	5
learns, to prevail against opposition in life.	1	2	3	4	5
is happy with what he/she has and is.	1	2	3	4	5
learns to avoid risks in life.	1	2	3	4	5
is happy with what he/she has and is.	1	2	3	4	5
is liked by others, is adorable.	1	2	3	4	5
shares with others.	1	2	3	4	5
is helpful, if someone gets injured, is ill or sad.	1	2	3	4	5
helps others (children, adults) voluntarily.	1	2	3	4	5

If you think at an ordinary week who - persons and institutions - except of you take care for your child (please indicate the corresponding)?

PERSONS:
◯ partner
◯ father/mother of the child who does not live with the family
◯ grandparents of the child
◯ older sisters and brothers of the child
◯ other relatives
◯ nanny
◯ other: .....

INSTITUTIONS:
◯ kindergarten
◯ social associations or organizations
◯ others: .....

◯ No, nobody.

Please insert the filled in questionnaire within Wednesday, 22nd of March in the corresponding box in the kindergarten.

THANK YOU FOR YOUR COLLABORATION!

CHAPTER 2

BUSY LITTLE BEES – AN EXPERIMENT ON DILIGENCE AND ENDOGENOUS TIME SCHEDULING IN EARLY CHILDHOOD

Abstract

Diligence, being able to work hard, is a positive predictor of educational success. In an experimental setting we analyze the development of diligence and the impact of exogenous versus endogenous time scheduling on effort provision in early childhood. A total of 429 children aged 3 to 6 worked on a real effort task for as long as they wanted to. Giving the children the possibility to decide when to work on the task - either today or tomorrow - shows that those who procrastinate the task to the following day provide significantly less effort. While younger children are more likely to procrastinate, the effort provision under procrastination is lower for all age groups. In addition, we find that children's ability to work extensively on a task relates to their willingness to challenge themselves in an unrelated task. Our results shed light on the development of diligence and can be used as a first step for designing interventions fostering the skill of being hard working in a real effort task.

This chapter is joint work with Matthias Sutter and Claudia Zoller.

2.1 INTRODUCTION

The importance of non-cognitive skills has become prevalent in the economics and psychology literature, emphasizing their influence on lifetime success, health, and education and asserting them as a fundamental requisite in the labor market (Heckman et al., 2006; Duckworth et al., 2007; Roberts et al., 2007; Almlund et al., 2011; Kautz et al., 2014). Among these non-cognitive skills grit has been proven to be highly indicative of educational achievement (Duckworth et al., 2007; Duckworth and Quinn, 2009; Burks et al., 2015), outperforming IQ as a predictor of success. Grit is defined as the ability to work persistently on a task, related to being self-disciplined, setting long-term goals, and pursuing them in response to negative performance feedback. Grit is highly correlated with long term success including higher earnings even when controlling for schooling (Diaz et al., 2013), predicts employees' likelihood to keep their jobs (Eskreis-Winkler et al., 2014), and is an indicator of innovativeness and success for entrepreneurs (Mooradian et al., 2016).

Alaoui and Fons-Rosen (2016) split grit into two components - tenacity, a nuanced interplay of perseverance and stubbornness, and diligence, the notion of being hard-working. We focus on the latter aspect, diligence, as it plays an especially critical role as a predictor of educational success - even more so than tenacity - during childhood and adolescence (Alaoui and Fons-Rosen, 2016). With this study we aim to further examine the development of diligence, the influence of time scheduling on effort provision, and its driving determinants. We extend our focus to early childhood, namely 3- to 6-year-old children, as this age span has already proven particularly relevant for the formation of preferences (Fehr et al., 2008).

Studying diligence in childhood is particularly important since non-cognitive skills have been proven to be malleable in these early years (Almlund et al., 2011; Kautz et al., 2014; Alan et al., 2016). In a recent study Alan et al. (2016), for instance, stage a successful school intervention for 10- to 11-year-old children aimed at improving children's beliefs about the malleability of their own ability and the role of effort in the skill accumulation

process. The authors find a significant increase in children's willingness to undertake a more challenging and rewarding task, a decrease in the likelihood to give up after failure, and an improvement of ability accumulation, consequently improving children's success and their payoffs.

We investigate how giving children the decision power of when to do a tedious task impacts their diligence in the task. While even very young children face increasing pressure of time schedules - be it for their leisure activities or later on class schedules - we analyze whether young children's effort provision can be increased by allowing them to choose the time setting for themselves. Companies, for instance, increasingly offer flexible work hour arrangements to their employees (Katz and Krueger, 2016). This practice could indicate a positive correlation between endogenous time setting and effort provision. However, little empirical evidence in this domain exists (Shepard et al., 1996; Wolf and Beblo, 2004; Mas and Pallais, 2017). We address this vacancy and investigate whether children's effort provision is dependent on having the decision power over their "work" schedule. Specifically, we consider whether children's level of diligence can be increased by giving them more flexibility on when to do a tedious task. For this purpose, we measure children's diligence in a real effort task under exogenously given or endogenously set scheduling.

We find that letting children decide autonomously when to work on the real effort task does not yield a more efficient outcome compared to exogenously imposing the schedule upon them. However, we also find evidence of a negative procrastination effect on diligence. Children who actively postpone the task to the next day display significantly lower levels of diligence by producing a lower output in the real effort task. Procrastination in our sample is mostly driven by younger children (3- and 4-year-olds), however the effort provision under procrastination is vastly lower irrespective of age. This is in line with findings by Alaoui and Fons-Rosen (2016), who show a higher likelihood of procrastination in less diligent adult subjects.

Furthermore, we consider the underlying determinants of diligence and whether the decision to challenge oneself serves as an indicator of more diligent behavior. Both diligence and the willingness to challenge oneself

are important components of grit. While experimental papers by Gerhards and Gravert (2015) and Alan et al. (2016) so far focus on grit as one single fundamental skill, we consider diligence and the willingness to take a challenge separately. By teasing these two aspects apart, we are able to examine whether children who are hard workers are also more willing to challenge themselves in an unrelated task. Both aspects have individually been proven to be important for later success. While we have emphasized the role of diligence for educational success, a study by Ashby and Schoon (2010) assesses the importance of the willingness to challenge oneself for life outcomes. Specifically, the authors find that young people for whom it is important to succeed at their job earn more money in adulthood compared to their less ambitious[1] peers. Niederle and Yestrumskas (2008) observe heterogeneity in students' willingness to seek a challenge. Male students chose a difficult task 50% more frequently than women did even when controlling for actual or expected performance levels. As Niederle and Yestrumskas (2008) implemented a fixed time span, they cannot make any inference regarding the correlation between seeking challenges and persistence in effort. This study is able to address this point.

In particular, we find support for more diligent children to be more likely to choose the challenging task over the easy option. In our experiment we presented children with two identical puzzles with different levels of difficulty, where the difficult puzzle yielded a higher reward. We find that the mere choice of the difficult puzzle is highly indicative of being more diligent in the real effort task. Children were also given the option on whether to actually follow through on their choice and complete the puzzle (by themselves, after the main part of the experiment was over) or whether to shirk from their decision. We find that those who follow through on their choice display more diligent behavior.

Literature concerning economic decision making during childhood and adolescence has highlighted the importance of time preferences for children. More patient children who are willing to wait for larger rewards have been

[1]Ambition stands for teenagers' willingness to be challenged in their job and to move up.

shown to have higher grades at school, better conduct, and are less likely to engage in health damaging behavior like smoking or drinking alcohol (Castillo et al., 2011; Golsteyn et al., 2013; Sutter et al., 2013; Alan and Ertac, 2018). We, therefore, elicited time preferences to investigate whether delay of gratification for a greater reward and working for a greater reward are interrelated. Our sample shows that younger children who are more impatient are also less diligent, but we fail to find an overall influence of patience on diligence. This is in line with findings of Non and Tempelaar (2016) who report no correlation between university students hypothetically elicited time-preferences and their study effort measured by the time they were logged in on an electronic learning platform, their number of solved exercises on this platform, the fraction of topics completed on the platform, as well as their participation in an online summer course. Similarly, Gerhards and Gravert (2015) report no significant correlation for adults between self-reported, unincentivized time preferences and the decision not to shirk in a real effort task.[2] Likewise, Burks et al. (2012) do not find any evidence on truck drivers' β , δ values (their discount rates for present and future delays) and their tendency to stay at least 6 months on the job.

Lastly, we consider family background as an explanatory variable for differing levels of diligence. Family background has been shown to heavily influence children's behavior. Socio-economic background, for instance, affects children's social-, time-, risk-, and competitive preferences (Bauer et al., 2014; Deckers et al., 2015; Almås et al., 2016, 2017; Deckers et al., 2017). We use a parental questionnaire to elicit demographic data, and self-assessed levels of diligence and procrastination, as well as incentivized time preferences. We provide suggestive evidence that parents with higher education have more diligent children. We find no evidence of parental diligence to be correlated with their children's diligent behavior. However, parents who report to procrastinate more frequently have children who procrastinate more often in the experiment.

[2]Gerhards and Gravert (2015) run a real effort task where students were asked to solve anagrams. They consider shirking as the decision to skip anagrams as well as the decision to switch to easier anagrams.

Overall, our paper sheds light on the development of diligent behavior in early childhood. The following sections give a detailed insight into the design and the procedures. Section 2.3 discusses the results, section 2.4 focuses on the influence of family background, and section 2.5 concludes.

2.2 DESIGN

We ran an experimental study with 3- to 6-year-old children in eight different kindergartens in Innsbruck, Austria.[3] Overall 429 children, among those 219 (51%) females, participated in our study (see Table 2.1). Children were paid in tokens which could be exchanged one-to-one for small presents like balloons, hair clips, key chains etc. Each child received one show-up token at the beginning of the experiment. All decisions were collected anonymously by assigning a code to children. The experimenter additionally emphasized that the child's answers were to be kept a secret to avoid spill-over effects.

Table 2.1: Number of observations

age	**male**	**female**	**total**
3	21	22	43
4	64	72	136
5	73	69	142
6	52	56	108
	210	219	429

We visited each kindergarten on two (or three - if the number of children was very large) consecutive days. Each session followed the same procedure. Children were asked by a trained experimental assistant (experimenter,

[3]The experiment was approved by the ethics committee of the University of Innsbruck and the municipal authorities of the city of Innsbruck. Kindergartens were informed about our study by the city authorities but were blind to the research question. 8 kindergartens participated in our study. 6 kindergartens used an opt-out option, where parents could inform teachers if they did not want their child to participate, which only one parent did. Two kindergartens required an opt-in option where over 70% of parents consented.

henceforth) whether they wanted to participate and were then accompanied to a separate "game" room (only two children opted out of participation). On day one all children were seated one-on-one with an experimenter for the first part of the experiment, namely the elicitation of time preferences, the puzzle task, and the explanation of the diligence task - a real effort task. On day two children were able to collect their payoffs for the delayed tokens of the time preference task. Additionally, children in the *tomorrow* treatment option were brought back to complete the diligence task (more information in subsection 2.2.3). At the end of each session, the child was able to trade her tokens for presents and was then accompanied back to the teacher. To ensure comprehension all children had to answer questions for each task (see appendix for more details).[4]

2.2.1 TIME PREFERENCES

We adapted the convex budget set procedure developed by Andreoni and Sprenger (2012) and presented children with three options of consumption allocation.

- Option 1 yielded 2 tokens today and none tomorrow.
- Option 2 yielded 1 token today and 2 tokens tomorrow.
- Option 3 yielded no tokens today and 4 tokens tomorrow.

By measuring children's preferences for delaying gratification to the next day we assess children's level of patience. Children opting for option 1 are, therefore, classified as impatient, whereas very patient children will delay all gratification to the next day doubling the amount of tokens received. To make the *tomorrow* payoff more salient, children collected their tokens for *today* in one bag and tokens for *tomorrow* were put in a separate bag with the child's name written on it. The *tomorrow* bags were returned to

[4]Overall, 91% of the children could answer all control questions correctly, indicating that we succeeded in sufficiently explaining the task even to the youngest children. Our results remain valid when we exclude all children who did not answer all control questions correctly from our sample.

children the following day to allow children to exchange the saved tokens for additional presents.

2.2.2 PUZZLE TASK TO CHALLENGE ONESELF

The puzzle task measured children's willingness to challenge oneself. Children were presented with two puzzles with the same picture which differed in piece size and number of pieces. The experimenter showed the child two identical puzzle boxes and took out one piece each to show the difference in piece size (see Figure A-1). The puzzle with larger pieces was identified as easy and the puzzle with smaller pieces as difficult. Children were asked to repeat why the puzzle was either easy or difficult. Taking into account that the difficulty of doing such a task differs with age, we presented children aged 3 and 4 with an easy puzzle with 6 pieces, while the difficult puzzle entailed 12 pieces. 5- and 6-year-olds were given the option of an easy puzzle with 12 pieces and a difficult puzzle with 24 pieces. For all age groups the completion of the easy puzzle yielded 1 token and the completion of the difficult puzzle paid 2 tokens. The tokens were presented next to the respective puzzle to make the payoff difference salient. During the experiment, children's choices of easy versus difficult were recorded and the chosen puzzle was handed over for the child to keep. Children were not required to complete the puzzle, however, they were informed that in order to receive the tokens for the puzzle, they had to complete it at the very end of the experiment by themselves.[5]

2.2.3 DILIGENCE TASK

As a third task children were introduced to the diligence task. It consisted of a real effort task where children were told to collect only yellow beads from a bowl of small, multicolored beads (see Figure A-2). Children could autonomously decide for how long they worked on this task and how many

[5] After children received their presents at the end of the experiment, the experimenter reminded them of the potential additional payoff if they completed the puzzle. Children could voluntarily do the puzzle by themselves and once an experimenter verified the puzzle, they received their additional present(s).

beads they collected. They were asked to notify the experimenter (e.g., by raising their hand) once they decided to stop working on the task.

First children were introduced to the task. The experimenter then showed them a bowl with 20 yellow beads and explained that they are worth 1 token. If they were to collect more yellow beads, they would receive more tokens. To control for ability, children were asked to practice the task for 30 seconds.[6] To avoid any external influences, children were seated in a separate "cubicle" while working on the real effort task (RET, henceforth).[7] This was done because of evidence showing that individual productivity is affected by the peers one is surrounded by (Falk and Ichino, 2006; Mas and Moretti, 2009). Once children signaled that they were "done" with the task, the experimenter weighed the sorted beads and paid children their earned tokens. Additionally, the time spent on the RET was recorded as a control measure.[8]

Treatments. To test the effect of time scheduling on diligence, we introduce a between-subjects treatment variation. After completing stages 1 and 2 of the experiment (time preferences and puzzle task) the subject pool was split into two groups where the timing of the RET was set either ***exogenously*** or ***endogenously***. Specifically, after the ability check children in the ***exogenous*** treatment were instructed to either work on the RET **today** or **tomorrow**, while children in the ***endogenous*** treatment autonomously decided whether to work on the RET **today** or **tomorrow**. We, hence, consider four different dimensions: a) *exogenous today*, b) *exogenous tomorrow*, c) *endogenous today*, and d) *endogenous tomorrow*.

Children that (actively or passively) postponed the task to tomorrow were fetched from their group by the experimenter the following day to work on the RET. It was common knowledge to children that the experimenter would return the following day. Children were then seated in their respective

[6]Children were not aware that they were being timed to avoid inconsistent results due to time pressure.

[7]Strict no talking and no peeking rules were enforced during the RET. A different experimenter oversaw the RET to reduce experimenter demand effects.

[8]Again, children were not aware of the time measure to avoid confusion about time pressure. The payoff relevant variable was the output and not the time spent on the task.

"cubicle" and again briefly reminded of the instructions. The remaining procedure closely followed the *today* condition, where children worked on the task for as long as they chose to and were then paid out according to the number of beads collected.[9]

2.2.4 PARENT QUESTIONNAIRE

In addition to collecting data from children, parents received a questionnaire asking for information on demographic variables like the socio-economic status of the family, parents' assessment of their child in terms of diligence, patience, and willingness to take a challenge, and parents' assessment regarding their own behavior. Specifically, parents were asked to fill out the grit scale[10] (Duckworth et al., 2007; Breyer and Danner, 2015), a questionnaire about their tendency to postpone tasks (Lay, 1986), and to state their time preferences. The latter task presented parents with a choice list where they could choose between a fixed amount of money (€50) earlier in time versus a higher monetary reward in the future (max €70). Among all participants five parents were randomly chosen to be paid out. To maintain anonymity, parents received a code that was matched to their child's decisions.

2.3 RESULTS

The result section first gives an overview of the decisions made in each task (see Table 2.2). Section 2.3.2 provides a detailed analysis for treatment results of the effect on time scheduling on diligence. Section 2.3.3 presents the regression analysis and discusses various driving factors of diligence in early childhood.

[9]Note that the order of tasks remained the same for all children for the following reasons: First, we did not want children to be distracted during the RET. Children might cut down on effort and time out of sheer curiosity for upcoming tasks. Second, as the RET is a non-cognitive task and the other tasks demand more cognitive skills we opted to start with the more demanding skills. As all children completed the tasks in the same order, our results also account for possible depletion effects across the experiment.

[10]By considering parents agreement with the statements "I am a hard worker" and "I am self-disciplined" - two items taken from the grit scale used in the PIAAC field trial (Tamassia and Lennon, 2013) - we infer parents' level of diligence.

2.3.1 DESCRIPTIVE ANALYSIS

Time preferences. In the time preferences task we measure children's patience. 40% of children chose the very patient option of delaying all tokens to tomorrow, forgoing all immediate rewards to maximize their payoff. 29% of the children display high impatience opting for two tokens today and nothing tomorrow. The same fraction of children split consumption between today and tomorrow (1 token today and 2 tokens tomorrow) choosing the intermediate option. For the data analysis we use a measure of impatience accounting for the total number of tokens claimed for immediate consumption. In line with previous literature on time preferences, we observe a slight decline of impatient choices with increasing age (p=.047, Cuzick's Wilcoxon-type test for trend) which is more pronounced for strictly impatient choices of two immediate rewards (p=.010, Cuzick's Wilcoxon-type test for trend) as shown in Figure 2.1. Females in our sample are slightly more impatient compared to males (33% vs. 26% for two tokens today; p-value=.084, Mann-Whitney-U-test), choosing the strictly patient option significantly less often (37% vs. 45% for four tokens tomorrow; p-value=.067, Mann-Whitney-U-test).

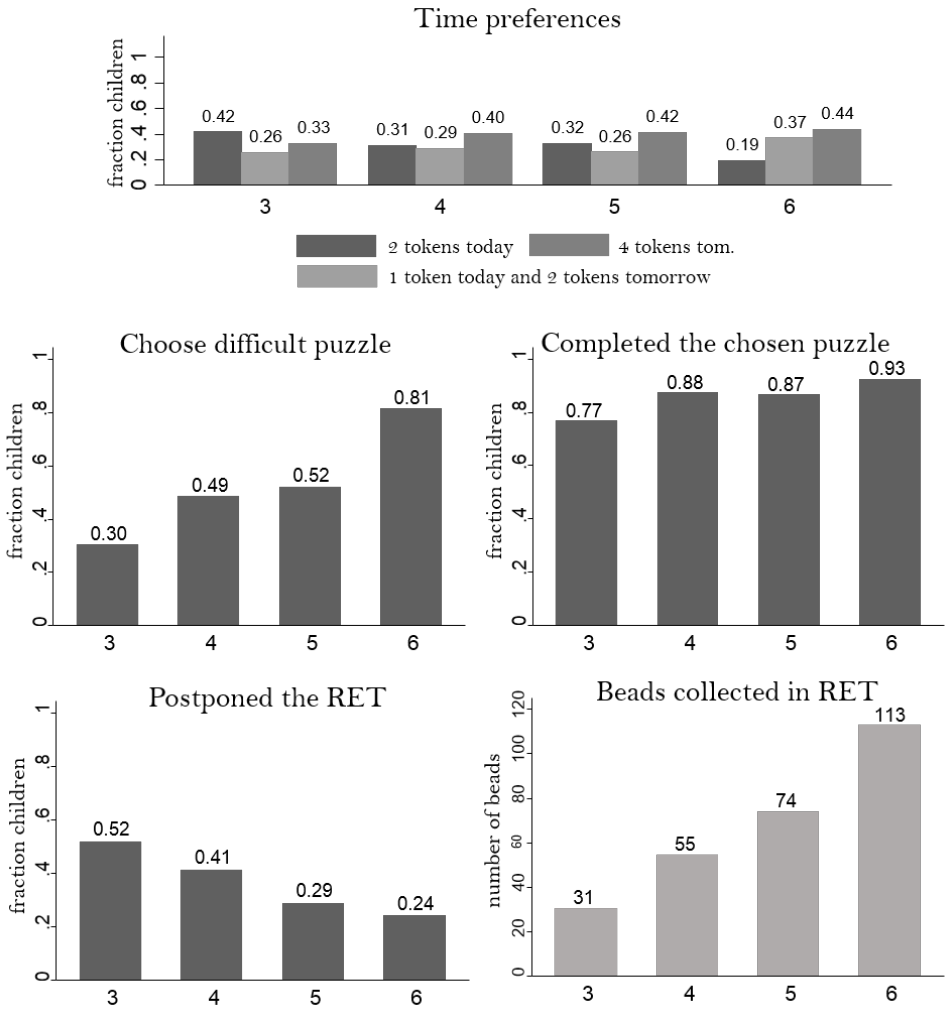
Puzzle task to challenge oneself. The second task measures children's willingness to challenge themselves. Children were asked to choose between an easy and a difficult puzzle. A little more than half (56%) of children opted for the difficult puzzle. As the number of pieces for the difficult puzzle varied between the younger (3- and 4-year-olds) and older cohort (5- and 6-year-olds), we consider children's choices in the respective age groups. We again observe a significant age trend, where the choice of the difficult puzzle increases with age with 30% of 3-year-olds taking the challenge compared to 49% of 4-year-olds (p=.036, Mann-Whitney-U-test), and 52% of 5-year-olds compared to 81% of 6-year-olds (p<.001, Mann-Whitney-U-test). Additionally, we find a significant gender effect for males being more likely to challenge themselves compared to females (61% vs 51%; p=.032, Mann-Whitney-U-test). This is in line with previous findings who also find males to be more willing to seek a challenge (Niederle and Yestrumskas, 2008).

