ESSAYS ON FAIRNESS, COORDINATION, AND DILIGENCE

Experimental Evidence From Children And Young Adults

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

2018

vorgelegt von
Mag. Claudia Elisabeth Zoller
aus Innsbruck, Österreich

Referent:	Prof. Dr. Matthias Sutter
Korreferentin:	Prof. Dr. Bettina Rockenbach
Tag der Promotion:	16.10.2018

To my family.

ACKNOWLEDGEMENTS

First, I would like to express my sincere gratitude to my supervisor Matthias Sutter, whose work on studies with children has sparked my interest in the field of experimental economics. I benefited greatly from his creativity, expertise, and patience, and he has been a substantial part in shaping my academic proficiency. His continuous support and guidance over the years is deeply appreciated.

I would also like to thank my secondary supervisor Bettina Rockenbach for helpful comments and for her support, and Oliver Gürtler for being part of my thesis committee.

My sincere thanks go to my co-authors Daniela Glätzle-Rützler and Ivan Soraperra for their valuable contributions to the papers and for sharing their expertise.

Special thanks go to Stefania Bortolotti, in whom I have not only found a great friend but a mentor, who always takes the time to help, discuss, and give feedback. I thank Anna Untertrifaller, who is not only the best office-mate one could hope for but a great friend who makes Cologne feel more like home. I especially thank Felix Kölle, Matthias Praxmarer, and Justin Buffat for sharing this experience with me. I feel very fortunate to have such wonderful colleagues both at the University of Cologne and at the MPI Bonn. I thank Suparee Boonmanunt, Markus Dertwinkel-Kalt, Lisa Dickmanns, Fabienne Fier, Uta Schier, Caroline Stein, Kristina Steinbrecher, Sebastian Tonke, Lukas Wenner, and Jarid Zimmermann at the University of Cologne and Zvonimir Bašić, Zita Coppinger, Zwetelina Iliewa, Sofia Monteiro, Shambhavi Priyam, Angelo Romano, Daniel Salicath, Ali Seyhun Saral, and Sebastian Schneider at the MPI for countless hours of insightful discussions and for all the fun times at and outside of the office.

Most importantly, I thank my parents and my partner Michael for their unfailing support and encouragement – without you I would not have made it this far. I dedicate this thesis to you and to my late grandmother Dr. Elisabeth Zoller – my namesake and a true economist.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Understanding economic preferences in human behavior is key in explaining economically relevant phenomena. Over recent years, a vast interest among economists has been to advance our understanding of how decisions are made. The influence of preferences on individual decision-making and decision-making in a social context has attracted a lot of attention, as economic preferences (such as i.e., time or risk preferences) have been shown to have wide-ranging implications for an individual's economic success (see e.g., Heckman and Smith, 1995; Fehr and Schmidt, 1999a; Bolton and Ockenfels, 2000; Charness and Rabin, 2002). Experimental economics has created effective methods to analyze human behavior by isolating decisions in a controlled environment. Tools in experimental economics allow for data elicitation in two ways: One setting is the laboratory experiment, where the experimenter is able to exercise tight control over incentives, information provided to subjects, and interactions between participants. Laboratory experiments use a standard subject pool of university students, testing game-theoretic models in a cost effective way. Field or lab-in-the-field experiments take place in a natural environment with non-standard subject pools, bridging findings from laboratory experiments and naturally occurring data (see e.g., Harrison and List, 2004; Levitt and List, 2007; Falk and Heckman, 2009; Levitt and List, 2009).

This thesis builds on two separate pillars, contributing to the literature on economic preferences by eliciting data in a laboratory and in lab-in-the-field experiments. The first part of this thesis analyzes preferences for the support of the welfare state. Increasing inequality provokes a lively debate about what can be considered a fair distribution of income. Preferences for redistribution typically depend on the perceived causes of income inequality and vary in the extent of the support for redistributive policies. Chapter 2 investigates the influence of potential immoral self-serving behavior on distributional preferences, where the source of the inequality is either determined by luck or cheating.

The second part of this thesis considers individual decision-making in children and adolescents. Chapters 3, 4, and 5 advance knowledge in the understanding of how behavior and economic preferences evolve with age by analyzing decision-making in children and adolescents in lab-in-the-field experiments. Chapter 3 provides an overview of the experimental literature with children, giving detailed insights into common behavioral patterns and the development of economic preferences with age. Chapter 4 investigates a strategic setting with children and adolescents in the context of a coordination game, measuring the influence of age and strategic uncertainty on efficient play. Chapter 5 focuses on the individual decision-making component in early childhood, analyzing effort provision and time preferences under either exogenously imposed or endogenously set time schedules.[1]

OVERVIEW AND MAIN FINDINGS

Chapter 2 (*Too lucky to be true – Fairness views under the shadow of cheating*) is joint work with Stefania Bortolotti, Ivan Soraperra, and Matthias Sutter.[2] We study how fairness views and the extent of redistribution are affected by a hitherto overlooked, but relevant factor: immoral self-serving behavior that can lead to increased inequality. We focus on situations in which the rich have potentially acquired their fortunes by means of cheating. In a laboratory experiment, we let third parties redistribute resources between two stakeholders who could earn money either by choosing a safe amount or by engaging in a risky, but potentially more profitable, investment. In one treatment, the outcome of the risky investment is determined by a random move, while in another treatment stakeholders can cheat to obtain the more profitable outcome. Although third parties cannot verify cheating, we find that the mere suspicion of cheating changes

[1] All authors contributed to each chapter to an equal extent.

[2] We gratefully acknowledge the financial support from the C-SEB Start-up Grant.

fairness views of third parties considerably and leads to a strong polarization. When cheating opportunities are present, the share of subjects redistributing money from rich to poor stakeholders triples and becomes as large as the fraction of libertarians – i.e., participants who never redistribute. Without cheating opportunities, libertarian fairness views dominate, while egalitarian views are much less prevalent. These results indicate that fairness views and attitudes towards redistribution change significantly when people believe that income inequality is the result of cheating by the rich.

Chapter 3 (*Economic behavior of children and adolescents – A first survey of experimental economics results*) is joint work with Daniela Glätzle-Rützler and Matthias Sutter. In recent years economic literature focusing on the decision-making process of children has progressed rapidly. The investigation of the development of economic preferences throughout childhood and adolescence identifies behavioral patterns, providing important insights for policy makers. We give a detailed overview of experimental studies in the field of economics conducted with children and adolescents, highlighting the development of behavioral findings by age cohorts, and investigating how interventions foster behavior. Children are already capable of making somewhat rational decisions and of drawing correct inferences about the partner's strategy, providing the basic requirements for a sophisticated decision-making process. Early childhood is mostly dominated by very selfish behavior, as children aim to maximize their own utility by keeping payoffs to themselves. As children grow older, they do not only display increasingly more patient and less risk-seeking behavior, but develop stronger other-regarding preferences. The majority of elementary school-aged children prefer the egalitarian choice, showing strong aversion to unequal payoff distributions even when the inequality is based on varying levels of exerted effort. In adolescence concerns for efficiency and social-welfare form, with subjects making increasingly more pro-social and altruistic choices. Differences by gender are especially prevalent when regarding competitive preferences, as females of any age group are less willing to compete in a tournament setting.