Following through on the challenge. Children only received tokens for the puzzle once they completed it at the very end of the experiment. While the easy puzzle paid one token, the difficult puzzle awarded two tokens. We use this as an additional commitment measure to test for who follows through on their choice. Overall, 87% of children chose to complete the puzzle and earn their additional reward. The fraction of children completing the puzzle does not differ over their choice of easy or difficult (86% vs 89%; p-value= .328, Mann-Whitney-U-test). Again, we find an age trend where the likelihood of completing the task increases with age (p=.025, Cuzick's Wilcoxon-type test for trend). Pairwise comparisons, however, do not yield statistically significant results. Specifically, 77% of the 3-year-olds completed the puzzle compared to 88% of 4-year-olds (p=.087, Mann-Whitney-U-test), and 87% of 5-year-olds compared to 93% of 6-year-olds (p=.133, Mann-Whitney-U-test).

Table 2.2: Summary statistics

Variable	**Mean**	**N**
PATIENCE		
2 tokens today	0.30	429
1 token today, 2 tokens tomorrow	0.30	429
4 tokens tomorrow	0.40	429
WILLINGNESS TO CHALLENGE ONESELF		
chose difficult puzzle	0.56	429
completed the chosen puzzle	0.87	429
DILIGENCE		
postponed the task	0.34	241
beads collected in ability check	10	429
beads collected in RET	73	429

Figure 2.1: Development by age



Diligence task. The third task assesses children's level of diligence in a real effort task measured by the output, namely the number of yellow beads collected. The output increases significantly from 31 beads collected by 3-year-olds, to 55, 74 and 113 beads for 4-, 5- and 6-year-olds respectively ($p<.001$, Cuzick's Wilcoxon-type test for trend). The maximum number of beads collected was 440 confirming great variance for the measured output. Girls display higher levels of diligence, sorting significantly more beads than their male counterparts (80 vs. 66, $p=.0163$, Mann-Whitney-U-test).

Average time spent on the task was 7.9 minutes, increasing with age from 5 minutes for 3-year-olds to 10 minutes for 6-year-olds (7 and 8 minutes for 4- and 5-year-olds respectively, p<.001, Cuzick's Wilcoxon-type test for trend).[11] As expected, the ability check also reveals different levels of ability between younger and older children. While 3-year-olds sorted on average 7 beads in the 30 seconds, this number steadily increased to 12 beads for 6-year-olds (p<.01, Cuzick's Wilcoxon-type test for trend).

The next sections discuss correlations between the experimentally elicited variables and show that children who sort more beads are (i) more likely to choose the difficult puzzle (p<.001, Mann-Whitney-U-test), (ii) more likely to follow through on their choice by completing the puzzle (p<.001, Mann-Whitney-U-test), (iii) less likely to procrastinate the sorting task to the next day (p<.001, Mann-Whitney-U-test), and (iv) more likely to score higher in the ability check (p<.01, Spearman's rank correlation). Concerning the last point, namely children's performance in the ability check, we observe that children who score higher in the ability check are more likely to choose the difficult puzzle (p<.001, Mann-Whitney-U-test), and to complete the puzzle independently of the chosen level of difficulty (p<.001, Mann-Whitney-U-test). The other measured choices are not correlated according to Spearman's rank correlation coefficients at a 5% significance level.

2.3.2 IMPACT OF TIME SCHEDULING ON DILIGENCE

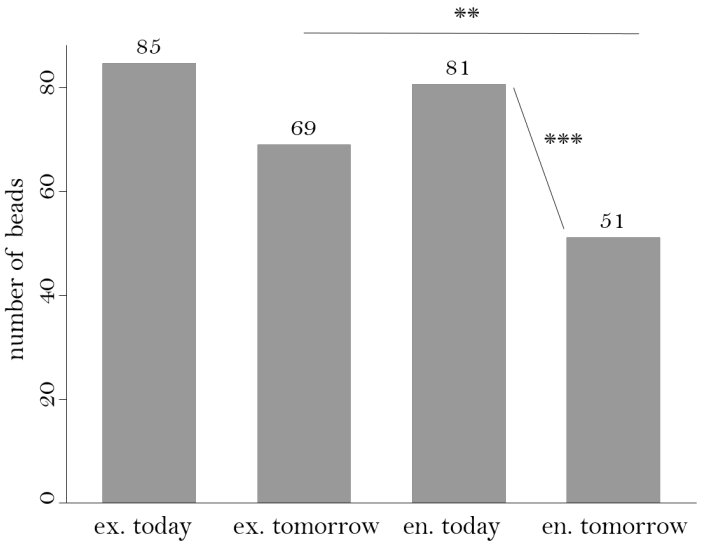
In this section we consider the treatment variations on exogenous and endogenous time scheduling of the diligence task. Children were randomly allocated to the different treatment groups[12]: 188 children participated in the **exogenous** treatment of which 91 were told to do the diligence task today (*exogenous today*) and 97 children were told to do the diligence task tomorrow (*exogenous tomorrow*). The remaining 241 children were assigned

[11] While we tried to time children as closely as possible to their actual working time this measure is not as exact as we wished and leaves some room for errors. Hence, we use the number of beads as our dependent variable for further analysis, relying on time as a separate control measure.

[12] Treatments were randomized at class level.

to the **endogenous** treatment, where they could decide whether to work on the diligence task today or tomorrow. About one third of the children (N=83) in the endogenous treatment decided to postpone the task to the next day (*endogenous tomorrow*), while 158 children preferred to work on the task today (*endogenous today*).

Figure 2.2: Diligence over treatments



First, we examine whether an exogenously given time schedule reveals differences in diligence compared to children who could autonomously decide the scheduling of the diligence task. Children who were given a schedule for either today or tomorrow collected on average 77 beads, while children who decided for themselves when to work on the diligence task collected on average 70 beads. Comparing the *exogenous* versus *endogenous* treatment does not yield any differences in diligence (p=.622, Mann-Whitney-U-test). Thus, letting children decide autonomously when to work on the real effort task did not yield a more efficient outcome compared to exogenously imposing the schedule upon children.

We now continue to split the treatments into their sub-groups of either *today* or *tomorrow*. This additional level of detail reveals diligence to be contingent on the time scheduling of the task (see Figure 2.2). While in the

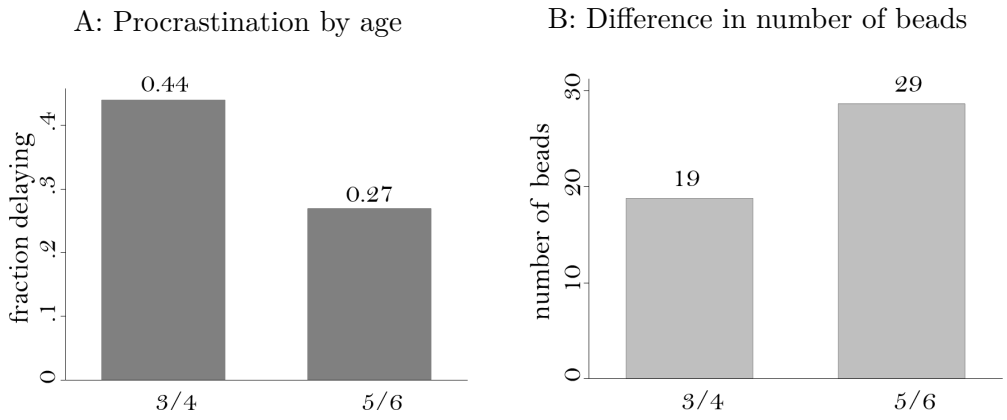
exogenous today sub-treatment children sort on average 85 beads, the level of displayed diligence slightly declines to 69 beads for the *exogenous tomorrow* sub-treatment. This difference in diligence, however, is not significant (p=.234, Mann-Whitney-U-test). Children's diligence levels when actively selecting to work today (*endogenous today*) are on par with children who were told to do the task today (81 vs 85 beads, p=.700, Mann-Whitney-U-test). Those children who actively procrastinated the task *endogenous tomorrow* sort significantly fewer beads (51 beads) compared to all other treatments (p=.002 for exogenous today vs endogenous tomorrow; p=.024 for exogenous tomorrow vs endogenous tomorrow; p<.001 for endogenous today vs endogenous tomorrow, Mann-Whitney-U-tests). We, hence, observe a self-selection effect of children who procrastinate to sort less beads.

Next, we examine age as driving factor for procrastination. We, therefore, split our data into a younger (3- and 4-year-olds) and older cohort (5- and 6-year-olds). We find that younger children in our sample are more likely to procrastinate the task to the next day (see Figure 2.3, Panel A). 44% in the younger cohort and 27% of the older cohort chose to postpone collecting beads (p=.006, Mann-Whitney-U-test).

Result 1. *Younger children are significantly more likely to procrastinate the real effort task to the next day.*

We examine whether lower effort provision under procrastinators is merely driven by the age effect of younger children producing less output. We can refute this claim by looking at the number of beads collected in the *endogenous tomorrow* sub-treatment again split by age groups. Panel B in Figure 2.3 displays the difference in output between *endogenous today* and *endogenous tomorrow*. Both age groups significantly reduce their performance under procrastination (p=.011 for 3- and 4-year-olds, p=.013 for 5- and 6-year-olds, Mann-Whitney-U-test) accounting for a difference in output of approximately 19 beads for younger children and 29 beads for older children. As an additional check, we standardized diligence over age which further supports the difference in output under procrastination (see Figure A-3 and Figure A-4 in the appendix). Children who self-selected into

Figure 2.3: Procrastination



Notes: Panel A reports the fraction of children delaying the RET by age group. The left bar stands for the younger (3- and 4-year-olds) and the right bar for the older (5- and 6-year-olds) cohort. Panel B reports the difference in collected beads between the endogenous today and endogenous tomorrow treatment by age group.

endogenous tomorrow performed significantly worse, independent of their age. Procrastination is, hence, an indicator of lower levels of diligence.

When controlling for other factors, ability seems to play a role in the decision to procrastinate. Table A-1 in the appendix shows that those who have a high ability in the task are less likely to choose to postpone it. Ability is, however, positively correlated ($p<.001$) with age and running separate regressions for the age groups reduces the significance of the ability effect on procrastination.

While effort provision under procrastination is significantly lower, we detect less procrastination behavior of older children. Hence, with increasing age children select more frequently into the *endogenous today* option. From this we can infer that with increasing age not only do children display more diligence in executing a task but also become more proficient in self-management by choosing not to delay the task itself.

Result 2. *Children who self-select into the procrastination option display significantly lower levels of diligence, independent of their age.*

2.3.3 INFLUENCING FACTORS

In this section, we investigate underlying driving factors of diligence by examining the influence of the willingness to challenge oneself and time preferences on the number of beads sorted. First, we consider diligence overall. We find the choice of the difficult puzzle and the willingness to complete the chosen puzzle to be indicators of higher levels of diligence. Then, we split the sample by age groups to deduce the development of the driving factors with age.

The willingness to challenge oneself was measured with the choice between two almost identical puzzles, differing only in number of pieces and therefore in the level of difficulty. The choice of the difficult puzzle ("difficult" dummy in Table 2.3) is representative of choosing the challenge over an easier, lower paying option. A little more than half of the children (56%) chose the difficult puzzle, with boys being more likely to take the challenge. The choice of the difficult puzzle is highly indicative of the level of diligence, sorting on average 20 additional beads as shown in column (1) in Table 2.3, which is equal to one token in payoff.

As an additional measure we control for those who have actually completed the chosen puzzle ("complete" dummy in Table 2.3). While completion of the puzzle was voluntary, the tokens for the puzzle task (1 token for easy, 2 tokens for difficult) were only paid out if the child completed the puzzle. Completing the puzzle is another strong indicator for providing more effort and exhibiting higher levels of diligence. Children who completed the puzzle collected on average additional 16 beads as shown in column (1) in Table 2.3. Irrespective of whether children chose the difficult or the easy task, those who persist and follow through on their choice also collect a greater number of beads. Overall, it is therefore not only beneficial for diligence to be willing to challenge oneself but also to follow through on the choice made.

The variable impatience has a negative coefficient, indicating that more impatient children exhibit lower levels of diligence. For the whole sample it is, however, not significant. Girls in our sample provide significantly more

Table 2.3: Diligence and influencing factors by age groups

Dep. var.: number of beads	(1)	(2)	(3)
	all	3/4 year olds	5/6 year olds
female	14.87***	8.123	20.13**
	(5.531)	(5.162)	(8.627)
ability	5.717***	3.383***	4.832***
	(0.867)	(0.839)	(1.305)
difficult	20.65***	7.841	25.80***
	(5.109)	(5.361)	(7.705)
complete	15.71***	8.836	24.01***
	(5.645)	(5.961)	(9.129)
impatience	-1.502	-7.480**	2.691
	(3.227)	(2.883)	(5.239)
Constant	-14.17	13.92	-13.36
	(9.298)	(8.900)	(15.63)
Observations	429	179	250

OLS regression with robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

output in the RET than boys. Moreover, high ability, which increases with age, shows that those who are more able also provide more output in the RET.[13] [14]

[13]Table A-2 replicates the regression using productivity (number of beads collected over time spent on task) as dependent variable. The influence of female, ability, difficult, and complete are robust to this measure. We consider the measure of number of beads as dependent variable as nonetheless more accurate measure for diligence compared to productivity. While the number of beads measures exactly the output (contingent for payment), productivity only considers the efficiency of children in the task.

[14]In Table A-3 we additionally control for treatment differences. Our results do not change. For the whole sample, we see that the children who decide to procrastinate the RET ("en. tomorrow") collect on average 25 beads less.

Result 3. *The choice of the (more rewarding) difficult task over the easy one is highly indicative of higher levels of diligence. Additionally, following through on that choice (irrespective of the level of difficulty) is highly correlated with higher diligence.*

Columns (2) and (3) in Table 2.3 consider the driving factors for diligence in the respective age groups for 3- and 4-year-olds and 5- and 6-year-olds. While the effect of ability on diligence is stable over all age groups, we observe some distinctions in the importance of the influencing factors between the younger and older age cohort. Most prominently, the influence of the willingness to take a challenge (difficult) and following through on the task (complete) on diligence both emerge at the ages of 5 and 6 but not for younger children. Those children of the older cohort who chose the difficult puzzle collected on average 26 beads in addition, which is more than one additional reward. Those who completed the puzzle show very similar effects of collecting 24 beads in addition. It seems the importance of challenging oneself and the decision not to shirk from completing the puzzle gains importance at age 5 and 6. Impatience, on the other hand, plays a significant role for the younger cohort resulting in a reduction of 7 beads in output. With increasing age, however, this effect disappears. Females become significantly more diligent with increasing age leading to an overall improvement of output of 20 beads.

2.4 PARENT DATA

In this section, we focus on information regarding family background in order to gain better understanding of its influence on children's behavior. Approximately half of parents (48%) agreed to our request and returned the filled in questionnaire.[15] As highlighted in Table 2.4, we find no differences in children's time preferences, their tendency to procrastinate, to choose the difficult puzzle, to actually complete the puzzle, and their performance in the RET task between the samples including parent data and excluding

[15]Some parental questionnaires did not contain full information on all questions, accounting for the difference in number of observations.

parents, making it a representative sample for the children subpool. First, we give an overview of the parent sample and their provided information. Next, we analyze the effects of socio-economic background of parents and the estimation of their child's behavior on children's level of diligence.

Table 2.4: Means by parental participation

Variable	**Parents participated** N=206	**Parents didn't participate** N=223	**p-value M.-Whit.-U-Test**
impatience	0.90	0.87	0.702
difficult puzzle	0.58	0.55	0.524
completion of puzzle	0.87	0.88	0.756
procrastination	0.32	0.37	0.377
ability	10	10	0.741
number beads	74	72	0.896

2.4.1 DESCRIPTIVE ANALYSIS

Among all participating parents 87% are mothers. Our sample of parents proves to be highly educated with 57% reporting at least one parent with a university degree. 71% of mothers report to work part-time compared to 7% of fathers, and 11% of mothers work full-time while 88% of fathers work full time. The vast majority (90%) of the participating families prevalently speak German at home. 10% are single-parents. On average families have 2 children and a monthly net income predominantly between €2,500-3,500 (35%) and in 28% of cases above €3,500.

Focusing on the role of parental education,[16] we observe that parents with higher education displayed more patience in the incentivized intertemporal

[16]The variable for education is coded as an ordinal scale where a value of 1 equals minimum amount of schooling required (9 years in Austria) and the maximum value of 5 equals a PhD.

choice task (p<.01, Cuzick's Wilcoxon-type test for trend). Specifically, we presented parents with a choice list between a fixed amount of €50 earlier in time or an increasing amount of money (between €50-70) three months later. While parents with vocational training or lower education required on average an additional €14 in order to be willing to wait three months for the reward, parents with a high-school degree required an additional €10 and those with a university degree an additional €8.

We find that parents with a higher level of education score higher on the (self-reported) grit scale (p<.001, Cuzick's Wilcoxon-type test for trend) confirming previous findings regarding higher levels of grit and educational achievement (Duckworth et al., 2007; Duckworth and Quinn, 2009; Burks et al., 2015).

2.4.2 PARENT DATA ON DILIGENCE

To investigate the influence of socio-economic background of the child's family on diligence, we asked for information on family composition, occupation, and education. Table 2.5 shows that we do not observe any influence of having at least one parent staying at home full time ("stay-at-home parent") or working part-time ("working part-time") on children's level of diligence. Also, the number of siblings or whether parents raise their children without a partner does not affect the number of beads children sorted. Families' highest obtained degree has a positive - albeit only weakly statistically significant - effect on children's diligence when controlling for parents' own behavior. This effect becomes significant when we additionally control for income in column (3) providing suggestive evidence for parents who hold a higher degree having children who display higher levels of diligence.[17]

[17]When asking parents about family's net monthly income we explicitly framed it as a voluntary disclosure in order not to be invasive. 80% of the parents agreed to answer this question.

Table 2.5: Diligence and parental data

Dep. var.: number of beads	(1)	(2)	(3)	(4)
age	13.78***	13.62***	12.01***	16.62***
	(4.456)	(4.530)	(4.546)	(5.413)
female	12.37*	15.88**	16.60**	18.26**
	(7.391)	(7.265)	(7.504)	(8.460)
ability	5.223***	5.236***	5.809***	5.757***
	(1.159)	(1.147)	(1.106)	(1.391)
difficult	17.48**	16.88**	17.68**	22.11**
	(6.753)	(6.799)	(7.323)	(8.552)
complete	16.86**	14.62*	13.18	13.78
	(8.025)	(8.493)	(9.730)	(12.56)
impatience	-1.424	0.431	0.338	3.577
	(4.406)	(4.453)	(4.550)	(5.065)
siblings		-3.581	-1.854	-6.688
		(4.714)	(4.873)	(6.076)
single		-10.61	-9.303	-20.18
		(12.12)	(14.04)	(19.57)
stay-at-home parent		-3.051	-1.274	-6.258
		(15.69)	(15.71)	(17.27)
working part-time		5.167	6.628	-3.188
		(13.70)	(14.03)	(17.27)
education		4.485	6.808*	9.579**
		(3.341)	(3.782)	(4.473)
parent's procrastination			0.433	0.268
			(0.603)	(0.682)
parent's diligence			-1.645	-5.086
			(3.455)	(3.549)
parent's patience			-1.563	-1.802
			(0.960)	(1.153)
income				-5.165
				(6.513)
Constant	-70.30***	-86.15***	-70.57*	-48.98
	(20.58)	(27.38)	(38.17)	(44.64)
Observations	206	206	197	157

OLS regression with robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Parents were furthermore prompted to fill in self-reported scales on grit (Duckworth and Quinn, 2009; Breyer and Danner, 2015), and procrastination (Lay, 1986) (see appendix for details). The grit scale - besides informing us about parents self assessed level of grit - also helps in inferring parents' self assessed level of diligence by considering two items of the grit scale, namely their agreement with the notions "I am a hard worker" and "I am self-disciplined". We find no evidence of neither parents' assessed level of grit (see Table A-4) nor diligence to influence their child's diligence (see Table 2.5). While parental self assessed procrastination has no effect on children's diligence, we see that parents who score high on the procrastination scale are also more likely to have children who procrastinate (p=.058, Mann-Whitney-U test). In this sense, procrastination behavior seems to be more transferable from one generation to the next than diligence.

Result 4. *Parents who score high on the procrastination scale are more likely to have children who procrastinate. While parental education seems to positively affect children's diligence, there is no evidence for parent's diligence to correlate with that of their child's.*

2.4.3 PARENT'S ESTIMATION OF CHILD'S BEHAVIOR

As a next step, we analyze the correlations between the parent's assessment of their child on an ordinal scale from 1 to 5 and a child's actual behavior. Parents' estimation about whether their child likes to spend a lot of time on a given task[18] and their believed child's level of patience significantly correlate with children's actual behavior in the RET task (p<.001, Spearman's rank correlation coefficients) and the time preferences elicitation (p=.014, Spearman's rank correlation coefficients). The same is true for parents' beliefs about their child's willingness to take a challenge and children's actual choice of the more difficult puzzle (p=.077, Spearman's rank correlation coefficients). Parents' assessed procrastination behavior of their child does not correlate with children's actual procrastination (p=.298, Spearman's

[18]We asked parents whether their child liked to spend a lot of time on any given task and did not specify the RET from the experiment.

rank correlation coefficients), while we have previously shown that parents' self-assessed and children's displayed procrastination behavior does.

Table 2.6: Parental assessment and children's behavior

Parental assessment	**Child's behavior**	**Spearman coefficient**	**p-value**
spends a lot of time on a task	no. sorted beads	.280	<.001
has a hard time waiting	impatience	.178	.014
likes challenging games	difficult puzzle	.128	.077
procrastinates unpleasant jobs	procrastination RET	.104	.298

2.5 CONCLUSION

Our study contributes to the emerging literature on grit - a skill found to be highly predictive of success in life. We focus on diligence, a crucial component of grit, in early childhood. We measure diligence as the output in a real effort task where children aged 3 to 6 were able to not only decide on how much effort to invest but also when to schedule the task. While it did not matter in terms of effort provision whether children were given an exogenous schedule or were able to decide for themselves when to do the task, there is a significant negative effect of self-selected procrastinators on exerted effort. While younger children are much more likely to procrastinate the task, effort provision under procrastination is lower irrespective of the age group. Procrastination behavior is, thus, more pronounced in 3- and 4-year-olds, while 5- and 6-year-olds are significantly more likely to do the task right away.

Considering the determinants and influencing factors of diligence, the willingness to choose the challenging and more rewarding task over the easy task serves as an indicator for the child to provide more effort, displaying more diligent behavior. Additionally following through on the choice, irrespective of level of difficulty, and not shirking is also highly indicative of

being more diligent. In this sense, we are able to show that diligent behavior, taking a challenge, and not shirking from a made choice - all important aspects of grit - highly affect each other. This effect is especially pronounced for older age cohorts.

Additionally we see an age effect, where effort provision increases significantly with age even when controlling for ability. Girls also display more diligent behavior, outperforming boys in the real effort task. When analyzing children's diligence and their family background we find a statistically weak effect of parents' education on their children's level of diligence. While parents self reported diligence does not correlate with children's diligence in the RET, we see that procrastination behavior of parents is correlated with that of children.

Overall, this paper is the first to study the development of diligence and how it is affected by time scheduling in early childhood. From a policy stand point, it is important to foster diligent behavior in children starting at early childhood. While we are able to show that with age children become more proficient, we also show that giving children the possibility to decide when to do a task might help identifying those children who are more in need of active support in conducting work persistently.

2.A APPENDIX

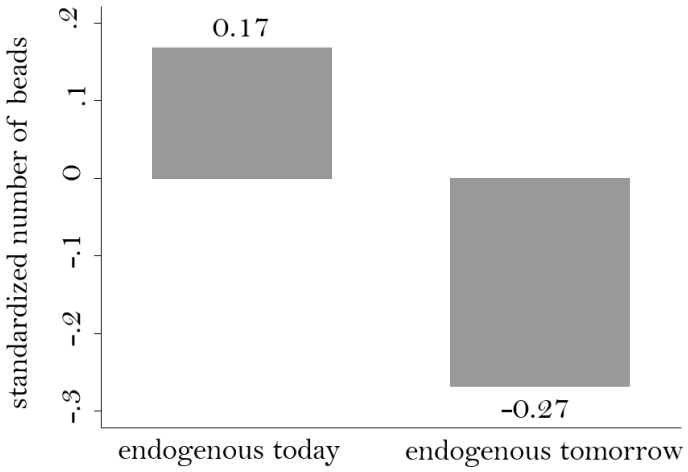
Figure A-1: Puzzle task to challenge oneself

Notes: Children could choose between two different puzzles with the same picture which differed in the piece size and consequently the number of pieces. We presented children aged 3 and 4 with an easy puzzle with 6 pieces while the difficult puzzle entailed 12 pieces. 5- and 6-year-olds were given the option of an easy puzzle with 12 pieces and a difficult puzzle with 24 pieces. The experimenter showed the child the two identical puzzle boxes and took out one piece each to show the difference in piece size.

Figure A-2: Setup diligence task

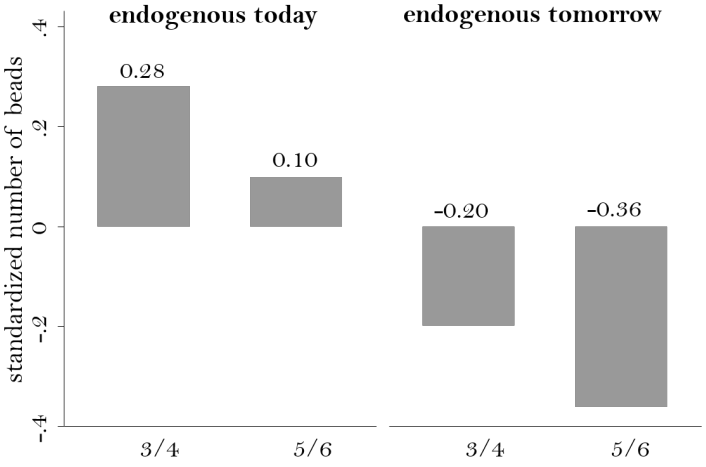
Notes: Children were seated in separate "cubicles" while working on the RET. The RET consisted of collecting the yellow beads out of the green bowl putting them into the yellow bowl.

Figure A-3: Standardized diligence in endogenous treatment



Notes: Diligence standardized over age to account for age and ability.

Figure A-4: Standardized diligence in endogenous treatment by age groups



Notes: Diligence standardized over age to account for age and ability.

Table A-1: Procrastination

Dep. var.: endogenous delay	(1)	(2)	(3)
	all	3/4 year olds	5/6 year olds
number of beads	-0.00172***	-0.00278**	-0.00106
	(0.000636)	(0.00137)	(0.000650)
ability	-0.0218**	-0.0184	-0.0172*
	(0.00847)	(0.0175)	(0.0102)
female	-0.0516	-0.00286	-0.0744
	(0.0586)	(0.0936)	(0.0734)
difficult	0.00137	0.0291	-0.0173
	(0.0612)	(0.0974)	(0.0780)
complete	-0.0643	0.0395	-0.161
	(0.0957)	(0.146)	(0.123)
impatience	0.0332	0.0539	0.00477
	(0.0342)	(0.0557)	(0.0436)
Observations	241	107	134

Probit regression with average marginal effects and robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A-2: Productivity and influencing factors by age groups

Dep. var.: productivity	(1)	(2)	(3)
	all	3/4 year olds	5/6 year olds
female	0.994**	0.729	1.308**
	(0.488)	(0.896)	(0.537)
ability	0.299***	0.0604	0.221***
	(0.0750)	(0.163)	(0.0730)
difficult	0.995**	0.140	1.228**
	(0.480)	(0.737)	(0.569)
complete	1.202**	2.102**	0.471
	(0.585)	(0.878)	(0.748)
impatience	-0.358	-0.571	-0.173
	(0.251)	(0.416)	(0.324)
Constant	5.817***	6.573***	7.655***
	(0.868)	(1.553)	(0.987)
Observations	426	178	248

OLS regression with robust standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Productivity defined as number of collected beads over time spent on task. We miss three observations because due to technical difficulties time recording did not work for those three children.

Table A-3: Diligence controlling for treatment

Dep. var.: number of beads	(1)	(2)	(3)
	all	3/4 year olds	5/6 year olds
female	14.06**	8.548	20.05**
	(5.504)	(5.184)	(8.666)
ability	5.360***	3.152***	4.615***
	(0.889)	(0.858)	(1.320)
difficult	20.95***	7.520	25.94***
	(5.099)	(5.312)	(7.731)
complete	16.86***	8.053	24.70***
	(5.500)	(5.866)	(8.787)
impatience	-1.418	-7.033**	1.784
	(3.234)	(2.937)	(5.254)
ex. tomorrow	-12.32	3.804	-16.17
	(8.774)	(9.259)	(12.38)
en. today	-9.067	10.77	-16.47
	(8.404)	(8.905)	(11.48)
en. tomorrow	-24.69***	-1.977	-31.29**
	(8.531)	(8.598)	(13.37)
Constant	-0.624	11.92	3.666
	(11.37)	(11.22)	(18.50)
Observations	429	179	250

OLS regression with robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A-4: Diligence and parental data

Dep. var.: number of beads	(1)	(2)	(3)	(4)
age	13.78***	13.62***	13.09***	17.25***
	(4.456)	(4.530)	(4.455)	(5.258)
female	12.37*	15.88**	16.80**	18.86**
	(7.391)	(7.265)	(7.311)	(8.365)
ability	5.223***	5.236***	5.452***	5.417***
	(1.159)	(1.147)	(1.147)	(1.445)
difficult	17.48**	16.88**	15.54**	20.90**
	(6.753)	(6.799)	(6.987)	(8.404)
complete	16.86**	14.62*	13.55	12.96
	(8.025)	(8.493)	(8.886)	(12.03)
impatience	-1.424	0.431	0.149	2.943
	(4.406)	(4.453)	(4.409)	(5.005)
siblings		-3.581	-3.702	-8.611
		(4.714)	(4.816)	(5.935)
single		-10.61	-10.64	-25.66
		(12.12)	(12.27)	(17.25)
stay-at-home parent		-3.051	1.629	-2.193
		(15.69)	(15.85)	(17.50)
working part-time		5.167	8.298	0.157
		(13.70)	(13.78)	(17.35)
education		4.485	6.651*	9.510**
		(3.341)	(3.660)	(4.246)
parent's procrastination			0.835	0.838
			(0.620)	(0.723)
parent's grit			0.712	0.446
			(0.771)	(0.896)
parent's patience			-1.814*	-2.030*
			(0.953)	(1.166)
income				-7.712
				(6.016)
Constant	-70.30***	-86.15***	-93.64***	-86.02**
	(20.58)	(27.38)	(29.06)	(34.74)
Observations	206	206	206	163

OLS regression with robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

INSTRUCTIONS

Instructions translated from German. German instructions available upon request.

Hello! My name is .... *[name of experimenter]*. What's your name? Would you like to play a game with us?

[Experimenter and child walk to experimenter room, showing the presents at the entrance and sit down in dedicated space. Two bags with the child's name are prepared.]
I will explain how the game works, so listen closely. In this game you can collect tokens which you can exchange for some presents afterwards. After I explain the game you will repeat it back to me, alright? And since you've been paying close attention I will already give you one token, which you can exchange for one present at the end of the game. Let's put the token over here and start the game. *[token placed in "today" bag]*

TIME PREFERENCES
In this game you can collect tokens for today and for tomorrow, which you can exchange for presents either today or tomorrow. Look, here I have pink and blue bowls. The tokens in the pink bowl can be exchanged into presents TODAY, the tokens in the blue bowl can be exchanged into presents TOMORROW.
Do you know what tomorrow means? Tomorrow means that you will sleep for one night to receive the tokens from the blue bowl and then you can exchange them into presents. I'm sure your kindergarten teacher has told you that we will be back tomorrow. We will be back tomorrow morning and also bring the same presents with us.

Now, the tokens from the pink bowl can be exchanged for presents today and the tokens from the blue bowl can be exchanged for presents tomorrow. Got it?

Here we have three possibilities and you can pick one: *[show three sets of cardboards with bowls]*

- Option 1: If you choose option 1, there are 2 tokens in the pink bowl and none in the blue bowl. That means you will receive 2 presents today and no presents tomorrow.
- Option 2: If you choose option 2, there is 1 token in the pink bowl and 2 tokens in the blue bowl. That means you will receive 1 present today and 2 presents tomorrow.
- Option 3: If you choose option 3, there are no tokens in the pink bowl and 4 tokens in the blue bowl. That means you will receive no presents today and 4 presents tomorrow.

You may now choose one of these three options. But first please answer some questions:

- How many options can you choose? *[child: one]*
- What happens if you choose option 1? How many tokens will you receive today and how many tokens will you receive tomorrow? *[let child count tokens in each bowl]*
- What happens if you choose option 2? How many tokens will you receive today and how many tokens will you receive tomorrow? *[let child count tokens in each bowl]*
- What happens if you choose option 3? How many tokens will you receive today and how many tokens will you receive tomorrow? *[let child count tokens in each bowl]*

[repeat instructions if child cannot answer correctly - repeat up to two additional times] Control questions asked in randomized order for each experimenter.

Well done! Now please choose one option.
Great, you have chosen option ... . That means you will receive X tokens

presents today and X presents tomorrow. Can you tell me why you chose this option? *[note down answer]* Let me put the tokens you will receive tomorrow into this bag. See, I wrote your name on it so I can save it for tomorrow. Tomorrow you will come back and exchange these tokens for presents. *[put tokens into today and tomorrow bags]*

PUZZLE TASK

[prepare puzzles: 3- and 4-year-olds - 6 and 12 pieces; 5- and 6-year-olds 12 and 24 pieces] You've done a great job so far! Would you like to play another game? Look, here I have two puzzles. Both puzzles have exactly the same picture. But one puzzle is more difficult and one puzzle is easier to do. This puzzle is difficult *[show puzzle with more pieces]*, because the puzzle pieces are smaller. This puzzle is easier because the puzzle pieces are bigger *[show one larger and one smaller puzzle piece for comparison]*. Do you see the difference? If you manage to do the difficult puzzle, you will receive two additional tokens. If you manage to do the easier puzzle, you will receive one additional token. You will have until lunch time to do the puzzle.

Now, before you decide which puzzle to keep, I have some questions for you.

- Which puzzle is more difficult?
- How many presents will you receive if you complete the easy puzzle? How many presents will you receive if you complete the difficult puzzle?
- Until when do you have to do the puzzle to still receive your presents for this task?

[repeat instructions if child cannot answer correctly - repeat up to two additional times] Great, now you can decide which puzzle you would like. You will be able to keep that puzzle and take it home with you afterwards.

REAL EFFORT TASK

Well done! Now I have one last task for you. You can collect some additional

tokens in this task, which you can again exchange for presents afterwards. Should I explain how the task works?
Look, here I have another bowl. There are many colorful beads in this bowl. The task is to collect ONLY YELLOW beads. The more yellow beads you collect, the more presents you will get. You can collect the yellow beads for as long as you want to. If you want to stop collecting beads, just say "stop" and we will count how many beads you have collected. Depending on the number of yellow beads you collect you will receive a few or many presents. If you, for example, collect this many yellow beads - these are 20 beads - *[show bowl with 20 yellow beads]* you will receive one present.

Do you understand how the task works?

- What do you have to do? *[child: collect yellow beads]*
- If you collect many yellow beads, will you receive more presents or fewer presents? *[child: more]*
- When will you stop the task? *[child: when I want to]*

Great! So let's do a trial round to see how it works. You can start picking yellow beads from the bowl. Ready? Go! *[stop child after 30 seconds, note number of beads]*
Great job!

[Read assigned treatment only!]
T1a (today): Since you've done such a good job, you can sit down over there right now and start collecting yellow beads from the bowl. You can collect yellow beads for as long as you want to. If you want to stop collecting yellow beads signal to .... over there and she/he will count your beads and exchange your tokens into presents.
T1b (tomorrow): Since you've done such a great job, you can do this task tomorrow. Tomorrow we will come back to this room and then you can sit in one of those spots over there and collect yellow beads. You will be able to collect yellow beads for as long as you want to. If you want to

stop collecting yellow beads you will signal to ... over there and she/he will count your beads and exchange your tokens into presents tomorrow. *[repeat explanation of RET when child comes back the next day]*
T2 (endogenous): Since you've done such a great job, you can decide when you want to do this task. You can either do the task right now or you can do it tomorrow - remember we will be back tomorrow with the same presents. For this task you will sit in one those spots over there and collect as many yellow beads as you want to. You can collect yellow beads for as long as you want to. If you want to stop collecting yellow beads signal to ... over there and she/he will count your beads and exchange your tokens into presents.

When would you like to do this task? Now or tomorrow? *[note down decision and read the according paragraph below]*

- You have decided to the task now. That means you can sit down over there and start collecting yellow beads. If you want to stop collecting yellow beads just signal to ... over there. Then you can exchange your tokens for presents afterwards.

- You have decided to the task tomorrow. That means we will pick you up tomorrow and take you to this room again. We're almost done for today! Now you get to exchange your tokens for presents and then I'll take you back to your class. Thank you for doing such a great job today!

RET
[different experimenter is responsible for supervising RET; child signals to stop the task, note down time and weigh beads on scale, convert into tokens]
You've done a great job! Can you tell me how much fun it was to collect beads? Look, here I have five smiley faces. This face is sad because it did not like the task at all. The face next to it didn't think it was that much fun either but not as bad as the first one. The face in the middle thought it was kind of ok. This face is smiling because it liked the task. And this face

here is laughing a lot because it really liked the task. How much did you like the task? Can you show me the face that fits you the most?
Alright that's it for today! Now let's exchange your tokens for presents! Then I will take you back to your class (if applicable: and I'll see you again tomorrow). Thank you for doing such a great job today! *[exchange tokens into presents with child; put chosen presents into bags, add parent questionnaire and seal them; take child back to class and leave bag at child's spot in wardrobe]*

PARENTAL QUESTIONNAIRE

Parental questionnaire translated from German. German version available upon request.

Dear parents, we kindly ask you to fill in this anonymous questionnaire. All answers are voluntary. Thank you for your collaboration!

Demographic information about your person

Gender:
O female
O male

Age: .... years

Occupation:
O full-time job
O part-time job
O momentarily unemployed

Highest educational degree:
O Mandatory schooling
O Vocational training
O High-School
O University (Bachelor/Master Degree)
O University (PhD)

Language, primarily spoken at your home (please indicate only one):
O German
O Turkish
O Serbian/Croatian
O other: .....

Number and age of children:
O 1 child, age: .... years
O 2 children, age: .... years & .... years
O 3 children, age: .... years & .... years & .... years
O 4 children or more, age: .... years & .... years & .... years & .... years & .... years

I'm raising my children:
O alone
O with my partner

Age of my partner: .... years

Partner's occupation:
O full-time job
O part-time job
O momentarily unemployed

Partner's highest educational degree:
O Mandatory schooling
O Vocational training
O High-School
O University (Bachelor/Master Degree)
O University (PhD)

Net monthly income of our family (voluntary disclosure):
O below €1,500
O €1,500- €2,500
O €2,500- €3,500
O above €3,500

We live in a:
O rental flat/house

O own flat/house

Self assessment

Please indicate for each of the following statements how well it describes you.

1= not at all to 5= to a very high extent

I am a hard worker.	1	2	3	4	5
I get enthusiastic about ideas for a short time but later lose interest.	1	2	3	4	5
I am self-disciplined.	1	2	3	4	5
I can cope with setbacks.	1	2	3	4	5
New projects sometimes distract me from previous ones.	1	2	3	4	5
I am good at resisting temptation.	1	2	3	4	5
I finish whatever I begin.	1	2	3	4	5
I have difficulty maintaining focus on projects or tasks that take more than a few months to complete.	1	2	3	4	5
I have trouble concentrating.	1	2	3	4	5

I often find myself performing tasks that I had intended to do days before.	1	2	3	4	5
I often miss concerts, sporting events, or the like because I don't get around to buying tickets on time.	1	2	3	4	5
Even with jobs that require little else except sitting down and doing them, I find they seldom get done for days.	1	2	3	4	5
In preparing for a deadline, I often waste time by doing other things.	1	2	3	4	5
New projects sometimes distract me from previous ones.	1	2	3	4	5
I usually return an RSVP request very shortly after receiving the invitation.	1	2	3	4	5
I often finish a task sooner than necessary.	1	2	3	4	5
I usually accomplish all the things I plan to do in a day.	1	2	3	4	5
I have difficulty maintaining focus on projects or task that take more than a few months to complete.	1	2	3	4	5
I am continually saying "I'll do it tomorrow".	1	2	3	4	5
I usually take care of all the tasks I have to do before I settle down and relax for the evening.	1	2	3	4	5

Lottery

Among the parents who will return the filled in questionnaire we will randomly pick five parents. In case you get picked, the following part will determine your earnings.

In the following lines (1-11) you are asked to choose between:
- receiving €50 directly after the end of our study (end of June), or

- receiving an amount between €50 and €70 at the end of September (hence, three months after the end of our study).

Please choose in each of the following lines, which of the two options you prefer. One of these lines will ultimately be randomly drawn for payment.

Example: If line 6 will be randomly picked for payment and in this line you have chosen the amount in the right column (€60) you will get €60 in three months.

	End of June	**or**	**End of September**
1)	€50	or	€50
2)	€50	or	€52
3)	€50	or	€54
4)	€50	or	€56
5)	€50	or	€58
6)	€50	or	€60
7)	€50	or	€62
8)	€50	or	€64
9)	€50	or	€66
10)	€50	or	€68
11)	€50	or	€70

In the envelope - which contained your questionnaire - you will find a code. This code facilitates your payment. Please transfer your personal code into the following box. Please ensure that you transfer the code correctly.

Participation code: ... ...

The codes of the winning participants will be announced via the information board in the kindergarten. The winners will be notified when to pick up their payments (either right away or in three months). In case you are one of the winners, you are kindly asked to present the piece of paper with your participation code in kindergarten. Please present this piece of paper at the communicated point in time in order to get your payment. [19] You will receive your payment in a sealed envelope. Please store your participation code in a safe place so we can verify your code for the payment.
For your information: The participation code ensures full anonymity.

Assessment of your child

In the following we will report several statements. Please indicate for each statement how well it describes your child.

1= not at all to 5= to a very high extent

If I ask my child to perform an unpleasant task he/she tries to postpone the task for as long as possible.	1	2	3	4	5
My child likes to dawdle.	1	2	3	4	5
My child likes to spend a lot of time on a given task.	1	2	3	4	5
My child gets easily distracted.	1	2	3	4	5
Whatever my child begins, he/she wants to finish.	1	2	3	4	5
It's hard for my child if he/she has to wait.	1	2	3	4	5
My child likes playing challenging games.	1	2	3	4	5

[19] In case your child has left kindergarten at this point in time, please notify us and we will send your payments.

During our project in kindergarten, we asked children to choose between three alternatives:

1) Option 1: two presents today, nothing tomorrow.
2) Option 2: one present today, and two presents tomorrow.
3) Option 3: nothing today, and four presents tomorrow.

We promised children to come back the next day with presents, which were as nice as the ones they could get on the first day. Independently of what your child might have chosen, which option would you prefer for your child? The decision you make now has no impact on your child.