Chapter 4 (*Coordination games played by children and teenagers – On the influence of age, group size and incentives*) is joint work with Daniela Glätzle-Rützler and Matthias Sutter. We investigate how increasing age, varying incentives, and group size affect the ability to coordinate on the efficient equilibrium. Over 800 children and adolescents, aged nine to eighteen, played two coordination games – a stag-hunt game – where the group size differs between two players or five players. Incentives are varied between-subjects, making either the efficient equilibrium more profitable or reducing risk by increasing the payoff in case of coordination failure. We find that coordination on the efficient equilibrium becomes less likely with increasing age, unless it is exceedingly profitable to do so. Coordination with one partner yields more efficient outcomes compared to when children have to coordinate in groups of five. Hence, children are able to account for strategic uncertainty when coordinating in large groups. This concern is more pronounced in older children as they choose the efficient strategy less frequently in large groups, which is driven by pessimistic beliefs about the other players' choices. Our results are largely in line with findings on incentives in coordination games in adult subject pools, and provide deeper insights into the development of coordination strategies with age.

Chapter 5 (*Busy little bees – An experiment on diligence and endogenous time scheduling in early childhood*) is joint work with Matthias Sutter and Anna Untertrifaller. In an experiment with over 400 children between the ages of three and six, we study the development of diligent behavior and how it is influenced by endogenous time scheduling. Displaying high levels of diligent behavior has been found to positively affect educational attainment and hence is an important determinant of economic success. To measure diligence children worked on a real effort task, collecting tiny yellow beads from a bowl filled with multicolored beads. Children were instructed to spend as much time and effort on the task as they wanted to. In a treatment variation a subset of children was able to decide when to do the real effort task – today or tomorrow. We find that children who postpone the task to the next day exert substantially less effort, displaying lower levels of diligence than children who choose to (or are told to) do the task right away. While procrastination behavior is more profound in the younger age cohort of three and four year-olds, procrastinators of all ages exert less effort. Children who are willing to challenge themselves in an unrelated task are furthermore more likely to show more diligent behavior in the real effort task. These results suggest that with age children become more proficient in their self-management.

CHAPTER 2

TOO LUCKY TO BE TRUE – FAIRNESS VIEWS UNDER THE SHADOW OF CHEATING

Abstract

Income inequalities within societies are often associated with cheating on the side of the rich. We study how fairness views and preferences for redistribution are affected when cheating may be the cause, but it cannot be proven. In our experiment, we let third parties redistribute income between a rich and a poor stakeholder. In one treatment, income inequality was only due to luck, while in two others rich stakeholders might have cheated. The mere suspicion of cheating changes third parties' fairness views considerably and leads to a strong polarization that is even more pronounced when cheating generates negative externalities.

This chapter is joint work with Stefania Bortolotti, Ivan Soraperra, and Matthias Sutter.

2.1 INTRODUCTION

The unequal distribution of income within countries has become a major issue in the academic as well as the public debate in recent years (Corak, 2013; Chetty et al., 2014; Piketty, 2014). While one side of the debate focuses on the reasons for the widening gap between haves and have-nots, another side addresses the important question of what would constitute a fair distribution of income between the rich and the poor (Cappelen et al., 2013). The answer to the latter question hinges to a large degree on the factors that generated the inequality in the first place, meaning that both sides of the debate are necessarily intertwined.

It has been shown that fairness views with respect to redistribution of income depend on whether or not income inequalities have been caused by differences in effort and hard work (Alesina and Angeletos, 2005), or whether someone can be held accountable for one's (mis)fortune (Konow, 2000; Cappelen et al., 2013; Möllerström et al., 2015; Akbaş et al., 2016; Lefgren et al., 2016; Tinghög et al., 2017). Judgments about the fairness of an income distribution are also highly sensitive to the available information about, and the subjective perception of, the income distribution (Kuziemko et al., 2015). Moreover, fairness views about the preferred extent of redistribution within a society have been found to be affected by concerns about procedural fairness or efficiency (Bolton et al., 2005; Balafoutas et al., 2013; Almås et al., 2016; Cassar and Klein, 2016), cultural background or political orientation (Rey-Biel et al., 2011; Almås et al., 2016; Cappelen et al., 2016; Konow et al., 2016; Klimm, 2018), or conflicts between social classes (Fehr, 2018).

It is noteworthy that none of the studies mentioned above has looked at how fairness views are affected if suspicion about the integrity of one's fortunes arises. This is surprising given the evidence from leaks about offshore financial activities such as the "Panama papers", showing that tax evasion rises with wealth (Alstadsæter et al., 2017), thus increasing the income inequality within societies in favor of the rich. Another prominent example where immoral and illegal behavior was the source of undeserved wealth was the common practice at banks such as Wells Fargo to open up fake bank accounts to meet the monthly targets of sales personnel (`http://fortune.com/2017/04/13/wells-fargo-report-earnings/`). Seen from a broader, macro-level perspective, there seems to be even a positive correlation between inequality in a given country – measured by the Gini coefficient – and the perceived level of cheating and corruption (see country-level evidence supporting this relationship in Figure A-1 in the Appendix). Yet, at the macro-level, this correlation could be influenced by several country-specific factors, e.g., the quality of formal institutions, the level of trust, or cultural differences. These factors make it difficult to assess properly how the perception of cheating affects fairness views and the desire to redistribute between rich and poor. For this reason, we are going to present a carefully controlled experimental study in which we have rich and poor subjects, where in some treatments the rich may, but need not, have acquired their income by cheating. We study how unbiased third parties redistribute income from rich to poor subjects in our experiment. We are particularly interested in situations in which it is untraceable for third parties whether the source of a person's high income may have originated from cheating or from honest behavior. We hence talk about redistribution *under the shadow of cheating* and not in the obvious presence of cheating. To our knowledge, we are the first to test the effect the shadow of cheating casts on fairness views, a topic of great relevance for the design of the welfare state and incentive schemes.

In our experiment, we let stakeholders choose between a safe option and a potentially more profitable, but risky investment. This choice is intended to capture many real-world situations, ranging from job choices to health or farming decisions, where subjects have to trade off risk and expected return. For instance, one could think of people having to select either into a low-pay, but secure job, or a potentially lucrative, but highly risky sector.

We exogenously manipulate the availability of cheating opportunities and hence the causes that can create income disparities among stakeholders. In one treatment – called the *Nature* treatment, which follows Cappelen et al. (2013) – the outcome of the risky investment is determined by a random computer draw that yields either a high or a low income for the stakeholder with 50% probability, respectively. The safe option, in contrast, yields an intermediate income for sure. In the other treatments – *Self-Report* and *Externality* – stakeholders face the same choice between a safe option and a risky investment. Here, the investment's outcome is resolved by the stakeholders themselves by flipping a coin and self-reporting the outcome. They are explicitly asked to report the outcome truthfully, which yields either the high or low income with 50% probability. Our request is, nonetheless, non-verifiable – mimicking many real-world situations in which monitoring is too costly.

Both *Self-Report* and *Externality* captures situations where cheating is possible.[1] Yet, both treatments differ from one another with respect to the costs cheating imposes on other players. *Self-Report* is a very mild manipulation as no harm is done to anybody in case of cheating. Misreporting the outcome of the lottery is a violation of a rule but such a rule is not meant to protect the other parties. The *Externality* treatment takes cheating a step forward: a stakeholder who cheats automatically appropriates resources that were allocated to another (idle) player. This mimics situations where the perpetrators benefit from their dishonest action at the expense of other people's income. Misreporting the high income here is a violation of a more important rule aimed at protecting vulnerable parties. Consider again the example of Wells Fargo where incentives and sales targets for employees led to the opening of fake bank accounts, engaging in harmful behavior towards the bank and its investors. This kind of illegal and immoral behavior led to high costs, imposing consequences not only on the offenders, but also on uninvolved parties.