Which option would you choose for your child?

◯ Option 1: two presents today, and no present tomorrow

◯ Option 2: one present today, and two presents tomorrow

◯ Option 3: no present today, and four presents tomorrow

We also played the following game with the children. Now we are interested in your opinion.

The children were told to collect only yellow beads from a bowl of small, multi-colored beads. The children could autonomously decide for how long they wanted to work on this task and how many beads they collected. Moreover, children could decide if doing the task right away or postponing it to the next day.

Which option would you choose for your child?

◯ "What is done, is done." My child should do the task right away.

◯ "Better late than never." My child should take his/her time and conduct the task on the following day.

Thank you for your participation.

CHAPTER 3

CONDITIONAL LYING

Abstract

Dishonesty is a widespread concern within and across companies. A characteristic feature of companies is that employees who collaborate with each other can observe each others' behavior. Observing that others are dishonest might impact one's own decision whether to be dishonest too. This might be especially the case in group settings where employees bear the consequences of each others' dishonesty. In a laboratory experiment, we find that irrespective of whether others' lies affect one's own payoff, one third of all subjects are dishonest if others are too. Having only one dishonest group member makes the vast majority of these *conditional liars* switch from being honest to being dishonest. The size of a lie increases with the number of dishonest group members that one faces. Overall, we find that *conditional liars* tell smaller lies than *always liars* do.

This chapter is joint work with Thomas Lauer.

3.1 INTRODUCTION

In their 2018 report the Association of Certified Fraud Examiners (ACFE, 2018) estimates a total loss of more than USD 7 billion worldwide due to fraudulent employees' behavior across countries and industries. Not only large business scandals like Enron and Volkswagen contribute to this staggering number. The ACFE reports that the largest share (89%) of fraud cases comes from so called asset misappropriation including cases in which employees inflated their expenses, working hours, or efforts. It comes as no surprise that companies and government agencies take actions to counteract or prevent this kind of dishonesty, and codes of ethics seem to be the means of choice. While Kaptein (2004) found that only about 53% of the two hundred largest companies in the world had a business code in 2001, the ACTE in their 2018 report indicates 80% of companies to have such a code of conduct. The rationale for these codes of conduct is to create an environment of honesty and prevent dishonesty from spreading within the company. We show that this concern is justified, i.e., that people condition their dishonesty on the dishonesty of others and that dishonesty spreads quickly within groups.

From the literature on lying behavior in individual settings we know that there are at least two different types of individuals: those who never lie even if they have to forgo substantial advantages and those who lie if they benefit from it (Gneezy, 2005; Sutter, 2009; Gneezy et al., 2013; Fischbacher and Föllmi-Heusi, 2013; Abeler et al., 2018). In a group context (e.g., inside a working team), however, there might be a third type following a different strategy. From the literature on cooperation in social dilemma situations we know that a substantial fraction of people condition their choices on the behavior of other people. They cooperate as long as they expect others to cooperate too and stop doing so when they learn that others do not cooperate (Fischbacher et al., 2001; Fehr and Fischbacher, 2003; Kocher et al., 2008; Fischbacher and Gächter, 2010). In a group setting, where people decide whether to be dishonest or not, people might follow a similar

strategy and make the decision to be dishonest or not depend on what they expect or observe others to do.

In the course of everyday working practice, employees face ample temptations to behave dishonestly. Observing other employees' fraudulent acts might impact one's own decision whether to be dishonest too. Think of a co-worker over-reporting his travel expenses, a colleague unjustifiably claiming to stay at home because of illness, or a team member overstating his performance in order to increase his chances to get promoted. An individual's tendency to respond to dishonesty with dishonesty might depend on the effect the other person's lie has on one. While in the first case the over-reporting does not directly harm the other employees, in the second case their work-load might increase to compensate for the colleague's absence. Ultimately, in the third case not only the other group members, but even the overall firm's efficiency might suffer because the dishonesty complicates the evaluation of employees' performances, potentially leading to the promotion of a wrong candidate. When deciding how to respond to other group members' dishonesty, individuals might take these varying effects lies have on one another into account.

This paper aims to answer the question on how other group members' dishonesty affects one's own willingness to be dishonest. We study this question in different group settings where dishonesty has increasingly harmful effects on one another. Fur this purpose, we run a laboratory experiment where we match participants into groups of four and ask them to individually work on a task that produces the group's joint payoff. Group members are then called to divide this payoff among each other based on individual reports regarding the own performance. When dividing the group's joint payoff each member can lie (i.e., inflate the own performance) to increase the own payoff. Here we systematically vary the number of dishonest group members a person approaches in order to study how this affects the individual willingness to be dishonest. In a between-subjects design, we, additionally, introduce three different group settings where we vary the mutual monetary effects that lying has on group members. In treatment WHITE, lying (besides benefiting oneself) has no monetary impact on the

other group members. Lying harms the other members in treatment GRAY, and additionally hurts the overall group payoff in treatment BLACK. We want to learn two things. Firstly, whether lying behavior in a group setting happens as a conditional response to the dishonesty of others and if so, which group setting fosters its emergence. Secondly, we want to gain a deeper understanding about the heterogeneity in the different types of liars.

Our results show that roughly one third of all participants condition their dishonesty on the dishonesty of their group members. Independent of the group setting and, thus, the mutual monetary effects that lying has on one another, we find a similar percentage of so called *conditional liars*. In this sense, none of our group settings particularly promotes or impedes the emergence of conditional lying. For the vast majority of *conditional liars*, having only one dishonest group member is enough to make them switch from honesty to dishonesty. While *conditional liars* distinguish themselves from *always liars* in their tendency to make their own dishonesty decision depend on others, we also see that *conditional liars* inflate their performance to a smaller degree than *always liars*. Both types of liars increase the size of their lies with the number of dishonest group members.

Our experimental design allows us to study the mechanisms that contribute to the emergence of conditional lying in a group setting. On the one hand, conditional lying might be grounded on the individual's perception that other people's rule violations make this kind of behavior become more acceptable (Wilson and Kelling, 1982; Cialdini et al., 1990; Keizer et al., 2008). Related to this, Kahan (1997) introduces the example of someone observing others committing a crime. If many commit a crime, the individual observing it will assume that the risk of being caught is low and little reputation costs are connected to committing a crime. Most importantly, the individual will also consider committing crime a more acceptable behavior. In this regard, Kahan (1997) proposes the concept of "social influence" as "individuals' perceptions of each others' values, beliefs, and behavior affecting their own conduct" (p. 350). We consider social influence to be one important driving factor for the emergence of conditional lying in a group setting.

Besides social influence, negative reciprocity might contribute to the emergence of conditional lying in a group setting if group members' lies have monetary drawbacks for one another. Negative reciprocity refers to people's general tendency to reciprocate monetary harms that others caused to them with (the same) unkind actions (Güth et al., 1982; Berg et al., 1995; Fehr and Gächter, 1998). If group members' lies have negative consequences on the remaining members, these remaining members might actually want to retaliate against the liar by telling lies themselves. Since in our experiment the mutual effects of lying differ across treatments, we can determine the contribution of social influence and negative reciprocity for the emergence of conditional lying. We find that social influence plays a much more pronounced role in the emergence of conditional lying than negative reciprocity.[1]

Overall, our paper contributes to research on the prevalence of rule violations and violators on individual lying behavior. Gächter and Schulz (2016) provide empirical evidence for the society as a whole shaping one's own dishonesty decision. They find that the amount of political fraud, tax evasion, and corruption in a society increases individuals' propensity to dishonestly report the outcome of a die roll in order to increase individual earnings. Complementing this, Gino et al. (2009) show that even a single individual impacts others' dishonesty decision if they share the same social identity. When students are asked to individually fulfill an exercise, having a student from their university in the room who obviously cheats on the experimenter increases the willingness of the remaining peers to cheat on the experimenter too. Likewise, Diekmann et al. (2015) find that the mere awareness that others might cheat is sufficient to trigger further dishonesty. We add to this literature by examining conditional lying as a mutual phenomenon in a group setting rather than focusing on unilateral behavior directed towards the experimenter. Additionally, we can differentiate between those who lie as a consequence of other liars and those who would have lied anyways. In

[1]For contributions to a public good Bardsley and Sausgruber (2005) show that mainly negative reciprocity (2/3) and to a lesser extent (1/3) social influence are responsible for the emergence of conditional free-riding.

this sense, our paper addresses the heterogeneity in types of liars and sheds light on the type specific peculiarities.

The remainder of this paper is organized as follows. Section 3.2 describes the experimental design and procedure. Section 3.3 presents the results of the study, concentrating first on the existence and emergence of conditional lying and second on the individual heterogeneity of different types of liars. Section 3.4 concludes.

3.2 EXPERIMENT

3.2.1 DESIGN

In a laboratory setting we randomly match subjects into groups of four. The experiment is then structured in two phases. In Phase I (Production), subjects are asked to work on a real effort task (slider task, Gill and Prowse, 2012). The subjects have 4 minutes to set as many sliders as possible on position 50. The more sliders the group members jointly solve, the bigger the group's joint payoff. A 2-minute non-incentivized trial phase proceeds this production phase. In the production phase, we purposefully make subjects work for their payoffs in order to create a feeling of entitlement (Gächter and Riedl, 2005; Kajackaite, 2018). After having completed the production phase, group members get individual feedback about their own number of correctly solved sliders but not the other players' performances.

In the Phase II (Reporting), subjects are asked to individually and simultaneously report to each other how many sliders they themselves have solved. Note that the instructions clearly state that subjects are asked to report the actual number of correctly solved sliders. This reporting phase is structured in two parts: Part A the unconditional reports and Part B the conditional reports. Subjects learn about the nature of these parts gradually. In Part A, subjects are asked to provide their unconditional reports, where all group members simultaneously report how many sliders they have solved. The more solved sliders a group member reports, the bigger her share of the group's joint payoff. For three out of four group members, this unconditional report is payoff relevant. For the forth member,

the conditional report provided in Part B of the reporting phase determines the own payoff. In Part B, group members are asked to report the number of sliders they solved conditional on their group members reporting a different than actual number of solved sliders. For this purpose, each subject faces a list with four different scenarios. Each group member is then asked to report the number of solved sliders for each of the possible scenarios: (i) no other group member reported a different than the actual number of solved sliders, (ii) one group member reported a different than the actual number of solved sliders, (iii) two group members reported a different than the actual number of solved sliders, and (iv) three group members reported a different than the actual number of solved sliders. Group members can, thus, either report the same number of solved sliders independently of the behavior of the others or adjust their own reported number of solved sliders based on the number of dishonest group members that they face (see appendix for more details).

We run three different treatments where we vary the monetary effect a lie has on the other group members. In treatment WHITE, a lie has a positive effect on one's own payoff but does not hurt the others because a player's own reported performance does not enter the payoff function of the other members. In this setting, conditional lying might arise due to observing others lying and not from reciprocating the harm another member's lie has caused one.[2] The specific payoff function in WHITE is given in equation (3.2.1). A player's payoff is determined by dividing one's own reported performance, R_i, by the sum of the other players' actual performances, $\sum_{j \neq i} P_j$, plus one's own reported performance. This ratio is then multiplied by the group's joint payoff, $\sum P$, which is based on the actual performance of the group.

$$\text{payoff}_{\text{i, WHITE}} = \frac{R_i}{R_i + \sum_{j \neq i} P_j} * \sum P \tag{3.2.1}$$

[2]While lying has no monetary effect on other group members, it changes the rank order of payoffs between the liar and the other members. In this sense, the preference for keeping the rank order could also explain the emergence of conditional lying in WHITE and the latter treatments. To reduce this concern, we do not inform subjects about their ranks.

In contrast, in treatment GRAY, a person's lie affects their own payoff positively and in addition harms the other members. This happens because inflating one's own performance decreases the other members' payoffs. Specifically, an inflated performance decreases the denominator in the payoff function of the other members as shown in equation (3.2.2). This means that in addition to social influence, reciprocity might explain the emergence of conditional lying because another member's lie harms oneself.[3] Individual payoffs are determined by dividing one's own reported performance by the sum of all reported performances ($R_i + \sum_{j \neq i} R_j$). This fraction is multiplied by the group's joint payoff, again based on the actual performances of group members.

$$\text{payoff}_{\text{i, GRAY}} = \frac{R_i}{R_i + \sum_{j \neq i} R_j} * \sum P \tag{3.2.2}$$

Lastly, in treatment BLACK, lying does not only transfer (GRAY) but also destroys part of the group's joint payoff. Specifically, the payoff function in BLACK is the same as in GRAY with the exception that we introduce deduction points for each wrongly reported performance level. As shown in equation (3.2.3), for each group member who wrongly reported her performance (where d is the number of dishonest members), we deduct 10 points from the group's joint payoff.[4] Again, conditional lying might result from two potential sources: social influence and reciprocity, where reciprocity might play an even more pronounced role than in GRAY due to the increased harm a member's lie causes to the others.

[3]We acknowledge that by changing the monetary effect the lie has on other group members by nature we also change the overall efficiency, i.e., individual lies no longer increase the joint payoff. This change in efficiency does not seem to play a fundamental role if we compare unconditional reports in WHITE and GRAY (see section 3.3.1).

[4]The deduction points are the same regardless of the size of the lie in order to express than lying - independently of its magnitude - hurts the overall group's efficiency.

$$\text{payoff}_{\text{i, BLACK}} = \frac{R_i}{R_i + \sum_{j \neq i} R_j} * (\sum P - d * 10) \tag{3.2.3}$$

At the end of the experiment, we ask participants to fill in a questionnaire. First, we ask how strongly people felt an inner conflict when reporting the number of solved sliders and how much they are morally and financially satisfied with their performance on a scale ranging from 1 to 7, where 7 indicates the strongest inner conflict/highest satisfaction. Next, we ask subjects to complete a psychological scale which elicits subjects' tendency to deceive or manipulate others for personal gain (Mach IV scale, Christie and Geis, 1970). This 20-statement scale captures the Machiavellianism of a person, namely her tendency to deceive and manipulate others for her own gain. Here, we want to study whether those categorized as *always liars* in our experiment are more likely to score above median on the Mach scale. In this sense, the Mach scale serves as a robustness check for our classification in different types of liars.

3.2.2 PROCEDURES

In total, 212 subjects participated in our experiment, 72 subjects were assigned to the WHITE, 68 subjects to the GRAY[5], and 72 subjects to the BLACK treatment. The experiment was conducted at the University of Cologne using zTree (Fischbacher, 2007). Participants were recruited via ORSEE (Greiner, 2015). To ensure comprehension - and especially to ensure that subjects understood the consequences that their lies had on others in the different group settings - we asked them to answer control questions during the course of the experiment.

3.3 RESULTS

This section will focus first on the performance of subjects in order to show that all subjects in our experiment had the possibility to lie, where lying

[5]Observation numbers are smaller in GRAY due to no-show ups.

means reporting a number of solved sliders that deviates from one's actual number of solved sliders. On the one hand, we will consider the binary decision to lie or not (extensive margin of lying) and on the other hand, the size of a lie conditional on lying (intensive margin of lying). When considering the intensive margin of lying, we calculate by how much a liar exaggerated her performance in relation to her actual performance. An intensive margin of lying equals to 0.1 means that a dishonest person increased her performance by 10%. We will then briefly give an overview of the unconditional reports provided in the different treatments, before we come to the main part of our study - the conditional reports. Where we will concentrate first on aggregate conditional reports and ultimately divide subjects into different types of liars.

3.3.1 PERFORMANCE AND UNCONDITIONAL REPORTS

Table 3.1 provides summary statistics of subjects' average performance and their extensive and intensive margin of lying when asked to provide unconditional reports. On average, participants managed to solve 32 sliders out of potential 96 sliders. This number does not statistically differ in the different group settings ($p>.378$, Mann-Whitney-U-test, pairwise comparisons). Effort levels of (unconditionally) honest or dishonest subjects also do not differ ($p=.442$, Mann-Whitney-U-test). The best performer managed to solve 53 sliders while the worst performer solved 10 sliders, showing that all subjects had the possibility to overstate their performance to a sizable extent.

A similar percentage of subjects lie in their unconditional reports in WHITE (39%) and in GRAY (37%) ($p=.862$, Fisher's exact test), while less lying occurs in BLACK (25%) (WHITE vs. BLACK, $p=.053$; GRAY vs. BLACK, $p=.093$; Fisher's exact tests). However, when considering the intensive margin of lying, we see that bigger lies are told in BLACK than in WHITE ($p=.021$, Mann-Whitney-U-test). Remember, in BLACK overstating by a little is financially unattractive because the deduction points for dishonesty are the same regardless of the size of the lie. The increased

Table 3.1: Average performance, extensive, and intensive margin of lying by treatment

Treatment	**Performance**	**Extensive margin**	**Intensive margin**
WHITE	32	39%	0.69
GRAY	32	37%	1.09
BLACK	32	25%	1.25

Notes: The extensive margin of lying captures the percentage of subjects being dishonest. The intensive margin of lying illustrates the size of a lie conditional on lying and in relation to an individual's performance.

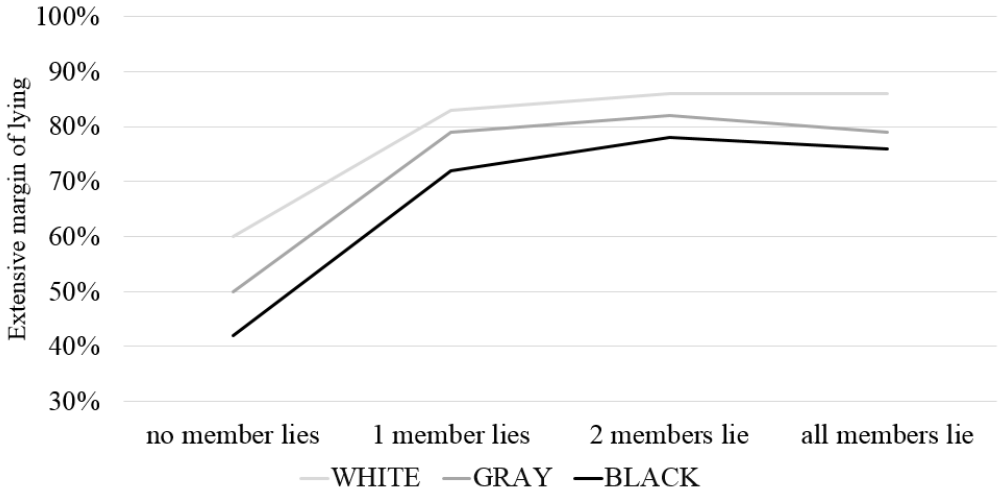
intensive margin of lying in BLACK might reveal that subjects have taken this design feature into account. In summary, we do not find unconditional lying to be considerably affected by the monetary harms that lying causes among group members. Only when the group's joint payoff is also affected does lying behavior occur significantly less often - and then the told lies are bigger.

3.3.2 CONDITIONAL REPORTS

In this section we focus on the conditional reports of subjects. These conditional reports illustrate if subjects let their lying decision depend on the behavior of their group members. In particular, we learn if subjects - given the behavior of their group members - are *conditional liars*, i.e., if they lie in response to the lies of others or whether they in principle either never or always lie. Figure 3.1 illustrates subjects' extensive margin of lying conditional on the number of dishonest group members that they face. The different lines indicate the three different treatments. Important to note is that all lines exhibit a positive slope, indicating that the fraction of dishonest subjects is higher if other members in the group are dishonest too. This provides first evidence that lying happens conditionally.[6] Since all

[6]The fraction of dishonest subjects is lower when all members lied than when 2 members lied because 10 subjects in our sample either switch from dishonesty to honesty when the number of dishonest group members increases or follow some other particular

Figure 3.1: Extensive margin of lying by treatment



treatments have similar slopes, this also indicates that individuals respond similarly to other members' dishonesty in the different group settings. This means that irrespective of the monetary consequences of a lie, conditional lying evolves similarly in the different group settings. The extensive margin of lying is statistically significantly different only in WHITE and BLACK when no member lied ($p=.045$, Fisher-exact test), while all other pairwise treatment comparisons are not statistically significant ($p>.160$, Fisher-exact tests).

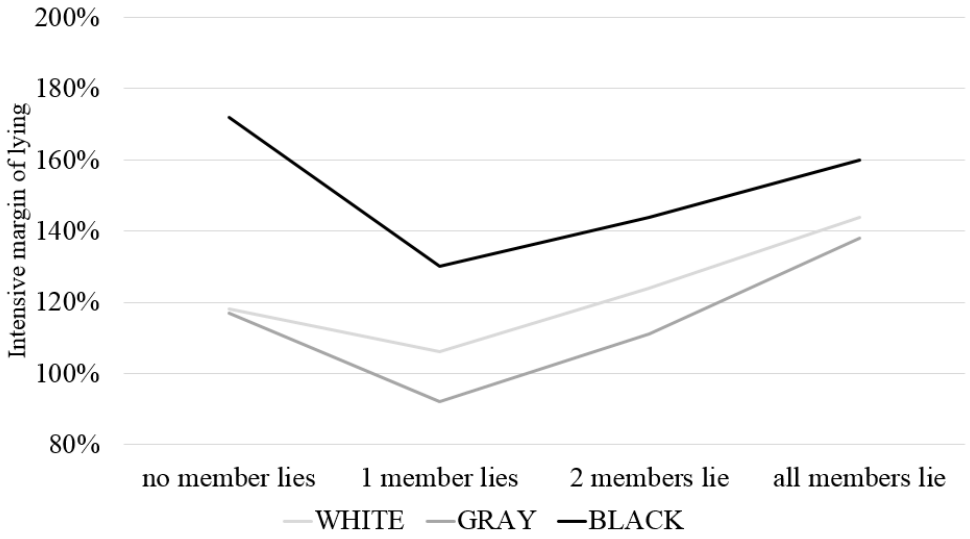
Figure 3.1 illustrates another interesting finding. In particular, we observe the extensive margin of lying to be higher even in the case of "no member lies" compared to the unconditional extensive margin of lying (see section 3.3.1). This difference cannot be driven by subjects who did not anticipate that they could lie when unconditionally reporting their number of solved sliders. All subjects had to correctly pass a series of control questions before reporting their number of solved sliders which one could only do when understanding that there exists a difference between actual and reported performance and

patterns (for instance, only lie if one or two members lie but not in the other conditions). We will not analyze the behavior of these subjects in more detail. If we exclude these subjects from our sample, the fraction of dishonest subjects if all other members lied is equal to 88% (WHITE), 83% (GRAY), 79% (BLACK).

that the reported not the actual performance was decisive for one's own payoff. However, what changes from unconditional to conditional reports is the salience of other group members' behavior. With unconditional reports the behavior of other members is not mentioned, while with conditional reports subjects are explicitly asked to condition their own behavior on the behavior of their group members. This means that the focus moves away from individual behavior (with unconditional reports) to the behavior of the group members (with conditional reports). If social influence plays a role in the decision whether to lie, it probably plays a more pronounced role with conditional reports when the salience rests on the behavior of the other group members. In line with this conjecture, we interpret the higher extensive margin of lying in conditional compared to unconditional reports as first evidence that social influence potentially matters in the decision to lie.