After the stakeholders' decisions, we let third parties – henceforth called *spectators* – redistribute the total sum of earnings within a pair of stake-

[1] This aspect of our design – i.e., the possibility for stakeholders to lie – links our project to the flourishing literature on deception and cheating (Gneezy, 2005; Fischbacher and Föllmi-Heusi, 2013; Gächter and Schulz, 2016; Abeler et al., 2018). Virtually all papers on this topic test how cheating depends on different contextual cues and conditions, including the structure of incentives (Conrads et al., 2013), loss avoidance (Grolleau et al., 2016), the nature of the task (Kajackaite, 2016), the costs associated with cheating (Gneezy et al., 2018), or the role of collaboration (Weisel and Shalvi, 2015). We take a completely different stand on the problem by not focusing on the *causes* of cheating, but rather on the *consequences of dishonesty on fairness views and distributional preferences of unaffected bystanders.*

holders. Spectators in the *Self-Report* and *Externality* treatment, however, cannot identify whether a high income of a stakeholder was the consequence of either a lucky coin toss or of misreporting the true (i.e., low) outcome of the coin toss. Although spectators can form expectations about the likelihood of cheating, there is no certainty. This ambiguity creates a difficult challenge for those spectators who want to eliminate income inequality whenever the high income resulted from dishonest behavior, but simultaneously want to refrain from redistribution if a stakeholder's high income was righteously acquired.

A recent paper by Cappelen et al. (2017) is related to ours as they also consider redistribution in groups composed of honest and dishonest stakeholders.[2] The authors report that spectators (i.e., third parties who decide upon redistribution) care more about rewarding the honest group members – thus accepting that dishonest group members also benefit from more redistribution – rather than being concerned with punishing dishonest group members at the expense of honest ones. Our experiment departs from Cappelen et al. (2017) in three main aspects. First, we consider situations in which cheating is on the side of the rich, while Cappelen et al. (2017) investigates the case in which cheating is on the side of the poor. We concentrate on cheating by the rich because the sentiment against the so-called elites in many countries is fueled by the poorer people's concerns that the rich have achieved their fortunes also through dishonest means. Second, we implement ambiguity about whether income inequalities are the consequence of cheating, while in Cappelen et al. (2017) there was full disclosure of cheating behavior at the group level, leaving no room for ambiguity as to whether cheating was involved or not. Finally, we enlarge the choice set by having a safe option, while in Cappelen et al. (2017) all group members were asked to work on a real effort task and were then paid according to a lottery system. This implies that a common ground of both papers is to consider a situation where two (groups of) stakeholders choose a risky option and one earns the high income and the other the low income;

[2]Their paper and ours were developed at the same time and independently of each other.

however, our set-up allows us to extend such a paradigm to include cases in which the stakeholders make different choices in the first place (i.e., one chooses a safe and the other a risky option), which is intended to mimic different types of behavior in the field, such as in educational or professional choices.

We report three main findings in ascending order of importance. First, in line with previous evidence about the incidence of dishonesty, we find cheating among stakeholders in the *Self-Report* and *Externality* treatment, but not to the full extent (Gneezy, 2005; Fischbacher and Föllmi-Heusi, 2013; Abeler et al., 2018; Gneezy et al., 2018). In line with stakeholders' behavior, we also provide evidence that spectators correctly expect cheating in this set-up.

Second, spectators are more likely to reduce inequality in the *Self-Report* and the *Externality* treatment than in the *Nature* treatment in cases where a stakeholder with high income – who might have cheated – is paired with a stakeholder who chose the safe option with the intermediate income. Interestingly, we do not observe any kind of *reward* for genuinely honest stakeholders who truthfully indicate a low income (through reporting an unlucky coin toss).

Third, and most importantly, we document a strong and significant shift in fairness views across treatments. We use a discrete choice random utility model to estimate three types of spectators (Cappelen et al., 2007, 2013): *Libertarians* are spectators who never redistribute, independent of the degree of income inequality between stakeholders. *Egalitarians* always redistribute resources equally, while *Choice Egalitarians* redistribute only among stakeholders who chose the risky investment, but do not redistribute if one stakeholder chose the risky option and the other the safe option.

The share of *Libertarians* is practically the same across all treatments, capturing about 40% of spectators. However, the other types differ sharply across treatments. The share of *Egalitarians* is three to four times larger in *Self-Report* and *Externality* than in *Nature*. Hence, if income inequalities *might* have been generated from cheating behavior, the fraction of egalitarian fairness views increases substantially. At the same time, *Choice Egalitarians* are much more frequent in *Nature* than in the treatments where the shadow of cheating prevails. Overall, therefore, the shadow of cheating leads to a strong polarization of fairness views where two diametrically opposing views – *Libertarians* who never redistribute and *Egalitarians* who always redistribute equally – become about equally strong in the population of spectators. As a consequence, the shadow of cheating might increase social tensions with respect to the question of how to deal with income inequalities. Interestingly, the polarization already arises in a situation where cheating has no negative externality on another, powerless third party. Yet, it is even more pronounced when cheating imposes such negative externalities, as in the *Externality* treatment. The latter treatment reflects a typical feature of real life as cheating often harms another party.

The remainder of the paper is organized as follows. Section 2.2 describes the experimental design and data collection. Section 2.3 presents the results for both stakeholders and spectators. Section 2.4 reports the results of a discrete choice random utility model estimating fairness views. Section 2.5 concludes.

2.2 EXPERIMENTAL DESIGN

Our design builds on Cappelen et al. (2013). We have two types of players – stakeholders and spectators – and two stages. We start by presenting the details of the first stage, in which stakeholders made their decisions. After that, we introduce the three experimental treatments. We then explain the second stage where spectators made a series of redistributive decisions. Finally, we describe the experimental procedures.

2.2.1 STAGE 1: STAKEHOLDERS' RISK-TAKING DECISIONS

Each stakeholder independently had to make five ordered decisions between a safe and a risky option, as shown in Table 21. The risky option paid either a high income of 800 tokens or a low income of 0 tokens, each with 50% probability. While the risky option remained fixed in all five decisions, the intermediate income paid by the safe option varied across decisions.

This amount increased linearly from 100 tokens in the first decision to 500 tokens in the fifth decision, making the safe option more attractive as the stakeholders proceeded through the five decisions. After all five decisions had been made, the risky option – if chosen – was resolved for each decision separately. The resolution of the risky option and its consequences depended upon the experimental treatment.

Table 21: Stakeholders' decisions

Decision	**Safe option**	**Risky option**	p
# 1	100	800 or 0	$p = .50$
# 2	200	800 or 0	$p = .50$
# 3	300	800 or 0	$p = .50$
# 4	400	800 or 0	$p = .50$
# 5	500	800 or 0	$p = .50$

Notes: p is the probability of earning the high income in the risky option. All amounts are expressed in tokens. Tokens are converted at the rate of 1 Euro = 300 tokens. In the *Nature* treatment, the outcome of the risky option is determined by a random draw of the computer. In the *Self-Report* and *Externality* treatments, the outcome is determined by a self-reported coin toss where Heads yields the high income of 800 tokens, and Tails the low income of 0 tokens.