The intensive margin of lying is captured in Figure 3.2, which displays the intensive margin of lying conditional on the number of dishonest group members. As above, different lines indicate the three different treatments. The U-shaped form of the intensive margin of lying illustrates that the size of a lie increases when one, two, or all members lie, but it is bigger when no member lies compared to a situation where one member lies. While, at first glance this might seem surprising, this pattern will become clearer in section 3.3.3. Section 3.3.3 addresses individual heterogeneity of different types of liars. It shows that individuals who lie when no other member is dishonest (and therefore drive up the intensive margin of lying in the "no other member lies condition" in Figure 3.2) overstate their performance to a higher degree compared to individuals who start lying only when others are dishonest too. In this sense, Figure 3.2 anticipates that different types of liars overstate their performance to a different degree. Again, we see similar patterns in all three treatments indicating that the (conditional) intensive margin of lying develops similarly regardless of the group setting. In line with our observation of unconditional treatment comparisons, we find that the intensive margin of lying is greater in BLACK than in WHITE ($p<.079$, Mann-Whitney-U-test

Figure 3.2: Intensive margin of lying by treatment

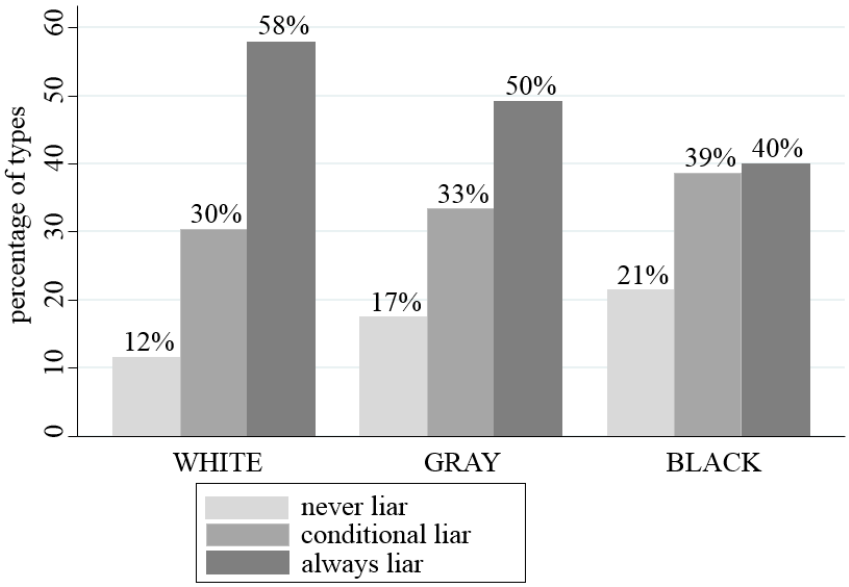


for pairwise comparisons of different conditions) and in GRAY ($p<.005$, Mann-Whitney-U-test for pairwise comparisons of different conditions).

For a better understanding of the heterogeneity of different types of liars, we classify subjects into three types: (i) subjects who never lie, (ii) subjects who switch from being honest to being dishonest as soon as one or more members lie, and (iii) subjects who always lie. Figure 3.3 reports the distribution of the different types of liars by treatment. The smallest fraction of participants in all treatments never lies (approximately 1 out of 6 subjects can be classified as *never liars*). In contrast, the largest fraction (slightly less than half of the subjects) are *always liars*. The rest of the subjects - about a third of all participants - condition their dishonesty on the dishonesty of their group members.

Focusing on the treatment comparisons, we see that the number of *never liars* is smaller in WHITE (12%) than in GRAY (17%) and BLACK (21%) ($p=.056$; (one-sided) Jonckerheere-Terpsta test). The opposite is true for *always liars* meaning that in treatment WHITE 58% of the subjects always lie, while this percentage is smaller in treatment GRAY (50%) and equals 40% in BLACK ($p=.023$, (one-sided) Jonckerheere-Terpsta test). These trends show

Figure 3.3: Distribution of the different types of liars by treatment



that some subjects do take the effects of their lies into account when deciding whether to be dishonest. In particular, less people always lie when their lies harm the other group members or the joint group payoff and, in turn, more people never lie in presence of negative effects on other group members or overall efficiency. The fraction of *conditional liars*, however, remains relatively stable across treatments (p=.144, (one-sided) Jonckerheere-Terpsta test). Specifically, conditional lying is present in WHITE (30%) when group members' lies do not have mutual monetary effects on one another, and its magnitude does not significantly change in GRAY (33%) and BLACK (39%) when we introduce mutual monetary effects of lying. From this we conclude that reciprocity seems to play a minor role for the emergence of conditional lying, since conditional lying happens irrespective of whether others' lies affect one's own payoff or not. Consequently, we infer that it is social influence that predominantly contributes to its emergence.

Result 1. *Roughly one third of all subjects are conditional liars. The fraction of conditional liars is unaffected irrespective of whether others' lies affect one's own payoff. From this we conclude that reciprocity seems to play*

a minor role for the emergence of conditional lying and, consequently, infer that social influence predominantly contributes to its emergence.

3.3.3 INDIVIDUAL HETEROGENEITY OF THE TYPES OF LIARS

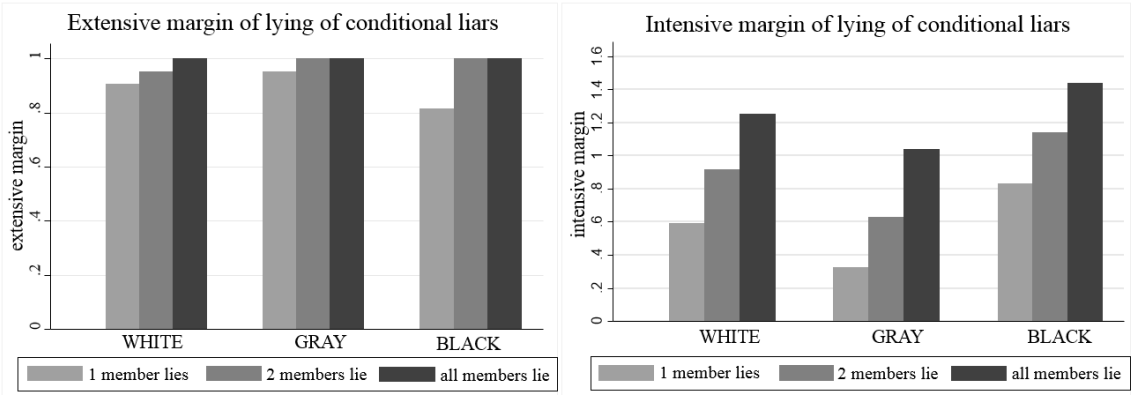
In this section we concentrate on the individual heterogeneity of the different types of liars. First, we focus on *conditional liars* only and examine their behavior when asked to unconditionally and conditionally report their number of solved sliders. Subsequently, we compare *conditional liars* to other types of liars and, ultimately, assess the validity of our classification in different types of liars by means of the Mach scale.

Conditional liars

Unconditional reports. Section 3.3.2 shows that reciprocity seems to play a minor role for the emergence of conditional lying, while it is mainly social influence which explains its emergence. Moreover, when contrasting unconditional and conditional reports, we expect the difference in the extensive margin of lying to be driven by the different role social influence plays with the two elicitation procedures. Putting these two findings together, *conditional liars'* unconditional reports might be affected by the weaker role social influence plays there. To test this expectation, we look at *conditional liars'* behavior when asked to unconditionally report their number of solved sliders. While we find a significant correlation between unconditional reports and average conditional reports for all subjects ($p<.001$, $\rho=.634$, Spearman correlation), we also find that only 10% of the *conditional liars* lie when unconditionally asked to report their number of solved sliders. Consequently and in line with our conjecture, 90% of *conditional liars* do not lie when social influence plays a presumably weaker role.

Conditional reports. Next, we focus on the number of rule violators in order to change *conditional liars'* individual willingness to lie. The left panel of Figure 3.4 illustrates the extensive margin of lying of *conditional liars* conditional on the number of dishonest group members. Again, we distinguish by treatment. In all three treatments we see similar patterns.

Figure 3.4: Extensive and intensive margin of lying of *conditional liars* in the different treatments



Notes: No bars for the case when no other member lies are provided because in this case *conditional liars* do not lie (their extensive margin of lying is equal to zero).

Namely, having only one dishonest group member drastically impacts *conditional liars'* extensive margin of lying. In fact, more than 80% of *conditional liars* behave dishonestly as soon as one other member lies. This percentage slightly but not statistically significantly increases for each additional liar ($p>.103$, Wilcoxon signed rank tests). We observe no difference in switching points from honesty to dishonesty across treatments ($p>.252$, Fisher's exact tests). This means that independently of the group setting, and thus the monetary effects lying has on one another, a single dishonest group member is enough to onset *conditional liars'* willingness to lie.

The right panel of Figure 3.4 illustrates the intensive margin of lying of *conditional liars* conditional on the number of dishonest group members by treatment. We again observe that the number of dishonest members impacts the intensive margin of lying. In all treatments *conditional liars* increase their intensive margin of lying the more group members lie ($p<.01$, Wilcoxon-signed rank tests).[7]

[7]The intensive margin of lying of *conditional liars* in treatment WHITE and GRAY does not systematically differ ($p>.098$, Mann-Whitney-U-tests). *Conditional liars*, however, tell smaller lies in GRAY than they do in BLACK ($p<.003$, Mann-Whitney-U-test) and in WHITE ($p<.086$, Mann-Whitney-U-test).

Result 2. *While in the vast majority of cases having one dishonest group member makes conditional liars switch from being honest to being dishonest, their intensive margin of lying gradually increases with the number of dishonest group members.*

Comparison of different types of liars

Next, we compare the different types of liars across each other. First, we focus on the comparison of *conditional liars* and *always liars* with regard to their intensive margin of lying (see Table 3.2). For all possible numbers of dishonest group members, *conditional liars* overstate their performance to a smaller degree than *always liars* do ($p<.003$, Mann-Whitney-U-tests). This is also true for the unconditional intensive margin of lying where we again observe that *conditional liars* tell smaller unconditional lies than always dishonest subjects do (0.338 vs. 1.025, $p=.060$, Mann-Whitney-U-test). In this sense, conditional and always dishonest subjects do not only distinguish themselves in the moment when they start lying but also in the size of their lies.

Table 3.2: Intensive margin of *conditional* and *always liars*

N of dishonest members	**conditional liar**	**always liar**
no member lies	0.000	1.311
1 member lies	0.591	1.416
2 members lie	0.917	1.534
all members lie	1.259	1.633

Notes: Based on the number of dishonest group members an individual faces, this table illustrates the intensive margin of lying of *conditional* and *always liars*.

As a next step, we focus on moral satisfaction and inner conflict of the different types of liars.[8] Table 3.3 reports the mean moral satisfaction and

[8]The question whether people felt an inner conflict when chosing how to behave in the experiment has previously been used in a study by Kocher et al. (2017).

inner conflict different types of liars report to have experienced in course of the experiment. Answers range on a scale from 1 to 7, where 1 stands for "low satisfaction/inner conflict" and 7 stands for "high satisfaction/inner conflict". Not surprisingly, *never liars* are morally satisfied the most and *always liars* are morally satisfied the least. While pairwise comparisons of moral satisfaction across different types of liars are always significant ($p<.006$, Mann-Whitney-U-tests), moral satisfaction does not differ by treatment condition, except for *always liars* who feel less morally satisfied in BLACK than in WHITE ($p=.002$, Mann-Whitney-U-test) or GRAY ($p=.003$, Mann-Whitney-U-test).

Table 3.3: Reported feeling of different types of liars

	never liar	**conditional liar**	**always liar**
moral satisfaction	6.06	5.19	4.43
inner conflict	2.29	2.97	2.71

Notes: Feelings range on a scale from 1 to 7, where 7 stands for the highest moral satisfaction/inner conflict.

Never liars are not only morally satisfied the most, they also report to have felt the smallest inner conflict when reporting their number of solved sliders. *Conditional liars*, instead, report the strongest inner conflict when reporting their performance. This means that *conditional liars*, although being sensitive to the influence of others, still seem to have inner troubles. This difference in experiencing an inner conflict is weakly significant if we compare *never liars* with *conditional liars* ($p=.056$, Mann-Whitney-U-test) and not different if we compare conditional and *always liars* ($p=.356$, Mann-Whitney-U-test).

Result 3. *Conditional liars tell smaller lies than always liars do. Conditional liars are more satisfied with their reporting decision than always liars but less satisfied than never liars. Conditional liars seem to experience a stronger inner conflict than never liars do.*

Robustness check of the validity of classification: Mach scale

As a robustness check for our classification in different types of liars we use the scores subjects obtained in the Mach scale. Since the Mach scale captures individuals' general willingness to deceive and manipulate others for one's own gain, we expect *always liars* to be high Mach types (score above median) and *never liars* to be low Mach types (score below median). This is in line with what we observe. *Always liars* are more likely to score above median (p=.027, Fisher exact test) and *never liars* are more likely to score below median in the Mach scale (p=.014, Fisher exact test). For *conditional liars* the categorization in high or low Mach types is more difficult. Their behavior depends on the behavior of others. In this light, we also do not find any evidence that *conditional liars* are more likely to be either low or high Mach types (p=.762, Fisher exact test), reinforcing the validity of our classification.

3.4 CONCLUSION

Dishonest behavior of employees costs companies billions of dollars every year. Understanding how dishonesty spreads in companies and more broadly in group settings is, therefore, of great importance to firms. In a laboratory experiment, we investigate to what extent the dishonesty of other group members fosters further dishonest behavior and whether the organizational setting affects the emergence of conditional lying.

We find that for a considerable amount of people - namely for a third - other people's dishonesty is the key factor for one's own decision whether to be dishonest too. Most importantly, the fraction of people conditioning the own dishonesty on other members' dishonesty does not change in different group settings. For *conditional liars* in the vast majority of the cases the presence of a single dishonest group member is enough to onset their propensity to lie. *Conditional* and *always liars* tell bigger lies, the more dishonest group members they face.

We consider these findings to offer important insights for companies. First, with regard to our finding that *conditional liars* care substantially

about the dishonesty of their group members while they have relatively little concern about the monetary effect of a lie, we attest the particular importance of clear codes of conduct. These codes of conduct, as emphasized in the introduction, may help in establishing honesty as the prevailing behavior inside a company and, therefore, may be able to prevent dishonesty from spreading around. Second and related to our finding that the presence of a single dishonest group member onsets *conditional liars'* propensity to lie, firms should put great care in deciding whom to hire. In course of the recruiting process, managers and human resource professionals may not only select employees based on their education and working experience but also based on their attitude towards honesty (Kerschbamer et al., 2017).[9] Of course, identifying those candidates for which honesty is important is challenging. However, even simple tasks like letting candidates fill in the Mach scale might suggest how important honesty is for a job applicant. Ultimately, our results show that other members' dishonesty affects not only *conditional liars* but also *always liars* by increasing their intensive margin of lying. Both types of liars adjust the size of their lies, the more dishonest group members they face. This, again, confirms the potentially devastating effects dishonesty causes in group settings and once again underlines the importance of strict guidelines and careful screening procedures. Overall, our results shed light on the presence of conditional lying, the heterogeneity of liars types and can be used as first experimental evidence for the study of dishonesty in group settings.

[9] Also Kerschbamer et al. (2017) emphasize the importance of careful screening procedures in the presence of credence good markets.

3.A APPENDIX

INSTRUCTIONS

Instructions translated from German. German instructions available upon request.

Welcome! You are about to participate in an experiment funded by the University of Cologne. Please switch off your mobile phone and remain quiet. It is strictly forbidden to talk to the other participants during the experiment. It is very important that you follow these rules. Otherwise we must exclude you from the experiment and from all payments. Please read the instructions, which are identical for all participants, carefully. Whenever you have a question or a concern, please raise your hand and one of the experimenters will come to your aid.

You will receive €4 for participating in this experiment. Beyond this you can earn additional money. How much you will earn depends on the decisions made by you and other participants. The participation fee and any additional amount of money you may earn will be paid to you in cash at the end of the experiment. Please notice: The amount of money you will collect during the experiment are denoted in "Taler". 5 Taler correspond to €1.

$$\boxed{5 \text{ Taler} = €1}$$

DETAILED INFORMATION REGARDING THE EXPERIMENT

For this experiment you will be randomly matched in groups of 4 people. You will not learn the identity of your group members, neither now nor at the end of the experiment. At the beginning of the experiment, we will ask all of you to complete a task. In the following, we explain the task in more detail.

PART I: SLIDER TASK

In this part, you are asked to posit as many sliders as possible on position 50. 48 sliders will appear on your screen. You can freely move the slider between position 0 and position 100. The starting position of each slider will be chosen randomly. You are asked to posit as many sliders as possible on position 50. A slider will only count if it is positioned exactly on 50. You will have 2 minutes of time for this task.

You will do this task twice. Therefore, you will have two times two minutes of time in order to posit as many out of the 48 sliders on position 50 as possible. In between the two trials a message will appear on your screen asking you to take a break of 20 seconds.

The more sliders you will posit in the two time two minutes (=4 minutes) on position 50, the bigger the group's joint payoff. For each correctly positioned slider the group's joint output increases by one Taler.

Example:
- The group jointly posit 140 sliders on position 50. The group's joint output corresponds to 140 Taler.

PART II: REPORT OF CORRECTLY POSITIONED SLIDERS

After each group member has solved the slider task and was informed about the own number of correctly positioned sliders, each of the members is asked to simultaneously inform the other group members about the own number of solved sliders. The report of one's own number of solved sliders determines a group member's share of the group's joint payoff. At the end of Part III all group members are informed about the reported numbers of solved sliders of her group members.

Your payoff depends on (1) group's joint payoff, (2) your reported number of solved sliders, and (3) the sum of solved/reported sliders of your group members. The computer automatically calculates how many sliders your

group has jointly solved. The following formula describes your payoff:

$$\text{PAYOFF} = \text{GROUP'S PAYOFF} \cdot \frac{\text{OWN REPORT OF SOLVED SL.}}{\sum \text{MEMBERS' SOLV./REPOR. SL.} + \text{OWN REPOR. SL.}}$$

Your payoff is greater:

- the greater is the group's joint payoff, thus, the more sliders you and your group mates have jointly and correctly solved;
- the greater is the relation between your own reported number of solved sliders and the number of solved/reported sliders of your group members added to your own number of reported sliders.

Example: The group has jointly solved 140 sliders. The group's joint payoff is 140 Taler. A group member (called member "G") reports 30 solved sliders. The computer automatically determines that the group members of "G" have correctly and jointly solved 110 sliders./The group member's of "G" report to have solved 110 sliders. "G"'s payoff equals:

$$\text{G's payoff} = 140 \text{ Taler} \cdot \frac{30}{110+30} = 30 \text{ Taler}$$

PART III: ONE FURTHER TASK

After all group members have reported their own number of solved sliders and before learning what their group members reported, each of the group members is asked to fulfill one more task. You will get direct information regarding this further task on your screen.

PAYOFF

Your final payoff will be calculated in the following way:

- Three out of four group members will be paid according to their reports in Part II.

- The fourth member will be paid according to one of her decisions in Part III.

A random draw will decide which group member will be paid according to one of her decisions in Part III. Each group member has the same chance to be paid based on her decision in Part III. Please make all decisions in Part II and III carefully as you don't know whether your payoff depends on Part II or III.

PRACTICE ROUND AND CONTROL QUESTIONS

Before the experiment starts, you will get the time to familiarize yourself with the slider task. For this, 48 sliders will appear on your screen. You have 2 minutes of time to posit as many sliders as possible on position 50. Please notice: The sliders you will posit in these initial 2 minutes on position 50 will not count for your payoff. The practice round only serves to get yourself familiar with the task.

After you got familiar with the task, we will ask you to answer some control questions, which will ensure that all participants have correctly understood the experiment.

Please stay quietly at your seats during the whole experiment. If you have questions, raise your hand. Please click on "OK" on your screen if you have read the instructions and you don't have questions.

INSTRUCTIONS PART III (HANDED OUT SEPARATELY AND AFTER COMPLETION OF PART II)

In Part III you are asked to report for each possible case how many sliders you have solved:

Example:
You have solved "x" sliders.

- *Case I:* None of your group members has reported a different than the actual number of solved sliders. How many sliders do you report to have correctly solved?
- *Case II:* How many correctly solved sliders do you report if one of your group members reports to have solved a different than the actual number of solved sliders?
- *Case III:* What if two group members report a different than the actual number of solved sliders?
- *Case IV:* How do you behave if three group members report a different than the actual number of solved sliders?

Your payment will be determined in the following way:

The software will randomly pick one group member (in the following we call that group member "M"). For this group member Part III is payoff relevant. The other three group members will be paid according to their reports in Part II. In addition, the reports provided by these three members will determine which case in Part III is payoff relevant for "M".

An example:

The group members "A", "B" and "C" have respectively solved "20", "40" and "50" sliders. In Part II they report the following number of solved sliders:

- Member “A” reports to have solved “20” sliders,
- Member “B” reports to have solved “80” sliders,
- Member “C” reports to have solved “50” sliders.

This means, 1 out of 3 group members has reported a different than the actual number of solved sliders. Now consider the number of solved sliders of member “M”, who has solved “30” sliders. How many solved sliders did “M” report in the case 1 out ouf 3 group members has reported a different than the actual number of solved sliders? Suppose “M” has in that case reported to have solved “30” sliders.

- The group’s joint payoff equals to: 20+40+50+30=140 Taler
(- The number of reported sliders of all group members equals to: 20+80+50+30 = 180)

- The group members earn the following payoffs:

- “A” earns: 140 Taler $\cdot \frac{20}{120+20/180}$ = 20/16 Taler
- “B” earns: 140 Taler $\cdot \frac{80}{100+80/180}$ = 62/62 Taler
- “C” earns: 140 Taler $\cdot \frac{50}{90+50/180}$ = 50/39 Taler
- “M” earns: 140 Taler $\cdot \frac{30}{110+30/180}$ = 30/23 Taler

Summary:
In each group three members will be paid according to the number of reported solved sliders in Part II. These reports of correctly solved sliders in Part II determine how many sliders member “M” reports to have solved in that specific case in Part III. Group member “M” will be paid according to one of her reports in Part III.

As soon as all members have read the instruction for Part III and there are no questions, click on “OK” on your screen. Before Part III starts you will be asked to answer some control questions.

MATCH IV TRUST SCALE

Answers are given on a 7-point scale ranging from strongly agree, somewhat agree, slightly agree, no opinion, slightly disagree, somewhat disagree, strongly disagree.

1. Never tell anyone the real reason you did something unless it is useful to do so.
2. The best way to handle people is to tell them what they want to hear.
3. One should take action only when sure it is morally right.
4. Most people are basically good and kind.
5. It is safest to assume that all people have a vicious streak, and it will come out when they are given a chance.
6. Honesty is the best policy in all cases.
7. There is no excuse for lying to someone else.
8. It is hard to get ahead without cutting corners here and there.
9. All in all, it is better to humble and honest than important and dishonest.
10. When you ask someone to do something for you, it is best to give the real reason for wanting it rather than giving reasons that carry more weight.
11. Most people who get ahead in the world lead clean, moral lives.
12. The biggest difference between most criminals and other people is that criminals are stupid enough to get caught.
13. Most men are brave.
14. It is wise to flatter important people.
15. It is possible to be good in all respects.

16. Barnum was very wrong when he said that there's a sucker born every minute.

17. Generally speaking, men won't work hard unless they're forced to do so.

18. People suffering from incurable diseases should have the choice of being put painless to death.

19. Most men forget more easily the death of their father then the loss of their property.

20. Anyone who completely trusts anyone else is asking for trouble.

CHAPTER 4

THE EFFECT OF ETHICAL RESPONSIBILITY ON PERFORMANCE

Abstract

In a laboratory real-effort experiment we study the effect of responsibility on performance. Specifically, we analyze whether being responsible for the decision of one's own work environment to be ethical or unethical affects workers' performance. We find that people who prefer to work in an ethical work environment perform better if they are responsible for this ethical work environment. Providing workers with responsibility for their work environment might, therefore, serve as a non-pecuniary incentive, however only for those workers who bear high ethical costs. If workers bear low ethical costs and, hence, choose to perform in an unethical work environment, we do not find the same positive effect of being responsible for one's own work environment. Furthermore, we see that workers care about their work environment even in situations where they are not responsible for this environment. Specifically, workers who are forced to work in an environment that violates their own ethical standards perform worse than workers whose own ethical standards are not violated by an imposed environment.

This chapter is joint work with Caroline Stein.

4.1 INTRODUCTION

Financial incentives are a well-studied and widely used mechanism to increase workers' productivity (Lazear, 2000; Shearer, 2004; Bandiera et al., 2005). At the same time, it is also widely acknowledged that monetary compensation is not the only factor that influences workers' motivation. In the past decades, economists and psychologists have emphasized the role of non-financial incentives and investigated how much the work environment affects workers' productivity. In this line of research, it has been shown that employees put more effort in what they do when they think their job is meaningful (Ariely et al., 2008; Chandler and Kapelner, 2013) or when their job has a mission, i.e., it generates donations to a charity (Fehrler and Kosfeld, 2014; Tonin and Vlassopoulos, 2015; Charness et al., 2016; Cassar, 2018). These last results suggest that workers are more motivated if they work in an environment which matches their ethical values and social standards. For a firm, however, adapting the work environment to a worker's values can be difficult and costly to implement. Therefore, it might be useful to leave employees more discretion to shape their work environment in a way which meets their own social and ethical standards.

Giving employees decisions rights to enable them to match their ethical and social values at their work place has become more common recently. One example is the US retailer Zappos. The current CEO Tony Hsieh introduced a no-script policy for those who work in the costumer service. He argues that employees should be able to "let their true personalities shine during every phone call" by giving them the freedom to interact freely with the customers. Thus, workers can make their own decision whether to consult in a way best for the clients or to sell unnecessary services. This is in contrast to the usual practice of firms to provide written scripts where workers are strongly encouraged to sell services by all means. We argue that by granting employees more room to shape their own work environment, workers might not only be more motivated because they act in accordance to their own values but also because they feel responsible for the work environment they act in.

In our paper, we investigate if this feeling of responsibility for one's own work environment can itself serve as an incentive. This question roots from psychological research arguing that the feeling to have self-determined one's own behavior and, thus, the feeling of being responsible for an environment plays a decisive role on individual's intrinsic motivation (Deci and Ryan, 1985; Ryan and Deci, 2000; Gagné and Deci, 2005). The literature suggests that behavior which is determined by one's own values or preferences makes the individual feel more coherent with one's self and, thus, more committed to the action one executes.

There is so far only little empirical evidence whether responsibility for one's own work environment affects performance. In a study, Fehrler and Kosfeld (2014) find that participants who choose themselves to contribute to a charity exert more effort compared to participants who are randomly assigned to contribute to a charity. The authors explain this effect by self-selection and do not consider the increase in performance to be additionally caused by the fact that individuals in the former case could choose and, thus, were responsible for the contributions to the charity. Separating these two explanations from each other is, indeed, difficult because they usually go hand in hand: As soon as an individual can choose and select into an option, she is also responsible for that option. Therefore, choosing an ethical or pro-social work environment might increase motivation through (i) the fact that the individual self-selects and, hence, acts according to her preferences or (ii) the feeling of being responsible.

In our experiment, we control for selection effects and focus on how being responsible for an ethical or unethical work environment affects performance. In particular, we match participants into pairs composed of one worker and one employer. The worker and the employer are each independently assigned a piece-rate that will determine the payoff for a task the worker eventually has to perform. The employer and the worker are then asked to report their own designated piece-rate. In this stage, participants usually have the opportunity to report a higher piece-rate than they were actually assigned to. The reported piece-rate will determine the payoff of both players in the same way. Consequently, the employer and the worker have both

the same incentive to over-report the piece-rate. At the same time, we explicitly declare over-reporting a violation of the rules of this experiment. In this sense, workers and employers face the trade-off between following the potential ethical principle of adhering to a stated rule or deviating from it for financial gains. In the following, we will define it an "ethical" work environment when the worker performs under the designated piece-rate. Conversely, we will call it an "unethical" work environment when the worker and the employer receive a payoff based on an over-reported piece-rate. Furthermore, we introduce a randomization procedure which either implements the reporting decision of the worker or of the employer. This creates our two major "treatment" conditions: *Responsibility* and *NoResponsibility*. In case the worker's work environment is implemented, we refer to the treatment condition as *Responsibility* and as *NoResponsibility* otherwise.

First, we find that in an ethical environment workers perform better when they are responsible for this environment compared to workers who perform in an ethical work environment the employer is responsible for. This result might potentially be driven by selection because the latter case additionally includes workers who prefer to work in an unethical work environment. We, however, want to focus on the effect of responsibility only. Therefore, we concentrate only on those workers who act in an ethical work environment implemented by the employer but would have anyway preferred to act in such an environment. When we compare their performance to the performance of workers who act in an ethical environment they are responsible for, we still find that performance is higher for workers who are responsible for the ethical work environment. This gives us reason to infer that responsibility has an effect on performance. However, we do not find the same effect for an unethical environment. In particular, workers who are responsible for an unethical environment do not seem to be more motivated. We explain this by another aspect coming along with responsibility: people who are fully responsible for an unethical work environment bear the full ethical cost connected to this action. In contrast, when someone else is responsible for the implementation of an unethical work environment ethical costs can be

shifted. Therefore, we argue that if a worker is responsible for an unethical work environment, the higher ethical costs counteract the motivational gain from being responsible. As a third result, we find that when workers have to perform in an unethical environment their employer is responsible for, ethical workers perform significantly worse than the unethical workers. We interpret this result as evidence for ethical costs to even incur in cases where one is not responsible for an unethical work environment.

With our research we add to a field of experimental literature studying the relationship between responsibility and effort provision. So far this relationship has only been investigated in gift exchange settings, where pro-social preferences explain positive effort choices. Charness (2000), for instance, finds that agents provide more effort when a random procedure determines agents' wages compared to a situation where a third party chooses agents' wages. The author explains this finding with differing degrees of responsibility occurring in the two situations. In presence of a random procedure, an agent feels solely responsible for final payoffs, while the same agent can shift part of this responsibility when a third party has determined the wage. Responsibility for final outcomes seems to enhance pro-social motivations and through this channel increase effort. In line with this result, Charness et al. (2012) show that when employers delegate their wage decision to workers, workers respond with higher effort. Controlling for other possible explanations such as reciprocity, the authors again show that feeling more responsible for the outcome seems to be the driving factor of the higher effort provision. Likewise, Falk and Kosfeld (2006) find that when principals restrict the effort choice set of the employees by setting a minimum effort level - and thus also employees' responsibility - employees perform worse than in case where choice sets are not restricted. In contrast to these studies, we approach the relationship between responsibility and effort provision not through the channel of pro-social concerns but concentrate on whether being responsible for an ethical work environment creates additional motivation, which also might lead to an increase in effort provision.

As a specific feature of our design, we are able to separate the effect of self-selection when making a decision from the effect of responsibility

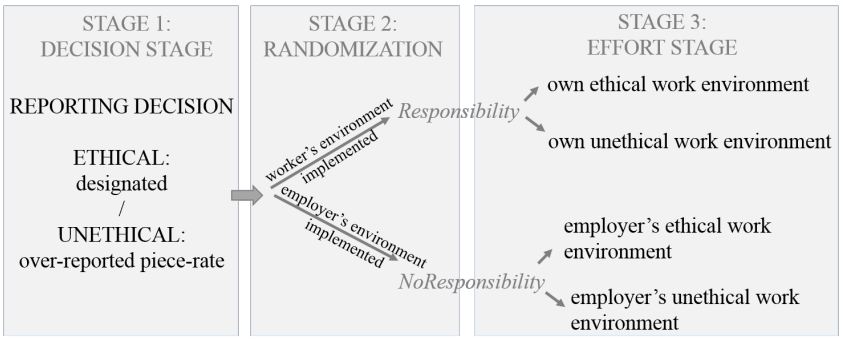
by following a procedure which is similar to the one used in a paper by Dal Bó et al. (2010). The authors in the mentioned paper study what effect democratically establishing an institution has on subsequent behavior. They find that people cooperate more if they can democratically establish an institution compared to having the same institution exogenously imposed on them. What the authors, in this context, call "endogeneity premium" can be interpreted in our setting as the valuation of being responsible. Also a study by Babcock et al. (2015) finds that the mere act of choosing and, thus, being responsible, might be motivating for individuals. In a field experiment, they find that people perform better if they can choose their incentive scheme compared to a situation where they are randomly assigned to it. Even if these studies do not consider any ethical or pro-social dimension, they present some evidence that the sole act of choosing might be motivating for individuals.

The remainder of this paper is organized as follows. Section 4.2 and 4.3 describe the experimental design and procedure. In section 4.4 we introduce a basic model that shall illustrate our behavioral hypotheses about the expected treatment effects. Section 4.5 presents the results of the study. Section 4.6 concludes.

4.2 EXPERIMENTAL DESIGN

We run a laboratory experiment where we group subjects into pairs and randomly assign one the role of an employer and the other the role of a worker. The experiment consists of three stages (see Figure 4.1).

Figure 4.1: Experimental design



Stage 1: Decision stage

In the decision stage, the employer and the worker are independently assigned a piece-rate.[1] They are then asked to simultaneously report this piece-rate, which in the effort stage will be used to calculate the earnings for both the employer and the worker. In particular, in the effort stage the worker performs a real effort task which in 50 percent of the cases is remunerated based on his own reported piece-rate and in 50 percent of the cases based on the piece-rate reported by the employer. Earnings are the same for employers and workers.[2] This means that a higher reported piece-rate implies higher earnings for both employer and worker.

For this reporting decision, both players find on their screens two different piece-rate options, namely 5 and 8 points per unit of effort, they can choose from. Here, those employers and workers who actually got a designated

[1]Please refer to the instructions for details. In synthesis: Subjects pick between five differently colored sealed envelopes, where each envelope contains a combination of piece-rate and real effort task. Subjects do not know the piece-rates or tasks contained in the different envelopes when picking an envelope. They also do not know whether the different envelopes yield different piece-rates or tasks. Only after picking an envelope, subjects learn about the designated piece-rate for the picked envelope. Workers are informed about the task of the picked envelope only shortly before performing the task, which happens in the third stage of the experiment.

[2]We purposefully choose to have employers and workers payoffs aligned in order to provide both with the same incentive to over-report.

piece-rate of 5 assigned - which happens in 4 out of 5 cases[3] - have the possibility to over-report. The experimental instructions explicitly clarify that reporting a higher piece-rate violates the rules of the experiment. Specifically, the instructions say: "Reporting a piece-rate that is different from your designated piece-rate is considered a violation of the rules. If you do so anyway, your earnings will be calculated based on your reported but not the designated piece-rate. You will, therefore, receive a piece-rate which does not correspond to the piece-rate designated for your task". In this sense, we define a work environment to be "ethical" when the earnings are multiplied by the reported, designated piece-rate and a work environment to be "unethical" when the earnings are multiplied by an over-reported piece-rate. Please note that workers will learn the nature of the real effort task only after the decision stage. By doing so, we can rule out that employers or workers condition the choice of acting in an ethical or unethical work environment on their belief about the worker's performance in the effort stage.

From now, we will only consider decisions of participants that were assigned a piece-rate of 5 points per unit of effort and, therefore, faced an actual trade-off between following the ethical principle of adhering to the stated rule or deviating from this principle for financial gains. This means that, in the following, in an ethical work environment the piece-rate always equals 5 points per unit of effort, while in an unethical work environment the piece-rate always yields 8 points per unit of effort.

Stage 2: Randomization

After both participants reported their individual piece-rate, the reported piece-rate of either the employer or the worker is implemented with an equal probability. The participants know this procedure beforehand. The randomization device creates our "treatment" conditions. In treatment

[3]Employers and workers have no information regarding the exact distribution of piece-rates. From the reporting decision on the screen they can infer that both piece-rates, 5 and 8 points per unit of effort, are possible.

condition *Responsibility*, the worker performs a task which is compensated based on a piece-rate he has reported himself. Thus, he either works in an ethical or unethical environment he is responsible for. In treatment condition *NoResponsibility*, the worker performs a task that is compensated based on a piece-rate the employer has reported. In this treatment, the worker is informed in addition to the employer's reported also about the employer's designated piece-rate in order to ensure that the worker - when performing the real effort task - knows whether the employer did over-report the piece-rate or not. Consequently, in the *NoResponsibility* treatment condition the worker acts in an ethical or unethical environment the employer is now responsible for.

Stage 3: Effort stage

In the effort stage, workers are asked to count the occurrences of "7s" that appear in sequential number sequences.[4] The payoffs for employers and workers are obtained by multiplying the worker's number of correctly solved sequences by the reported piece-rate. The number of correctly solved tasks will be our measure for performance.[5]

Table 4.1 displays all relevant conditions[6] in which the worker can eventually perform the task. First, we distinguish between *Responsibility* and *NoResponsibility*, which we exogenously vary through the randomization stage and, therefore, call our "treatment conditions". In case of *Responsibility*, the worker performs under his own decision. Based on the worker's own decision, he can be responsible for either an ethical or an unethical work

[4]The real effort task is similar for the different colored envelopes. The only thing which changes is the color in which the number "7" appears in the sequences of numbers.

[5]We use the number of correctly solved tasks as our measure of performance as this measure is incentivized. In the appendix we additionally report the results when we use the total number of attempted matrices as our measure of performance. It seems that the effect of responsibility does not reveal in the quantity of work but in the quality of work.

[6]In case of *Responsibility*, the worker does not receive any information about the decision the employer made before she performs the task. The instruction make this aspect clear to the participants. We, therefore, in *Responsibility* do not distinguish between the employer's decision as we do for *NoResponsibility*.

environment.[7] In case of *NoResponsibility*, the worker performs under the employer's decision. He then works in an environment he is not responsible for. In addition, we distinguish in this latter "treatment condition" between the cases where the worker performs according to his preferences or not. We call it a matched worker when a worker who prefers to act in an (un)ethical environment actually performs in an (un)ethical environment. Conversely, we call it a mismatched worker if he acts in an environment he does not prefer. At the end of the experiment, the employer and the worker receive information regarding the number of correctly solved sequences by the worker, the reported piece-rate, and the earnings of the pair.

4.3 EXPERIMENTAL PROCEDURES

We ran our experiment at the University of Cologne using the software ztree Fischbacher (2007) in January and April 2017. The participants earned €1 for 25 points. On average, the participants earned €10.50 including the show-up fee. The instructions were common knowledge and read out loud at the beginning of each session. In total, 312 subjects participated in our experiment.[8] We have 68 observations in our treatment condition *Responsibility* and 57 observations in *NoResponsibility*. Table 4.1 displays the exact numbers of observations for each condition as described in our experimental design. We preregistered our experiment at AEA RTC Registry, Nr. 1956.

[7] Reporting a higher piece-rate can also be considered an ethical behavior because it not only increases one's own but also the other participant's payoff. If employers and workers consider over-reporting ethical, we will, in the result section, under-estimate our results for the unethical case because the effect of ethical costs opposing the effect of responsibility will be reduced.

[8] We have to drop the observation of one participant who wrote in the questionnaire: "I was tired and, unfortunately, did not know which role I was." This means that this person did not know if his effort would actually count for the payoff calculation, since employers were also asked to perform a real-effort task. However, employers' performance did not count for final payoffs. This was clearly stated in the instructions.

Table 4.1: Relevant conditions in which worker possibly performs the task

	NoResponsibility Employer's environment implemented		**Responsibility** Worker's environment implemented
	Ethical worker	**Unethical worker**	
Ethical environment implemented	Worker is *not responsible* for the *ethical environment* and acts *according* to his preferences N=29	Worker is *not responsible* for the *ethical environment* and does *not* act *according* to his preferences N=12	Worker is *responsible* for the *ethical environment* N=48
Unethical environment implemented	Worker is *not responsible* for the *unethical environment* and does *not* act *according* to his preferences. N=9	Worker is *not responsible* for the *unethical environment* and acts *according* to his preferences. N=7	Worker is *responsible* for the *unethical environment* N=20

4.4 CONCEPTUAL FRAMEWORK

We provide a basic model in order to evaluate which effect being responsible for an ethical or unethical work environment might have on performance. First, individuals decide between piece-rate p^H, p^L with $p^H > p^L$ and where the decision for p^H implies a rule violation, and, thus, creates an unethical work environment. After the workers have reported a piece-rate $p^j \in \{p^L, p^H\}$, they choose an effort level e^* which maximizes:

$$U_i(e, \pi_i, \delta_i, \alpha) = p^j e - c_i e^2 + r\pi_i e - [rv\delta_i + (1-r)v(1-\alpha)\delta_i]e$$

Our utility function is composed of the produced outcome, $p^j e$, minus a standard effort cost function with increasing marginal costs, $c_i e^2$. We extend this utility function now by two additional components. First, we allow for a utility gain from being responsible for a decision, $r\pi_i$, where $\pi_i \in [0, \overline{\pi}]$ is the value of responsibility and where $r \in \{0, 1\}$ denotes the treatment condition. We assume that this value π_i only applies in a situation where a worker acts in a work environment he is responsible for, which is only the case in our treatment condition *Responsibility* ($r = 1$). Furthermore, we assume a constant marginal gain from responsibility. We do not claim that this assumption is generalizable, but we think it derives from our design. In our experiment, the worker can be responsible for the piece-rate which applies for each unit of effort. In this way, the feeling of being responsible for acting in an ethical or, respectively, unethical work environment reproduces for every unit of effort.

Second, we introduce an ethical cost, δ_i, where $\delta_i \in [0, \overline{\delta}]$. We assume that this ethical cost applies in situations where the worker acts in an unethical work environment, where $v = 1$. We again assume that acting in an unethical environment increases effort costs at a constant marginal rate. This again derives from our experiment, where the unethical environment results from an over-reported piece-rate the worker earns for each unit of effort. Therefore, the violation of a stated rule reproduces for each unit of effort and creates the ethical cost again and again. To capture the phenomenon of shifting ethical costs, we allow the cost of being unethical to be deduced by α, where $0 \leq \alpha \leq 1$, in case someone else is responsible for the unethical work environment. In particular, we refer with this additional parameter to previous literature showing that the costs of being unethical or unfair can be (partly) shifted to the person who actually made the decision (Bartling and Fischbacher, 2011; Oexl and Grossman, 2013). The worker then chooses an optimal effort level e^* dependent on the condition of our design (see Table 4.2).