2.2.2 EXPERIMENTAL TREATMENTS

We ran three between-subjects treatments: *Nature*, *Self-Report*, and *Externality*. The first two – *Nature* and *Self-Report* – differ only in the way in which the outcome of the risky option was determined. The *Externality* treatment builds on *Self-Report* and imposes negative consequences on another player in case of cheating.

- In the *Nature* treatment, the outcome of the risky option was determined by a random draw performed by the computer. The probability of the high or low income was 50% for each level. The outcome of each random draw was shown to the stakeholder after all decisions between the safe and risky option had been made.

- In the *Self-Report* treatment, the outcome of the risky option was determined by a self-reported coin toss. The coin tosses had to be performed only after all five decisions had been made. Stakeholders were asked to get a coin or to use an online website (`justflipacoin.com`) to flip a coin for each decision in which they had chosen the risky option. We explicitly requested stakeholders to report the results of the coin tosses truthfully (see Instructions in the Appendix). Misreporting was hence a clear violation of the rules. Under the assumption of honest reporting, our procedure guarantees the same likelihood (of 50%) of earning the high income across treatments, conditional on choosing the risky option. However, our request for honest reporting could not be enforced as there was no possibility to detect lies at the individual level (for further details, see the experimental procedure below). This set-up mimics situations in which rules are not enforceable or the cost of monitoring is too large.

- In the *Externality* treatment, the outcome of the risky option was determined by a self-reported coin toss as in *Self-Report*. This time, we built in an externality of cheating. Each stakeholder was matched with an idle player whose earnings *might* depend on the choices of the stakeholder. The idle player was informed about the rules but was not eligible to make any decision. Stakeholders knew that the idle player received a bonus payment of either 0 or 800 tokens which was determined as follows:

  (i) if the matched stakeholder chose the *safe* option, the earnings of the idle player were independently determined by a random draw of the computer and both outcomes (0 and 800 tokens) were equally likely;

  (ii) if the matched stakeholder chose the *risky* option, the earnings of the idle player were negatively correlated with the ones of the stakeholder. This means that if the stakeholder reported a winning coin toss, the matched idle player received the low income of 0 tokens. If the stakeholder reported a losing coin toss, the matched idle player received the high income of 800 tokens.

While the latter treatments – *Self-Report* and *Externality* – both allow for cheating, it is important to remark the differences between the two. In the *Self-Report* treatment, one might see cheating as a rather *innocent* deviation from the rules, as it does not entail any loss for any other player. In the *Externality* treatment, falsely reporting a winning coin toss yields a high income for the stakeholder at the expense of the idle player's earnings. In this case, stakeholders are imposing a negative externality on someone to (illegitimately) seek their own profit.[3]

Note that in all treatments the rules used to resolve the outcome of the risky option were common knowledge from the beginning of the experiment – i.e., before stakeholders made any decision. In line with Cappelen et al. (2013), we also informed stakeholders at the beginning of the experiment that the study comprised two stages. For comparability, we use the same wording as in the reference paper (Cappelen et al., 2013): "Stage 2 of the experiment concerns the distribution of earnings from Stage 1. Details of the second stage will be provided after the first stage is complete." Only at the end of a session were stakeholders informed about the rules of stage 2.

2.2.3 STAGE 2: SPECTATORS' REDISTRIBUTION DECISIONS

In stage 2, spectators decided how to redistribute the sum of earnings within a pair of stakeholders. For each pair, the spectator was informed about the stakeholders' choices in each decision (see Table 21) and of the outcomes if the risky option had been chosen. In the *Externality* treatment,

[3]It is not impossible to imagine that a fair minded stakeholder might want to choose the risky option and purposefully report a losing coin toss to ensure the high income for the idle player. A stakeholder with very high costs of cheating and high sensitivity to inequality might choose this option. However, in a Fehr and Schmidt (1999b) framework, a stakeholder would never choose such an option, unless the reputation cost of being mistakenly associated as a cheater is larger or equal to the high income itself. Yet in this case a stakeholder might simply choose the safe option. Downward cheating to favor the idle player seems unlikely, also in light of the fact that the idle players are not especially deserving, as they are drawn from the same population and do not make any decision.

for stakeholders who chose the risky option, the earnings of the idle player were reported. If the stakeholder chose the safe option, the consequences for the idle player read: "outcome determined by the computer." No further information about the exact amount for the idle player was given in this case. When redistributing money between the two stakeholders (not between the idle players with which the two stakeholders were paired), spectators could redistribute the sum of the earnings of the two stakeholders from stage 1 in steps of 25 tokens. The payment for spectators themselves was a fixed amount and they were not affected by the stakeholders' decisions, and hence had no material self-interest at stake. This avoids any kind of personal self-serving bias on the part of spectators and allows us to elicit impartial and unbiased fairness views (Konow, 2000, 2009).

In the *Nature* treatment, spectators were informed that the mechanism used to determine the outcome of the risky option was a random draw performed by the computer, yielding the high or low income with equal probability. In the *Self-Report* and *Externality* treatment, instead, they were informed that stakeholders had to self-report the outcome of a coin toss to resolve the risky option. In the *Externality* treatment, spectators were also informed about the idle players and how their earnings were determined.

Each spectator was exposed to only one treatment and had to make 20 redistribution decisions (see Table A-1 in the Appendix for further details). One of these redistribution choices was payoff-relevant for a pair of stakeholders, but spectators were not informed which one was relevant.

2.2.4 EXPERIMENTAL PROCEDURES

STAKEHOLDERS. Stakeholders were recruited via Amazon Mechanical Turk (MTurk, henceforth) using the behavioral research platform TurkPrime (Litman et al., 2016).[4] For our study, we recruited a total of 600 online par-

[4]MTurk has gained momentum among social scientists and it is increasingly regarded as a valid alternative to other data collection techniques, with over 1,000 peer-reviewed papers relying on the platform (Litman et al., 2016). MTurk participants – often referred to as "workers" – represent a massive dataset of potential participants from a wide range of countries and with a diverse backgrounds. Monetary incentives for MTurk workers are often lower (at least in absolute terms, much less so in relative terms) than in the

ticipants - 120 in *Nature* and 240 each in *Self-Report*[5] and in *Externality*. In *Externality*, 120 online participants were assigned to the role of stakeholders and the other 120 to the role of idle players. The latter were asked to read the instructions given to the stakeholders and to answer the same control questions. They were made aware of how their potential bonus payment was generated but had no decision to make.

Participation was restricted to subjects from the U.S. with a high completion rate to minimize attrition.[6] Decisions were collected via SoSci (Leiner, 2014). Only participants who were able to answer all control questions correctly were allowed to participate. After two incorrect trials, stakeholders were automatically excluded from the study and were prevented from re-taking it. The stakeholders' average payment was about $2, including a $0.60 participation fee. The task lasted on average 8 minutes (implying an average hourly rate of about $15, which is comparable to many laboratory experiments). Idle players spent on average 7 minutes until completion, and earned on average $1.20, including a $.30 participation fee.

We believe the task in Stage 1 is particularly well suited for MTurk for two main reasons. First, the task is extremely simple and short, hence reducing potential concerns about understanding and concentration. Second, conducting the experiment on MTurk grants a degree of privacy to participants that would be difficult to achieve in the lab. Stakeholders were identified by a code and they completed their assignment over the internet from home or a place of their choice. Hence, there was no possibility for the experimenter to observe the result of the coin toss used in the *Self-Report* and *Externality* treatments to determine the outcome of the risky option. Given the complete separation between participants and

laboratory; however, there is evidence that reduced incentives have little or no effect on behavior (Buhrmester et al., 2011; Horton et al., 2011; Litman et al., 2015).

[5] Initially, we had 120 online stakeholders in *Self-Report*, as in the other two treatments. After feedback from seminar participants, we added 120 more participants in *Self-Report* because we wanted to collect beliefs among spectators. That also required additional data collection of stakeholders on MTurk.

[6] We recruited only experienced online workers. All of them had taken part in at least 50 previous assignments and had successfully completed at least 95% of these assignments. The average completion rate was 97.5%.

experimenter, stakeholders could easily infer that the experimenter had no way to detect cheating.

SPECTATORS. We recruited 237 students from the University of Cologne to act as spectators. Two sessions, with a total of 57 subjects, were assigned to the *Nature* treatment.[7] Four sessions, with a total of 120 subjects, were assigned to the *Self-Report* treatment. Two further sessions, with a total of 60 participants, were assigned to the *Externality* treatment. All sessions were conducted at the Cologne Laboratory for Economic Research (CLER) a few days after collecting data on MTurk. Subjects were recruited using ORSEE (Greiner, 2015) and the experiment was programmed using z-Tree (Fischbacher, 2007). Upon arrival, subjects were randomly assigned to a cubicle and no form of communication was allowed. A paper copy of the instructions was distributed to spectators and instructions were read aloud to assure common knowledge (see Appendix). Spectators could proceed to the proper experiment only after having answered all control questions correctly. Socio-demographic characteristics and personality traits (HEXACO Personality Inventory-Revised, Ashton and Lee 2009) were collected at the end of the experiment. Spectators were paid a fixed amount of €10 for the redistribution part, including a show-up fee of €4. The average session lasted about 45 minutes.

In two of the four *Self-Report* sessions (60 spectators) and in both *Externality* sessions (60 spectators), we additionally elicited beliefs and risk aversion. After making their redistribution choices, spectators were asked to answer the two following questions about the stakeholders:

- What is the percentage of participants in the online assignment that chose the risky option?
- Consider now only the online participants who have chosen the risky option: what is the percentage of participants who reported Heads?

[7]Due to a low show-up rate in one *Nature* session, we have only 57 spectators in this treatment. The number of pairs of stakeholders from MTurk was instead 60. The three extra-pairs were paid exactly the amount they had earned in Stage 1, as if there was no redistribution.

Please recall that Heads yielded an income of 800 tokens and Tails 0 tokens.

For the sake of simplicity, we elicited beliefs only for a safe level of 300 tokens. Beliefs were incentivized with a stepwise quadratic scoring rule (see Instructions in the Appendix) and six randomly selected spectators per session – of 30 subjects each – were paid based on one of the two questions.

To elicit a spectator's risk aversion (as a control variable for the redistribution choices), we followed the task proposed by Eckel and Grossman (2008). Spectators were presented with five options, of which they had to pick one. In each option, there was a 50% chance of a low payoff and a 50% chance of a high payoff. The low and high payoffs changed for each option. Higher expected payoffs were associated with higher risk. One randomly selected spectator per session was paid for this task. On average, spectators earned additionally €4 from the belief-elicitation and risk task.

2.3 RESULTS

In this section, we first present the results for the stakeholders' behavior. Then we continue with the discussion of redistribution patterns among spectators.

2.3.1 STAKEHOLDERS' RISK-TAKING AND CHEATING BEHAVIOR

The bars in Figure 21 show the relative frequencies with which stakeholders choose the risky option, conditional on the income from the safe option (ranging from 100 to 500 tokens) and on the treatment (left for *Nature*, right for *Self-Report*, bottom for *Externality*). We observe a clear downward trend in the relative frequencies of choosing the risky option in all treatments, dropping from 75-79% for a safe income of 100 tokens to 17-27% for a safe income of 500 tokens. A series of χ^2 tests fail to reveal any significant difference in risk-taking between *Nature* and *Self-Report* as well as between *Nature* and *Externality* for any safe income level (the p-values ranges from

$p = 0.122$ to $p = 0.875$).[8] For the comparison between *Self-Report* and *Externality*, χ^2 tests reveal a significant difference only for the safe income level of 500 ($p = .042$), while the other four comparisons are insignificant. The fraction of stakeholders who always choose the safe option is also similar across all three treatments, with 17% in *Nature*, 20% in *Self-Report* and 18% in *Externality.* Moreover, stakeholders display a high degree of consistency in their choices in all treatments, as less than 7% switch more than once between the risky option and the safe option.

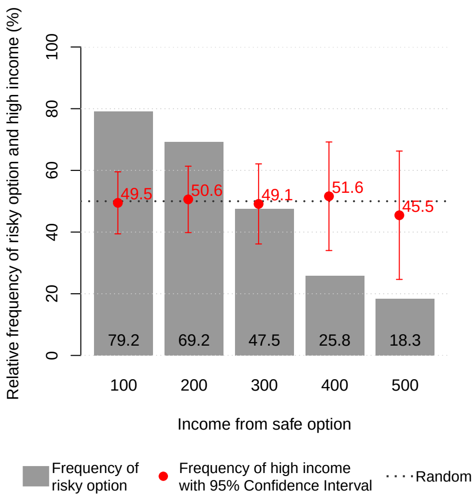
Result 1. *The relative frequency of choosing the risky option is not significantly different between Nature and the other two treatments. In all treatments, risk-taking drops as the income from the safe option increases.*

Figure 21 also shows the relative frequency of getting the high income from the risky option (see circles for averages and whiskers for confidence intervals). On the left-hand side, we see that in the *Nature* treatment this relative frequency is not significantly different from 50%, due to the fact that the outcome of the risky option was determined by a random computer draw. On the right-hand side of Figure 21, we note instead that in the *Self-Report* treatment stakeholders report having been lucky in their coin toss significantly more often than chance would predict (see Confidence Intervals in Figure 21). In fact, conditional on choosing the risky option, they claim the high income in 72% of cases. Among all stakeholders in *Self-Report*, 32% declared having been lucky in all instances where they chose the risky option. In the bottom part of Figure 21, we can observe that the behavior in the *Externality* treatment is remarkably similar to *Self-Report.* Among the stakeholders who chose the risky option in *Externality*, 76% claimed the

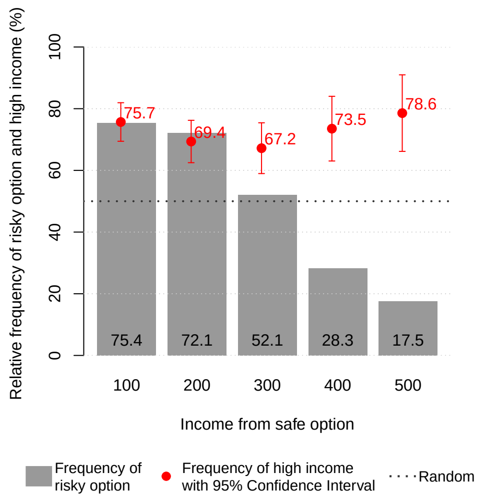
[8] It is interesting to note that the possibility to report a favorable outcome at one's discretion does not induce a change in risk-taking behavior in *Self-Report* and *Externality.* Compared to *Nature*, one would think that stakeholders in the two treatments with cheating opportunities only switch from the safe option to the risky one, but not vice versa. However, if subjects have a preference for being seen as honest, even risk-lovers could prefer to choose the safe option to avoid looking dishonest when reporting a lucky draw. In line with the evidence by Abeler et al. (2018) and Gneezy et al. (2018), our results suggest that direct costs of lying and reputation concerns are not negligible for a sufficiently large fraction of stakeholders, which could explain that the relative frequency of choosing the risky option does not differ between treatments.