Table 4.2: Optimal effort levels by condition

	NoResponsibility Employer's environment implemented		**Responsibility** Worker's environment implemented
	Ethical worker	**Unethical worker**	
Ethical environment implemented	$e^* = \frac{p^L}{2c_i}$	$e^* = \frac{p^L}{2c_i}$	$e^* = \frac{p^L + \pi_i}{2c_i}$
Unethical environment implemented	$e^* = \frac{p^H - (1-\alpha)\delta_i}{2c_i}$	$e^* = \frac{p^H - (1-\alpha)\delta_i}{2c_i}$	$e^* = \frac{p^H + \pi_i - \delta_i}{2c_i}$

Backward induction reveals the threshold of ethical costs for which individuals are indifferent between choosing a higher piece-rate by making an unethical decision and being honest with the lower piece-rate.

$$\hat{\delta} = p^H - p^L$$

To give the intuition, this threshold simply illustrates that individuals are predicted to choose an ethical work environment only if the marginal cost from deviating from a stated rule (δ_i) exceeds its marginal benefit, such that $\delta_i > \hat{\delta}$.

If we assume all parameters to be zero, we would not find any ethical work environments put in place. For a standard selfish decision-maker, ethical costs would not apply. Therefore, all would prefer to act in an unethical work environment because only the marginal benefit of deviating from a stated rule would be taken into account. Furthermore, we would find no treatment difference in performance because being responsible would not affect effort costs of workers. We get the same prediction for any form of outcome-based social preference models. Choosing the highest piece-rate by deviating from the stated rule would increase efficiency without distorting equity because the payoffs of the employer and worker are aligned.

We will now discuss the predictions that follow if we assume the parameters we additionally introduced to be non-zero. Let us first look at

the ethical costs that would apply when a worker performs in an unethical work environment. If we assume δ_i to be distributed over a sufficiently large interval,[9] we would find a fraction of workers that chooses not to deviate from the stated rule and consequently chooses to perform in an ethical work environment. Furthermore, since the threshold does not depend on the worker's productivity, the decision itself to perform in an ethical or unethical work environment would not lead to differences in performance between *Responsibility* or *NoResponsibility*.

If we also assume $\pi_i > 0$, workers would be motivated by the mere fact that they act in a work environment they are responsible for.[10] If this is the case, we would find a higher effort level for ethical workers when they act in an ethical environment they are responsible for compared to ethical workers that act in an ethical work environment their employer is responsible for. For the unethical case, the model does not give a clear prediction because here we have two possible effects coming along with responsibility that go in different directions. Comparing unethical workers between treatment conditions, the motivational gain from being responsible might be counteracted by ethical costs that are higher if one is responsible for an unethical work environment. Therefore, the observable effect of responsibility on performance, in case of an unethical work environment, also depends on how much workers can shift the ethical cost to the employer who chose to implement an unethical work environment in the *NoResponsibility* condition.

Hypothesis 1. *Workers who choose to report truthfully exert more effort when they are responsible for the ethical work environment than when the employer is responsible for the ethical work environment.*

Let us now consider what happens when ethical costs can be shifted. We, first, assume $\alpha = 1$ which illustrates the case where workers can fully

[9]Specifically, we need to assume here that δ_i is distributed over an interval $[0, \overline{\delta}]$, where $\overline{\delta} \geq \hat{\delta}$.

[10]For the sake of simplicity, we assume $\pi_i > 0$ to be uncorrelated with the type of the work environment - whether it is an ethical or unethical work environment. Alternatively, one could argue that responsibility has a (stronger) impact only in a situation, where the individual chooses an ethical work environment. Our results cannot rule out this alternative approach.

shift their ethical costs δ_i to the employer in *NoResponsibility*. Unethical workers with *Responsibility* would add the component $\frac{\pi_i - \delta_i}{2c_i}$ to their optimal effort provision compared to unethical workers in treatment condition *NoResponsibility*. The direction of the effect, therefore, depends on the relative magnitudes of π_i and δ_i. Furthermore, with $\alpha = 1$ we would find no difference in effort provision between ethical and unethical workers who act in an unethical work environment the employer is responsible for (*NoResponsibility*).

In contrast, if we assume $1 > \alpha > 0$ workers would not be able to fully shift their ethical costs in the *NoResponsibility* condition. When an unethical work environment is implemented, the ethical costs would, thus, be lower in *NoResponsibility* than in *Responsibility*. This would decrease the performance of the unethical worker in *Responsibility* compared to *NoResponsibility*, potentially offsetting the motivational gain from being responsible. It would also mean that unethical workers in the treatment condition *NoResponsibility* would choose a higher optimal effort level than ethical workers when an unethical environment is implemented. This happens because those workers who prefer an ethical environment revealed a $\delta_i > \hat{\delta}$ and therefore higher ethical costs than workers who prefer an unethical work environment, where $\delta_i < \hat{\delta}$.

In our experiment, we expect workers to not be able to fully shift ethical costs in case of *NoResponsibility* ($1 > \alpha > 0$) because, even if the employer chooses the unethical work environment for them, workers have to perform in this environment. Therefore, workers cannot fully dissociate from the unethical behavior. A partial shift of ethical costs would also be consistent with results from Oexl and Grossman (2013) which show that observers blame individuals who carry out an unethical behavior even if they did not initiate it themselves and were not able to correct it.

Hypothesis 2. *In case workers act in an unethical work environment the employers are responsible for, those workers who actually prefer to act in such an environment perform better than workers who would prefer to act in an ethical work environment.*

4.5 RESULTS

In the following, we compare the workers' mean performance levels between the different conditions distinguished by treatments and workers' types as described above. We will only compare conditions with the same incentive scheme. We use the non-parametric Mann-Whitney-U test, which is suitable for independent observations and small samples. In the first section, we will look at aggregated outcomes without controlling for selection. To address the hypotheses we derived from our framework, we will then look at the specific types of workers to see (i) if responsibility itself has an effect on performance and (ii) if ethical costs can be (partly) shifted in case of *NoResponsibility*.

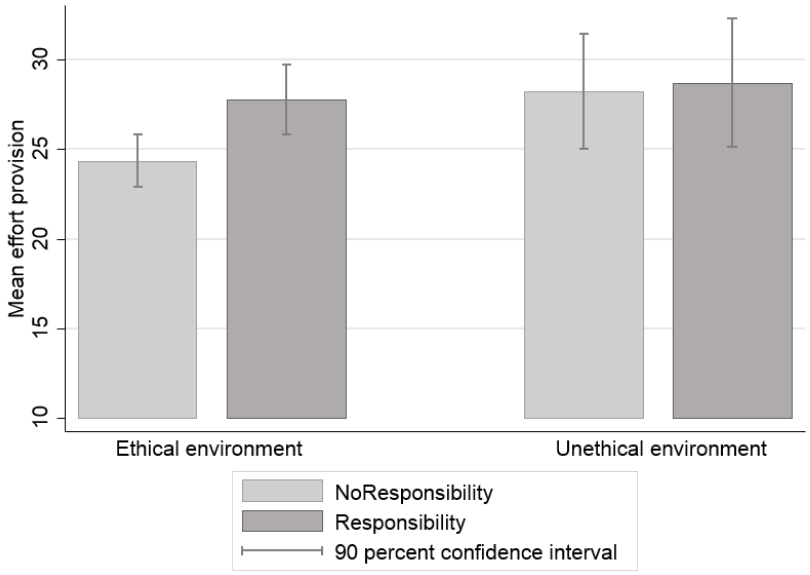
4.5.1 *RESPONSIBILITY* VERSUS *NORESPONSIBILITY*

Figure 4.2 compares the mean performance of workers who act under *Responsibility* (dark gray bars) to workers that act under *NoResponsibility* (light gray bars). The left panel illustrates the workers' performance when an ethical work environment is implemented, whereas the right panel displays the mean performance when the workers act in an unethical environment. First, the figure shows that apparently some participants do prefer to act in an ethical work environment. Specifically, we find a remarkable fraction of 70 percent of the workers who do not deviate from the stated rule and report the piece-rate truthfully. This suggests that people do incur ethical costs from reporting a higher piece-rate in our setting.[11]

We now focus on the workers who act in such an ethical work environment and, consequently, consider the left panel of Figure 4.2. We see a clear increase in performance when a worker himself is responsible for the ethical work environment compared to a situation where the employer is responsible for the ethical work environment. With an increase of 15% in mean performance, this difference is economically and statistically significant (p=.049, Mann-Whitney-U-test). We interpret this as first evidence that being responsible for one's own ethical work environment increases

[11] We find the same fraction of employers choosing to implement an ethical work environment.

Figure 4.2: Performance of workers in an ethical and unethical work environment by treatment condition



performance. However, we will soon show that this effect is not yet pure of self-selection issues.

The right panel of Figure 4.2 displays workers' performance in case an unethical work environment was implemented and, thus, under high incentives. We see no increase or decrease in performance when the worker is responsible for the unethical work environment compared to a situation where the worker is not responsible for it. We interpret this as first evidence for the positive effect of responsibility being offset by the full ethical costs one bears in case one is responsible for the implementation of an unethical work environment. In contrast, when someone else is responsible for the unethical work environment, ethical costs can be shifted. We observe that the higher ethical cost under *Responsibility* counteract the motivational gain from being responsible. Again, also this effect is not yet pure of self-selection issues.

Result 1. *Workers who perform in an ethical work environment they are responsible for perform better than workers who perform in an ethical work*

environment they are not responsible for. We do not find a similar effect for unethical work environments.

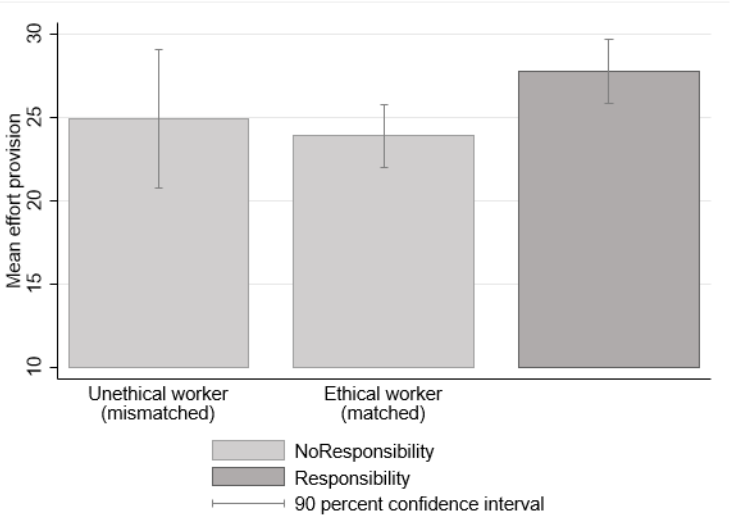
4.5.2 TYPE SPECIFIC COMPARISONS

As anticipated, Figure 4.2 does not reveal the pure effect responsibility has on performance but potentially includes self-selection issues. In particular, workers in the *Responsibility* condition may self-select and, therefore, be different from those workers who act in the same work environment in the *NoResponsibility* condition. While workers in case of *Responsibility* act, by nature, in a work environment which matches their own preferences, under *NoResponsibility* an employer's preferred work environment might match or mismatch the worker's preferred environment. This could have an effect on workers' motivation and, consequently, on their performance. However, since we also expect responsibility to contribute to the increase in performance, the positive effect of being responsible for an ethical work environment should persist even when controlling for self-selection.

In order to exclude self-selection as a potential driver for Result 1, we classify workers into ethical and unethical workers based on their own decision whether or not to follow the ethical principle of adhering to the stated rule of reporting truthfully. Then, we distinguish between situations where employer's ethical or unethical work environment matches or mismatches the worker's preferred environment. To sum up, only if we observe that matched ethical workers under *NoResponsibility* perform better than ethical workers under *Responsibility*, we can assess that responsibility and not self-selection explains our Result 1.

For this purpose, we introduce Figure 4.3 which displays mean performance of workers that act in an ethical work environment. As in Figure 4.2, we distinguish between the two conditions *Responsibility* and *NoResponsibility*. In addition, Figure 4.3 also separates the types of worker in the *NoResponsibility* condition. Thus, the two light gray bars display the mean performance of matched ethical (left gray bar) and, respectively, mismatched unethical workers (right gray bar).

Figure 4.3: Performance in an ethical work environment by treatment conditions and worker's type



We can now compare mean performance of the same types between the two different treatment conditions in order to address any possible selection effects. We, thus, compare workers acting in an ethical work environment they are responsible for (dark gray bar) to ethical workers acting in an ethical work environment they are not responsible for (right light gray bar). We still find that the difference between the dark gray bar and the left light gray bar is statistically significant (p=.037, Mann-Whitney-U-test). This confirms our Hypothesis 1 and shows that performance increases when one is responsible for one's own ethical work environment.

We now turn to the left light gray bar which displays mean performance of mismatched unethical workers in the *NoResponsibility* case. When we compare the two light gray bars, we find that unethical types perform as well as ethical types when performing in an ethical work environment the employer is responsible for (p=.518, Mann-Whitney-U-test).

Result 2. *In an ethical work environment, ethical workers perform better under Responsibility than under NoResponsibility. This result shows that*

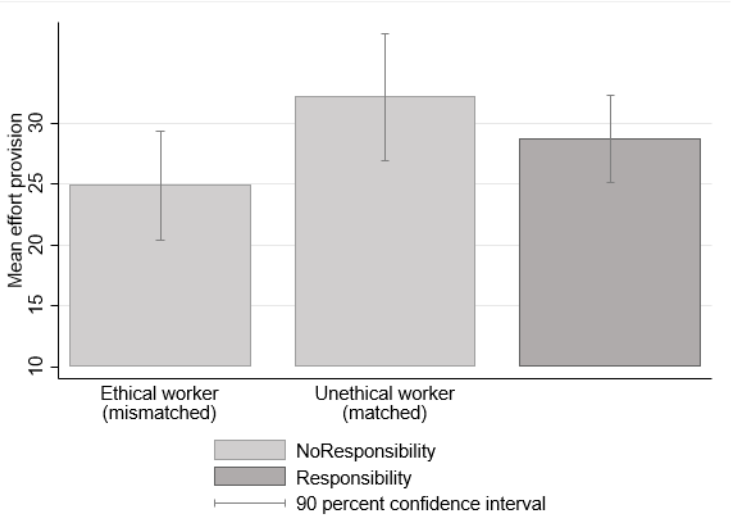
the mere fact of being responsible for one's own ethical work environment positively affects performance.

Figure 4.4 shows the mean performance of workers when acting in an unethical work environment. Again, we distinguish between the treatment conditions. The dark gray bar represents mean performance under *Responsibility* and the light gray bars stand for mean performance under *NoResponsibility*. As before, we distinguish between types, which are the mismatched ethical workers (left light gray bar) and the matched unethical workers (right light gray bar).

We first look at the workers acting in an unethical work environment they are responsible for (dark gray bar) and compare their mean performance to unethical types acting in an unethical work environment they are not responsible for (*NoResponsibility*, right light gray bar). The difference in mean performance is, if anything, going in the opposite direction compared to the ethical case. This means that in the unethical case workers perform worse under *Responsibility* than under *NoResponsibility*. This difference in performance between the two treatment conditions, however, is not statistically significant (p=.331, Mann-Whitney U-Test). Even if we cannot draw a clear conclusion from this result, it still gives some suggestive evidence for the hypothesis that under *Responsibility* workers face higher ethical costs that counteract the motivational gain from responsibility.

Lastly, Figure 4.4 also allows us to compare how ethical and unethical workers perform when they both are asked to perform in an unethical work environment they are not responsible for. To this aim, we compare the two light gray bars in Figure 4.4. We find a significant difference in performance (p=.034, Mann-Whitney-U-test). In particular, it shows that ethical workers perform significantly worse than unethical workers even if they are not responsible for the unethical work environment. This suggests that in case of *NoResponsibility*, workers are not able to fully shift their ethical costs to the employer. Consequently, the differing ethical costs ethical and unethical workers bear - although being partially shifted on to the employer - might reflect in differing performances of the two types of workers.

Figure 4.4: Performance in an unethical work environment by treatment conditions and worker's type



Result 3. *Our results provide suggestive evidence for responsibility in an unethical work environment to not increase performance as much as it does in an ethical work environment. If the employer implements an unethical work environment, ethical workers perform significantly worse than unethical workers do.*

4.6 CONCLUSION

We study whether the feeling of being responsible for one's own work environment might serve as an incentive to increase workers' performance. For this, we let workers choose between an ethical or unethical work environment. In the field, having the choice about one's own work environment usually implies not only responsibility but also allows people to sort into environments they prefer. Both mechanisms might affect workers' motivation. In a laboratory experiment and using a specific randomization technique, we can separate the first from the second possible effect. Specifically, we can compare workers that both act according to their preferences once with and once without being responsible for this environment.

We find that responsibility for a work environment might serve as an non-pecuniary incentive. However, this effect depends on the way the work environment is shaped. In particular, it seems that the incentive effect becomes effective only if workers are responsible for an ethical work environment. In this case, workers perform better than those who also prefer an ethical environment but are not responsible for it. In contrast, workers who choose an unethical work environment do not respond to the same extent to responsibility. In particular, the incentive effect of being responsible for an unethical work environment seems to be counterbalanced by ethical costs which arise when acting in a way that hurts ethical standards. From this finding, one might conjecture that responsibility can effectively increase performance only for workers with high ethical costs. These workers have a preference for working in an ethical environment even if this environment offers lower monetary incentives. Consequently, these types of workers might be more likely to be found in organizations with strong ethical and social missions because workers rather sort into environments they prefer. Those types of organizations often face tight budget constraints. Thus, for them increasing the responsibility of their employees to shape their work environment in an ethically way might be particularly relevant, as it provides organizations with a cost-saving tool to improve performance.

Our results might also have interesting implications for the field of compliance management, which installs strict ethical guidelines to effectively ensure workers' behavior to be compliant to legal and ethical standards. Our results show that monitoring might come at costs that haven't been accounted for so far. In the presence of strict ethical guidelines, the positive effect from being autonomously responsible for one's own ethical conduct cannot be exploited. This means that workers will provide less effort when they are forced to follow ethical guidelines compared to a situation where they autonomously choose to behave ethically. Apart from the important role of guidelines, a firm, in this light, might be interested in putting increased emphasis on screening carefully for workers with high ethical standards. Screening for employees with high ethical standards and granting them the possibility to actively shape their work environment might not only save

monitoring costs but also increase the motivation of workers if they can feel responsible for their own ethical work environment.

Lastly, our results show that imposing a work environment which does not match workers' ethical standards might have deteriorating effects on their performance. Coming back to our example from the beginning - most of the costumer service departments enforce very detailed scripts workers are required to exactly follow and mostly serve to up-sell services by all means. In case this procedure violates a worker's ethical values, we find that this work environment might lead to performance reduction. Specifically, we observe that workers perform worse if they prefer to work in an ethical work environment but are forced to act against their ethical standards compared to workers who themselves were willing to install an unethical work environment. Our results show that even higher monetary incentives cannot compensate for this mismatch in ethical standards. This might provide further justification for allowing employees to shape their work environment in a way that meets their ethical standards. Only in this way firms can ensure that employees with high ethical standards will be able to fully unfold their potential.

4.A APPENDIX

PROOF

$U_i(e, \pi_i, \delta_i, \alpha) = p^j e - c_i e^2 + r\pi_i e - [rv\delta_i + (1-r)v(1-\alpha)\delta_i]e$

$\frac{\partial U}{\partial e} = p^j - 2c_i e + r\pi_i - rv\delta_i - (1-r)v(1-\alpha)\delta_i$

Optimal effort $\Rightarrow \frac{\partial U}{\partial e} = 0$

Ethical decision implemented with Responsibility (r=1, v=0):
$p^L - 2c_i e + r\pi_i = 0 \Rightarrow e^* = \frac{p^L + \pi_i}{2c_i}$

Ethical decision implemented with NoResponsibility (r=0, v=0):
$p^L - 2c_i e = 0 \Rightarrow e^* = \frac{p^L}{2c_i}$

Unethical decision implemented with Responsibility (r=1, v=1):
$p^H - 2c_i e + r\pi_i - \delta_i = 0 \Rightarrow e^* = \frac{p^H + \pi_i - \delta_i}{2c_i}$

Unethical decision implemented with NoResponsibility (r=0, v=1):
$p^H - 2c_i e - (1-\alpha)\delta_i = 0 \Rightarrow e^* = \frac{p^H - (1-\alpha)\delta_i}{2c_i}$

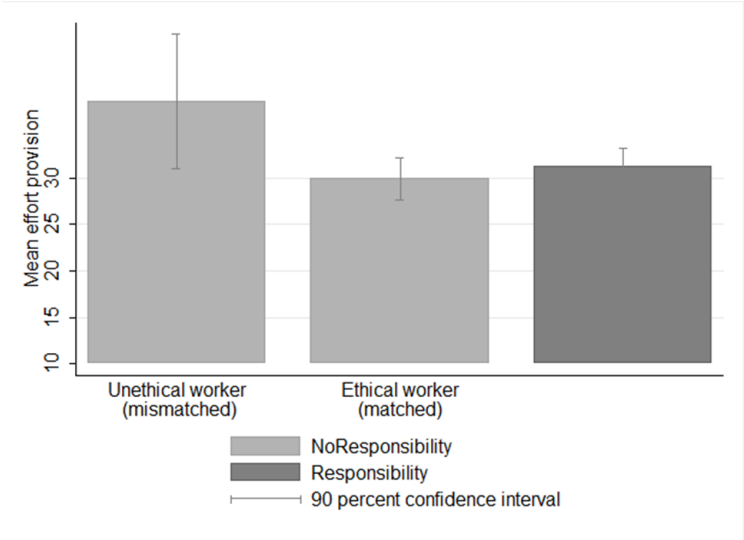
Backward induction yields $\hat{\delta}$ for which individuals prefer being honest with the lower piece-rate than choosing a higher piece-rate by making an unethical decision:

$p^L(\frac{p^L+\pi_i}{2c_i}) - c_i(\frac{p^L+\pi_i}{2c_i})^2 + \pi_i(\frac{p^L+\pi_i}{2c_i}) \geq$
$p^H(\frac{p^H+\pi_i-\delta_i}{2c_i}) - c_i(\frac{p^H+\pi_i-\delta_i}{2c_i})^2 + \pi_i(\frac{p^H+\pi_i-\delta_i}{2c_i}) + \delta_i(\frac{p^H+\pi_i-\delta_i}{2c_i})$

$\frac{1}{4c_i}(p^L + \pi_i)^2 \geq \frac{1}{4c_i}(p^H + \pi_i - \delta_i)^2$

$\delta_i \geq p^H - p^L$

Figure A-1: Number of attempted matrices with the ethical work environment being implemented by treatment conditions and worker's type



When using the total number of attempted matrices - in contrast to the total number of correctly solved matrices - we find no effect of responsibility on performance. Ethical workers acting under *Responsibility* do not try to solve more matrices compared to ethical workers acting in an ethical work environment imposed by the employer (31.6 vs 32.4; p=.256, Mann-Whitney-U Test). In this sense, not the quantity of the work but the quality of the work improves in the presence of responsibility.