Figure 21: Relative frequency of risky choices and high income

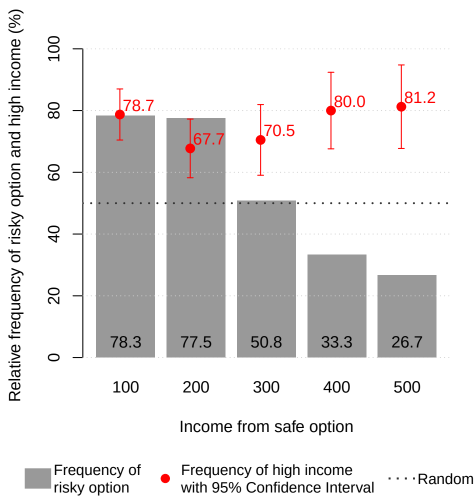
(a) *Nature*



(b) *Self-Report*



(c) *Externality*



high income (significantly more than predicted by a fair coin) and 35% of the stakeholders report to always have been lucky. This seems to suggest that the presence of negative externalities did not deter stakeholders from cheating.

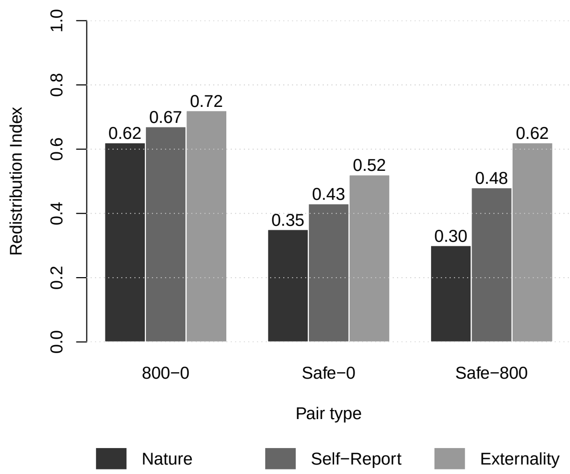
Result 2. *Many stakeholders cheat in the Self-Report and Externality treatments as the observed fraction of stakeholders reporting the high income from the risky option is significantly larger than 50% in these two treatments. Yet, stakeholders do not cheat to the full extent, as in roughly one quarter of the cases they report the low income (of zero tokens) when they could have easily claimed the high income.*

2.3.2 SPECTATORS' REDISTRIBUTION DECISIONS

Our 237 spectators made a total of 4,740 redistribution decisions. In 316 cases, there was nothing to redistribute because both stakeholders had earned the low income of zero tokens from the risky option. In addition, there are 1,133 cases where both stakeholders had the same positive income (either by having chosen the same safe option or by having earned the high income from the risky option). In virtually all of these instances (97.3%), spectators did not redistribute any income from one stakeholder to the other, as they had the same income to begin with. Therefore, we have a total of 3,291 (out of 4,740) cases with income inequality between the two stakeholders. In the majority of these cases (59%), spectators modified the initial distribution of earnings and their intervention was almost always aimed at reducing disparities. Only in less than 5% of cases (169 in total) did they increase inequality.

For the cases with strictly positive inequality, we can distinguish between three types of pairs in order to provide a more detailed analysis of redistributive behavior. In the first pair, henceforth called *800-0*, both stakeholders chose the risky option, but only one stakeholder earned the high income of 800 tokens, while the other earned the low income of zero tokens. The second pair is labeled *Safe-0*. One stakeholder in such a pair chose the safe option and received an intermediate income in the range from 100 tokens to

Figure 22: Redistribution Index by pair composition and treatment



Notes: The Redistribution Index compares the extent of redistribution within each pair of stakeholders before and after redistribution. An index of zero indicates no redistribution, while an index of 1 indicates that spectators have equally split the pair's total earnings.

500 tokens. The other stakeholder earned zero tokens after having chosen the risky option. The third pair is denoted as *Safe-800.* In this pair, one stakeholder earned the high income of 800 tokens from the risky option, and the other one chose the safe option with an intermediate income from 100 tokens to 500 tokens.

Figure 22 reports the extent of redistribution, depending on the type of stakeholders' pair. The vertical axis presents a measure of redistribution that we call Redistribution Index (RI) which is defined as follows:

$$RI = \frac{\pi^R_{pre} - \pi^R_{post}}{\pi^R_{pre} - X/2}$$

where π^R_t are the earnings for a stakeholder R before ($t = pre$) or after ($t = post$) the redistribution stage, and X are the pair's total earnings. We indicate as the richer stakeholder R the person in the pair with the larger earnings before the redistribution stage.[9] An index $RI = 0$ indicates no redistribution at all, while $RI > 0$ indicates that spectators have shifted resources to the poorer stakeholder. An $RI = 1$ means that spectators have implemented an equal split of the pair's total earnings. For $RI > 1$ the pre-redistribution ranking of stakeholders is reversed – what used to be the richer stakeholder before the redistribution is now the poorer one.[10]

In Figure 22, we observe the highest degree of redistribution – $RI = 0.67$ on average – in the pairs of the type *800-0*. Furthermore, in these pairs there are only minor differences in the extent of redistribution across the three treatments.

In the *Safe-0* pairs, the redistribution is much smaller, $RI = 0.42$ on average, compared to the case in which both stakeholders made the same ex-ante decision. In *Safe-0* pairs we observe only a moderate increase in redistribution in *Self-Report* compared to *Nature*. This is noteworthy because stakeholders reporting the low income of zero tokens in the *Self-Report* treatment are almost certainly honest subjects who resisted the temptation of cheating about their income. Spectators would have had the chance to *reward* such honest stakeholders by redistributing more money in their favor in the *Self-Report* treatment than in the *Nature* treatment, since spectators can infer honesty in the *Self-Report* treatment, but not in the *Nature* treatment. The introduction of externalities triggered a somewhat higher level of redistribution ($RI = 0.52$ versus $RI = 0.35$ in *Nature*). This could be explained by the fact that a stakeholder reporting a low income in the *Externality* treatment automatically grants a high income to an idle player. Such an action can hence be interpreted as an additional sign of kindness and not only of honesty, suggesting that spectators reward kindness.

Although spectators do not seem to reward honesty, they strongly react to potential dishonesty. We see this in the right-most bars in Figure 22,

[9]Please recall that we only consider cases with pre-redistribution inequality and therefore can always uniquely identify the richer stakeholder in the pair.

[10]For *800-0* and *Safe-0* pairs the RI can only take values between 0 and 2. Negative values, as well as values larger than 2, are possible only for the *Safe-800* pairs. Overall, we observed a negative RI index in 5.45% of the *Safe-800* pairs and a $RI > 2$ in 5.25% of the *Safe-800* cases.

where the redistribution in pairs of the type *Safe-800* is shown. Here we note that the average redistribution index is only 0.30 in *Nature*, but that it increases to 0.48 in *Self-Report* and even to 0.62 in *Externality*. Hence, when a high income of 800 tokens is potentially caused by cheating and the respective stakeholder is paired with someone who chose the safe option, spectators are much more willing to take away money from such a stakeholder than when they can be sure that the stakeholder had no means of earning such a high income through dishonest behavior.

Table 22 reports a series of GLS estimations providing statistical support for the evidence in Figure 22. In Table 22, the dependent variable is the Redistribution Index and the main explanatory variables of interest are dummies for the *Self-Report* treatment, the *Externality* treatment, for *Safe-0* pairs, and for *Safe-800* pairs, thus taking *800-0* pairs as benchmark. In addition, we are interested in the interaction between *Self-Report*, respectively *Externality*, and the dummies for pair composition. Model 1 in Table 22 only considers treatment dummies. Both for *Self-Report* and *Externality* the coefficients are positive, but only for *Externality* it is significant – showing that the level of redistribution gets larger when cheating is associated with externalities on powerless third parties. In Model 2, we add dummies for *Safe-0* and *Safe-800* pairs and both coefficients are negative and highly significant, thus suggesting less inequality reduction as compared to *800-0* pairs. In Model 3, the positive and highly significant coefficient for the interaction between the treatments (*Self-Report* as well as *Externality*) and *Safe-800* confirms the observation from Figure 22 that there is more redistribution in these pairs in the *Self-Report* and *Externality* treatments than in the *Nature* one. Redistribution remains significantly larger in *Externality* than in *Nature* (see post-estimation test at the bottom of the Table), but the differences between *Self-Report* and *Externality* fail conventional significance (with p-values slightly larger than 0.1; see bottom of Table 2). These results from models 1 to 3 are robust after controlling for socio-demographic characteristics and personality traits (Model 4) and after introducing fixed effects for the safe income level (Model 5).

Result 3. *Redistribution is lowest in Nature, intermediate in Self-Report, and highest in Externality. This general pattern is most pronounced in pairs where one rich stakeholder (with high income of 800) is paired with a poorer one who chose the safe intermediate amount. In these pairs, the possibility of cheating on the side of rich stakeholders triggers significantly more redistribution through spectators than when cheating is not an option (in Nature).*

2.4 ESTIMATION OF FAIRNESS VIEWS

Our experimental design allows estimating spectators' fairness views based on their 20 redistribution choices. We are going to introduce a discrete choice random utility model (following Cappelen et al. 2007, 2013) and then present the distribution of fairness views, showing how the shadow of cheating leads to a strong shift in this distribution. At the end of this section, we will examine how the estimated fairness views depend on spectators' beliefs, their personality traits, and their political orientation.

For the estimation of different types, we assume spectators are only motivated by fairness views, because self-interest does not play a role in our set-up, given the flat payment of spectators. Specifically, if X is the total income in the pair of stakeholders to which a spectator is assigned, we assume that the spectator's utility from giving y to the first and $X - y$ to the second stakeholder is given by:

$$V(y; \cdot) = -\beta(y - F^k)^2/2X$$

where F^k is the fair amount allocated to the first stakeholder according to the spectator's fairness view k and where β is the weight attached to fairness. A spectator's utility is decreasing in the distance between the amount (y) allocated to the first stakeholder and the fair amount F^k prescribed by the fairness view k.

Spectators can differ along two dimensions: (i) how much they care about fairness (β); and (ii) their fairness views (F^k). In line with previous papers, we consider three possible types of fairness views:

Table 22: Determinants of the Redistribution Index

Dependent variable: Redistribution Index (RI)	Model 1	Model 2	Model 3	Model 4	Model 5
Self-Report treatment *(d)*	0.117	0.115	0.052	0.079	0.065
	(0.087)	(0.086)	(0.093)	(0.093)	(0.094)
Externality treatment *(d)*	0.221**	0.219**	0.114	0.158	0.146
	(0.100)	(0.099)	(0.107)	(0.108)	(0.108)
Safe-0 pair *(d)*		-0.233***	-0.268***	-0.268***	-0.308***
		(0.028)	(0.050)	(0.050)	(0.051)
Safe-800 pair *(d)*		-0.195***	-0.319***	-0.319***	-0.359***
		(0.021)	(0.047)	(0.047)	(0.049)
Self-Report x *Safe-0*			0.033	0.033	0.057
			(0.065)	(0.065)	(0.066)
Self-Report x *Safe-800*			0.132**	0.132**	0.160***
			(0.056)	(0.056)	(0.061)
Externality x *Safe-0*			0.053	0.053	0.078
			(0.076)	(0.076)	(0.077)
Externality x *Safe-800*			0.201***	0.201***	0.227***
			(0.063)	(0.063)	(0.067)
Constant	0.412***	0.559***	0.618***	0.994***	0.976***
	(0.072)	(0.072)	(0.077)	(0.327)	(0.326)
Controls	No	No	No	Yes	Yes
Safe level fixed effect	No	No	No	No	Yes
Post-estimation F-tests					
Self-Report vs. *Externality* †	–	–	$p = .107$	$p = .108$	$p = .187$
Nature vs. *Externality* §	–	–	$p < .001$	$p < .001$	$p < .001$
N.obs	3291	3291	3291	3291	3291
R^2 (overall)	0.011	0.025	0.027	0.049	0.053

Notes: GLS regression with individual random effects. Symbols $***$ and $**$ indicate significance at the 1% and 5%, respectively. The regressions only include pairs with strictly positive pre-redistribution inequality. Dummy variables are indicated by *(d)*. Controls include the following variables: *Male* takes value 1 for males and 0 for females. Political orientation was measured on a scale from 1 (left) to 10 (right) in the final questionnaire. *Center* takes value 1 for participants that indicated a value between 4 and 7, and 0 otherwise. *Right* takes value 1 for participants that indicated a value between 8 and 10, and 0 otherwise. The Honesty-Humility score is based on the HEXACO-PI. "Persons with very high scores on the Honesty-Humility scale avoid manipulating others for personal gain, feel little temptation to break rules, are uninterested in lavish wealth and luxuries, and feel no special entitlement to elevated social status." (`http://hexaco.org/scaledescriptions`) *Inequality (questionnaire)* is a self-reported variable ranging from 1 (a society should aim to equalize incomes) to 10 (a society should **not** aim to equalize income). † *Self-Report* vs. *Externality*, test of joint significance: *Safe-0* = *Safe-800* & *Self-Report* x *Safe-0* = *Self-Report* x *Safe-800* & *Externality* x *Safe-0* = *Externality* x *Safe-800*. § *Nature* vs. *Externality*, test of joint significance: *Externality* = 0 & *Externality* x *Safe-0* = 0 & *Externality* x *Safe-800* = 0.

- ***Libertarians*** never support redistribution, and no matter what the severity of, and the reasons for, the inequality are, they leave the earnings within a pair of stakeholders unaltered. If x is the income of the first stakeholder before redistribution, we have $F^{Libertarians} = x$, which yields the optimal choice $y = x$.

- ***Egalitarians*** always eliminate inequality within a pair and split the earnings equally: $F^{Egalitarians} = X/2$, which yields the optimal choice $y = X/2$.