INSTRUCTIONS

Instructions translated from German. German instructions available upon request.

Welcome to this experiment! Please read the first page of the instructions and the part containing detailed information regarding the experiment carefully. We will afterwards read all parts aloud.

If you read everything carefully, you can, additionally to the €4 show-up fee, earn extra money based on your decisions and the decisions of another participant in the experiment. It is, thus, very important that you read the instructions carefully. If you have questions, please raise your hand. We will come to your place and answer your questions. During the experiment, you are not allowed to talk to the other participants or to use your phone.

During the experiment we will not refer to euros but points. Your earnings will, hence, be expressed in points and only at the very end of the experiment converted into euros. The following exchange rate applies:

$$25 \text{ points} = €1$$

At the end of the experiment, you will get the €4 show-up fee plus the money you additionally earned paid out in cash.

Neither before the experiment nor afterwards will you learn the identity of the participants you will interact with during this experiment. The same applies for the other participants who as well will not learn anything about the identity of the other participants.

On the following pages, we will provide detailed information regarding the experiment.

Detailed information regarding the experiment

For this experiment you will be paired in two. One member will take on role "A" and the other member will take on role "B". You will learn about your role shortly before the beginning of the experiment by reading this information on the screen. If you are member "A" you will get a member with role "B" assigned. If you are member "B" you will get a member with role "A" assigned.

Task of member "B":

Member "B" has to execute a task. For this task, "B" has 10 minutes of time. "B" will fulfill one of the following five possible tasks:
- task "green",
- task "blue",
- task "orange",
- task "red",
- task "yellow".

Each of these tasks is divided in smaller subtasks. Member "B" will learn about the nature of these subtasks and the assigned piece-rate only shortly before conducting the task. The more subtask "B" solves, the higher the earnings of the pair.

Choice of the task:

"A" and "B" choose simultaneously which task "B" should execute. With 50% probability the chosen task of "A" and with 50% probability the chosen task of "B" will be implemented.

For the choice, which task "B" should execute, an experimenter will approach each of you and present you 5 closed differently colored envelops. Each envelop stand for the corresponding colored task. Each participant

can choose one out of these five envelops:

- task “green”,
- task “blue”,
- task “orange”,
- task “red”,
- task “yellow”.

Please notice! When picking a task you will neither know the nature nor the associated piece-rate of the task.

Communication of task and piece-rate:

As soon as you have chosen a task, you are asked to open the envelop. The envelop contains three pieces of information:
- a password you type to inform the computer about our chosen task,
- the designated piece-rate for your chosen task, on which you click on your screen to let the computer calculate your earnings,
- a number code which will become relevant later.

Please insert the first two pieces of information on your screen:
1) Make sure that the password is correct. Only if the password is correct, the correct task can be displayed on the screen.
2) Make sure that the piece-rate is correct. The task can only be multiplied by the designated and not a different piece-rate if the piece-rate has been reported correctly.

Earnings:

Please notice: The earnings are the same for both members of the pair and are obtained by multiplying the number of correctly solved subtask of “B”

by the reported piece-rate.

Earnings= number of solved subtasks by “B” * reported piece-rate

This means:
- The more subtasks “B” correctly solves, the higher are the earnings.
- The higher the reported piece-rate, the higher are the earnings.

Reporting a different than the designated piece-rate represents a violation of the rules. Your earnings will be calculated based on your reported not the designated piece-rate. You will, therefore, earn a piece-rate which does not correspond to your chosen task.

Randomization:

Random assignment determines which choice (chosen task of “A” or “B”) will be selected and therefore implemented by “B”.

a) In case “A’s” choice has been selected:
- “B” reads on the screen which task “A” has selected and which piece-rate “A” has reported.
- Additionally “B” receives the picked envelop of “A”. This envelope contains a sheet of paper where as a third bullet point (behind the password and the designated piece-rate) a number code appears.
- “B” is asked to type in this number code. Only afterwards, will “B” be able to proceed with the execution of the task.

For your information: By typing the number code “B” cannot reverse the chosen task of “A” nor “A’s” reported piece-rate. Inserting the password guarantees that “B” is informed about “A’s” chosen task, “A’s” reported piece-rate, and the designated piece-rate for “A’s” chosen task. It ensures

that “B” is ready to begin with the execution of the task.

b) In case “B’s” choice has been selected,
- “B” is asked to insert the number code which “B” finds written on the sheet contained in “B’s” picked envelop as a third bullet point (behind the password and the designated piece-rate).
- Only afterwards, will “B” be able to proceed with the execution of the task.

By typing the number code “B” cannot reverse “B’s” own chosen task or reported piece-rate. Inserting the password ensures that “B” is ready to begin with the execution of the task.

Task of member “A”:

While “B” is executing “A’s” or “B’s” chosen task, “A” is asked to fulfill another task. “A’s” performance in this task is not decisive for the earning of the two members of the pair. “A” will learn about the nature of the task directly on the screen.

Control questions:

Before the experiment starts and each participant is informed about his or her role, all participants are asked to answer some control question in order to ensure their understanding.

REFERENCES

Abeler, J., Nosenzo, D., and Raymond, C. (2018). Preferences for truth-telling. *Econometrica, Forthcoming*.

ACFE (2018). Report to the nations - 2018 global study on occupational fraud and abuse. Technical report. Available at SSRN: https://www.acfe.com/report-to-the-nations/2018.

Alan, S., Baydar, N., Boneva, T., Crossley, T. F., and Ertac, S. (2017). Transmission of risk preferences from mothers to daughters. *Journal of Economic Behavior & Organization*, 134:60–77.

Alan, S., Boneva, T., and Ertac, S. (2016). Ever failed, try again, succeed better: Results from a randomized educational intervention on grit. Technical report, HCEO Working Paper.

Alan, S. and Ertac, S. (2018). Fostering patience in the classroom: Results from a randomized educational intervention. *Journal of Political Economy, Forthcoming*.

Alaoui, L. and Fons-Rosen, C. (2016). Know when to fold 'em: The grit factor. Working Papers 899, Barcelona Graduate School of Economics.

Almlund, M., Duckworth, A. L., Heckman, J., and Kautz, T. (2011). Personality Psychology and Economics. In *Handbook of the Economics of Education*, volume 4, chapter 1, pages 1–181. Elsevier.

Almås, I., Cappelen, A., Salvanes, K. G., Sørensen, E., and Tungodden, B. (2016). Willingness to compete: Family matters. *Management Science*, 62(8):2149–2162.

Almås, I., Cappelen, A. W., Salvanes, K. G., Sørensen, E., and Tungodden, B. (2017). Fairness and Family Background. *Politics, Philosophy & Economics*, 16(2):117–131.

Almås, I., Cappelen, A. W., Sørensen, E. ., and Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science*, 328(5982):1176–1178.

Andreoni, J. and Sprenger, C. (2012). Estimating time preferences from convex budgets. *American Economic Review*, 102(7):3333–56.

Ariely, D., Kamenica, E., and Prelec, D. (2008). Man's search for meaning: The case of legos. *Journal of Economic Behavior & Organization*, 67(3-4):671–677.

Ashby, J. S. and Schoon, I. (2010). Career success: The role of teenage career aspirations, ambition value and gender in predicting adult social status and earnings. *Journal of Vocational Behavior*, 77(3):350 – 360.

Babcock, P., Bedard, K., Charness, G., Hartman, J., and Royer, H. (2015). Letting down the team? Social effects of team incentives. *Journal of the European Economic Association*, 13(5):841–870.

Bandiera, O., Barankay, I., and Rasul, I. (2005). Social preferences and the response to incentives: Evidence from personnel data. *The Quarterly Journal of Economics*, 120(3):917–962.

Bardsley, N. and Sausgruber, R. (2005). Conformity and reciprocity in public good provision. *Journal of Economic Psychology*, 26(5):664–681.

Bartling, B. and Fischbacher, U. (2011). Shifting the blame: On delegation and responsibility. *The Review of Economic Studies*, 79(1):67–87.

Bauer, M., Chytilová, J., and Pertold-Gebicka, B. (2014). Parental background and other-regarding preferences in children. *Experimental Economics*, 17(1):24–46.

Ben-Ner, A., List, J. A., Putterman, L., and Samek, A. (2017). Learned generosity? An artefactual field experiment with parents and their children. *Journal of Economic Behavior & Organization*, 143:28–44.

Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior*, 10(1):122–142.

Branas-Garza, P., Cobo-Reyes, R., Espinosa, M. P., Jiménez, N., Kovářík, J., and Ponti, G. (2010). Altruism and social integration. *Games and Economic Behavior*, 69(2):249–257.

Breyer, B. and Danner, D. (2015). Grit scale for perseverance and passion for long-term goals. *Zusammenstellung sozialwissenschaftlicher Items und Skalen*.

Brummelman, E., Thomaes, S., Nelemans, S. A., Orobio de Castro, B., and Bushman, B. J. (2015a). My child is god's gift to humanity: Development and validation of the parental overvaluation scale (POS). *Journal of Personality and Social Psychology*, 108(4):665.

Brummelman, E., Thomaes, S., Nelemans, S. A., Orobio de Castro, B., Overbeek, G., and Bushman, B. J. (2015b). Origins of narcissism in children. *Proceedings of the National Academy of Sciences*.

Burks, S., Carpenter, J., Götte, L., and Rustichini, A. (2012). Which measures of time preference best predict outcomes: Evidence from a large-scale field experiment. *Journal of Economic Behavior & Organization*, 84(1):308 – 320.

Burks, S. V., Lewis, C., Kivi, P. A., Wiener, A., Anderson, J. E., Götte, L., DeYoung, C. G., and Rustichini, A. (2015). Cognitive skills, personality, and economic preferences in collegiate success. *Journal of Economic Behavior & Organization*, 115:30 – 44.

Camerer, C. F. and Loewenstein, G. (2011). *Advances in Behavioral Economics*, chapter 1, Behavioral economics: Past, Present, Future, pages 3–51. Princeton University Press.

Cassar, L. (2018). Job mission as a substitute for monetary incentives: Benefits and limits. *Management Science*.

Castillo, M., Ferraro, P. J., Jordan, J. L., and Petrie, R. (2011). The today and tomorrow of kids: Time preferences and educational outcomes of children. *Journal of Public Economics*, 95(11):1377 – 1385. Special Issue: International Seminar for Public Economics on Normative Tax Theory.

Chandler, D. and Kapelner, A. (2013). Breaking monotony with meaning: Motivation in crowdsourcing markets. *Journal of Economic Behavior & Organization*, 90:123–133.

Charness, G. (2000). Responsibility and effort in an experimental labor market. *Journal of Economic Behavior & Organization*, 42(3):375–384.

Charness, G., Cobo-Reyes, R., Jiménez, N., Lacomba, J. A., and Lagos, F. (2012). The hidden advantage of delegation: Pareto improvements in a gift exchange game. *American Economic Review*, 102(5):2358–79.

Charness, G., Rigotti, L., and Rustichini, A. (2016). Social surplus determines cooperation rates in the one-shot prisoner's dilemma. *Games and Economic Behavior*, 100:113–124.

Christie, R. and Geis, F. L. (1970). *Machiavellianism*. Academic Press, Incorporated.

Cialdini, R. B., Reno, R. R., and Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of personality and social psychology*, 58(6):1015.

Cipriani, M., Giuliano, P., and Jeanne, O. (2013). Like mother like son? Experimental evidence on the transmission of values from parents to children. *Journal of Economic Behavior & Organization*, 90:100–111.

Crawford, V. P., Gneezy, U., and Rottenstreich, Y. (2008). The power of focal points is limited: Even minute payoff asymmetry may yield large coordination failures. *American Economic Review*, 98(4):1443–58.

Currarini, S., Jackson, M. O., and Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4):1003–1045.

Czermak, S., Feri, F., Glätzle-Rützler, D., and Sutter, M. (2016). How strategic are children and adolescents? Experimental evidence from normal-form games. *Journal of Economic Behavior & Organization*, 128:265 – 285.

Dal Bó, P., Foster, A., and Putterman, L. (2010). Institutions and behavior: Experimental evidence on the effects of democracy. *American Economic Review*, 100(5):2205–29.

Deci, E. L. and Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, 19(2):109–134.

Deckers, T., Falk, A., Kosse, F., Pinger, P. R., and Schildberg-Hörisch, H. (2017). Socio-Economic Status and Inequalities in Children's IQ and Economic Preferences. Technical report, IZA Discussion Paper No. 11158.

Deckers, T., Falk, A., Kosse, F., and Schildberg-Hörisch, H. (2015). How Does Socio-Economic Status Shape a Child's Personality? Technical report, IZA Discussion Paper.

Diaz, J., Arias, O., and Tudela, D. V. (2013). Does Perseverance Pay as Much as Being Smart?: The Returns to Cognitive and Non-cognitive Skills in urban Peru. Technical report, World Bank.

Diekmann, A., Przepiorka, W., and Rauhut, H. (2015). Lifting the veil of ignorance: An experiment on the contagiousness of norm violations. *Rationality and Society*, 27(3):309–333.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2011). The intergenerational transmission of risk and trust attitudes. *The Review of Economic Studies*, 79(2):645–677.

Duckworth, A. and Quinn, P. (2009). Development and validation of the short grit scale. *Journal of Personality Assessment*, 91(2):166–174. PMID: 19205937.

Duckworth, A. L., Peterson, C., Matthews, M. D., and Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6):1087–1101.

Eskreis-Winkler, L., Duckworth, A. L., Shulman, E. P., and Beal, S. (2014). The grit effect: Predicting retention in the military, the workplace, school and marriage. *Frontiers in psychology*, 5:36.

Falk, A. and Ichino, A. (2006). Clean evidence on peer effects. *Journal of Labor Economics*, 24(1):39–57.

Falk, A. and Kosfeld, M. (2006). The hidden costs of control. *American Economic Review*, 96(5):1611–1630.

Fehr, E., Bernhard, H., and Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, 454(7208):1079.

Fehr, E. and Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425(6960):785.

Fehr, E. and Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, 25(2):63–87.

Fehr, E. and Gächter, S. (1998). Reciprocity and economics: The economic implications of homo reciprocans. *European economic review*, 42(3-5):845–859.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868):137.

Fehrler, S. and Kosfeld, M. (2014). Pro-social missions and worker motivation: An experimental study. *Journal of Economic Behavior & Organization*, 100:99–110.

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

Fischbacher, U. and Föllmi-Heusi, F. (2013). Lies in disguise - an experimental study on cheating. *Journal of the European Economic Association*, 11(3):525–547.

Fischbacher, U. and Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review*, 100(1):541–56.

Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, 71(3):397–404.

Friedman, D. and Sunder, S. (1994). *Experimental methods: A primer for economists*. Cambridge University Press.

Gächter, S. and Riedl, A. (2005). Moral property rights in bargaining with infeasible claims. *Management Science*, 51(2):249–263.

Gächter, S. and Schulz, J. F. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature*, 531(7595):496.

Gagné, M. and Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, 26(4):331–362.

Gerhards, L. and Gravert, C. (2015). Grit trumps talent? An experimental approach. Working Papers in Economics 629, University of Gothenburg, Department of Economics.

Gill, D. and Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review*, 102(1):469–503.

Gino, F., Ayal, S., and Ariely, D. (2009). Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel. *Psychological Science*, 20(3):393–398.

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

Gneezy, U., Rockenbach, B., and Serra-Garcia, M. (2013). Measuring lying aversion. *Journal of Economic Behavior & Organization*, 93:293–300.

Golsteyn, B., Grönqvist, H., and Lindahl, L. (2013). Adolescent Time Preferences Predict Lifetime Outcomes. *The Economic Journal*, 124(580):F739–F761.

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

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of economic behavior & organization*, 3(4):367–388.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–482.

Herrmann, B., Thöni, C., and Gächter, S. (2008). Antisocial punishment across societies. *Science*, 319(5868):1362–1367.

Kahan, D. M. (1997). Social influence, social meaning, and deterrence. *Virginia Law Review*, pages 349–395.

Kajackaite, A. (2018). Lying about luck versus lying about performance. *Journal of Economic Behaviour & Organization*, 153:194–199.

Kaptein, M. (2004). Business codes of multinational firms: What do they say? *Journal of Business Ethics*, 50(1):13–31.

Katz, L. F. and Krueger, A. B. (2016). The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015. Working Paper 22667, National Bureau of Economic Research.

Kautz, T., Heckman, J. J., Diris, R., ter Weel, B., and Borghans, L. (2014). Fostering and measuring skills. Technical Report 110, OECD Education Working Papers.

Keizer, K., Lindenberg, S., and Steg, L. (2008). The spreading of disorder. *Science*, 322(5908):1681–1685.

Kerschbamer, R., Sutter, M., and Dulleck, U. (2017). How social preferences shape incentives in (experimental) markets for credence goods. *The Economic Journal*, 127(600):393–416.

Kocher, M. G., Cherry, T., Kroll, S., Netzer, R. J., and Sutter, M. (2008). Conditional cooperation on three continents. *Economics Letters*, 101(3):175–178.

Kocher, M. G., Martinsson, P., Myrseth, K. O. R., and Wollbrant, C. E. (2017). Strong, bold, and kind: Self-control and cooperation in social dilemmas. *Experimental Economics*, 20(1):44–69.

Lay, C. H. (1986). At last, my research article on procrastination. *Journal of Research in Personality*, 20(4):474 – 495.

Lazear, E. P. (2000). Performance pay and productivity. *American Economic Review*, 90(5):1346–1361.

Lergetporer, P., Angerer, S., Glätzle-Rützler, D., and Sutter, M. (2014). Third-party punishment increases cooperation in children through (misaligned) expectations and conditional cooperation. *Proceedings of the National Academy of Sciences*, pages 6916–6921.

Mas, A. and Moretti, E. (2009). Peers at work. *American Economic Review*, 99(1):112–45.

Mas, A. and Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, 107(12):3722–59.

Mooradian, T., Matzler, K., Uzelac, B., and Bauer, F. (2016). Perspiration and inspiration: Grit and innovativeness as antecedents of entrepreneurial success. *Journal of Economic Psychology*, 56:232 – 243.

Niederle, M. and Yestrumskas, A. H. (2008). Gender differences in seeking challenges: The role of institutions. Working Paper 13922, National Bureau of Economic Research.

Non, A. and Tempelaar, D. (2016). Time preferences, study effort, and academic performance. *Economics of Education Review*, 54:36 – 61.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314(5805):1560–1563.

Oexl, R. and Grossman, Z. J. (2013). Shifting the blame to a powerless intermediary. *Experimental Economics*, 16(3):306–312.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., and Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socio-economic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2(4):313–345. PMID: 26151971.

Ryan, R. M. and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1):68–78.

Sally, D. (1995). Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality and Society*, 7(1):58–92.

Shearer, B. (2004). Piece rates, fixed wages and incentives: Evidence from a field experiment. *The Review of Economic Studies*, 71(2):513–534.

Shepard, E. I., Clifton, T. J., and Kruse, D. (1996). Flexible work hours and productivity: Some evidence from the pharmaceutical industry. *Industrial Relations: A Journal of Economy and Society*, 35(1):123–139.

Sutter, M. (2009). Deception through telling the truth?! Experimental evidence from individuals and teams. *The Economic Journal*, 119(534):47–60.

Sutter, M., Feri, F., Glätzle-Rützler, D., Kocher, M. G., Martinsson, P., and Nordblom, K. (2018). Social preferences in childhood and adolescence. A large-scale experiment to estimate primary and secondary motivations. *Journal of Economic Behavior & Organization*, 146:16–30.

Sutter, M., Kocher, M. G., Glätzle-Rützler, D., and Trautmann, S. T. (2013). Impatience and Uncertainty: Experimental Decisions Predict Adolescents' Field Behavior. *American Economic Review*, 103(1):510–31.

Sutter, M., Yilmaz, L., and Oberauer, M. (2015). Delay of gratification and the role of defaults - An experiment with kindergarten children. *Economics Letters*, 137:21–24.

Tamassia, C. and Lennon, M. L. (2013). Technical Report of the Survey of Adult Skills (PIAAC). Technical report, OECD.

Tonin, M. and Vlassopoulos, M. (2015). Corporate philanthropy and productivity: Evidence from an online real effort experiment. *Management Science*, 61(8):1795–1811.

Wilson, J. Q. and Kelling, G. L. (1982). Broken windows. *Atlantic monthly*, 249(3):29–38.

Wolf, E. and Beblo, M. (2004). Does work time flexibility work? An empirical assessment of the efficiency effects for german firms. Technical report, SSRN.

CURRICULUM VITAE

PERSONAL DETAILS

Name	Anna Untertrifaller
Place of birth	Bozen (Italy)

EDUCATION

Sine 2017	Research fellow, Max-Planck-Institute for Research on Collective Goods, Bonn
Since 2014	Ph.D. student, University of Cologne
2014	Master in Economics, University of Verona, Italy
2012	Bachelor in Business and Economics, University of Verona, Italy