- ***Choice Egalitarians*** eliminate inequality only when the disparities are generated by luck in case two stakeholders have chosen the same option (i.e., the risky option), but do not redistribute if inequality reflects differences in choices (safe option vs risky option):

$$F^{ChoiceEgalitarians} = \begin{cases} X/2 & \text{if } C_1 = C_2 \\ x & \text{if } C_1 \neq C_2 \end{cases}$$

  where C_i takes value 1 if stakeholder i chooses the risky option and 0 if he/she chooses the safe option with the safe income level.

Looking at the descriptive data, we observe that 73.1% of all decisions correspond exactly to one of the three types (68.8% in *Nature*, 75.4% in *Self-Report* and 71.5 % in *Externality*).[11] These fractions correspond to the number of decisions consistent with the action prescribed by at least one fairness view, and they do not indicate the fraction of spectators being classified as *pure types*.

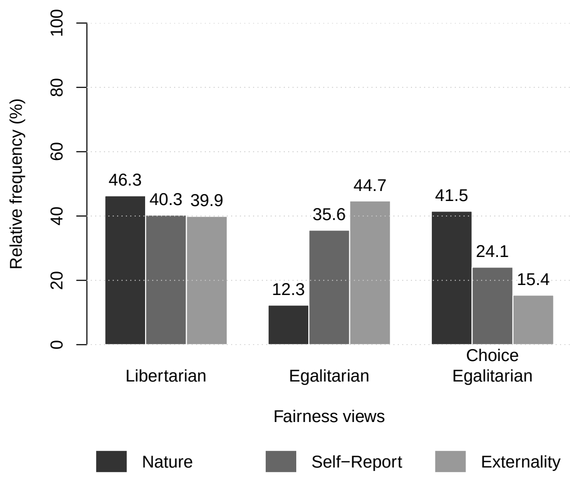
Since we let all spectators make 20 redistribution decisions, we can estimate the likelihood with which a spectator belongs to any of the three different types of fairness views. Given a spectator's fairness view k, we consider a discrete choice random utility model of the form

$$U(y; \cdot) = V(y; \cdot) + \varepsilon_{iy} \qquad \text{for } y = 0, 25, \ldots, X \tag{2.4.1}$$

[11] Note that our fraction of 73.1% of decisions that match at least one type is remarkably similar to the 71.1% reported in Cappelen et al. (2013) for their experiment in Norway.

where ε_{iy} is assumed to be i.i.d extreme value distributed and, to control for individual heterogeneity in noisy behavior, β is assumed to be log normally distributed with $\log(\beta) \sim \mathcal{N}(\zeta, \sigma^2)$. Denoting by $L_{i,k}$ the individual likelihood conditional on being of type k, we can obtain the total likelihood of an individual by considering the finite mixture of types $L_i = \sum_k \lambda^k L_{i,k}$, where λ^k is the probability of being of type k.[12]

Figure 23: Estimation of fairness views



Notes: The bars report the share of λ^k (in %). The results are based on a discrete choice random utility model.

Figure 23 and Table A-2 in the Appendix report the estimated proportion of types, λ^{k}[13]. *Libertarians* account for a fairly large share of the spectators in all treatments, ranging from 40% in *Self-Report* and *Externality* to 46% in *Nature*. Apart from this similarity, the distribution of fairness types differs substantially between *Nature* and the two treatments where cheating is possible (likelihood ratio tests: *Nature* versus *Self-Report*, $\chi^2(4) = 13.696$; $p = .008$; *Nature* vs. *Externality*, $\chi^2(4) = 17.468$, $p = .002$).

[12]For further details on the estimation strategy, please refer to section 2.5 in the Appendix or see Cappelen et al. (2013).

[13]Numerical integration is performed using 100 Halton draws for each observation (Train, 2009)

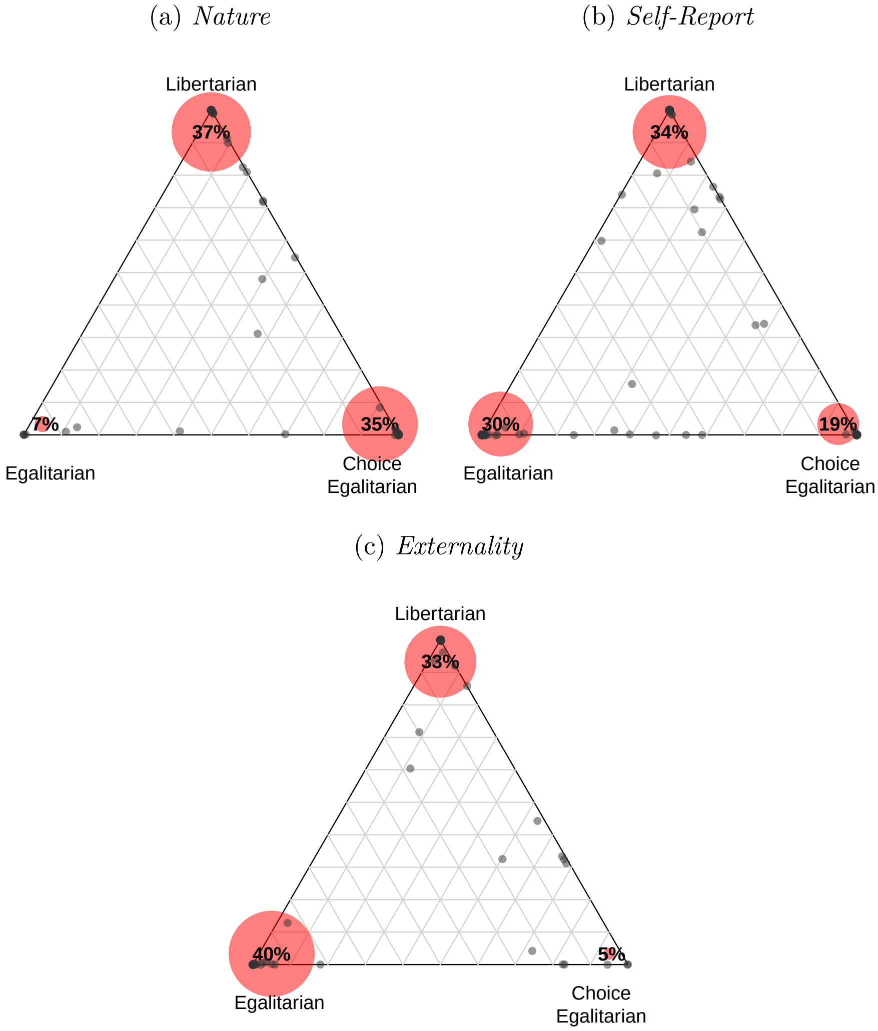
In *Nature*, only 12% of spectators are *Egalitarians*, while a large share of spectators (41%) are *Choice Egalitarians.* This pattern is completely reversed when cheating is possible. In *Self-Report*, 36% of spectators are classified as *Egalitarians* and only 24% as *Choice Egalitarians.* When cheating comes with an externality, the reversal becomes even more striking, as in *Externality* we classify 45% as *Egalitarians*, and only 15% as *Choice Egalitarians.* In other words, under the shadow of cheating spectators are much less likely to condition their redistribution decision on whether the two stakeholders chose the same action – i.e., the risky option – or not. Rather, unconditional egalitarianism becomes much more prominent. As a consequence, there are two diametrically opposed fairness views that dominate in the treatments with a possibility to cheat (to get rich): *Libertarians* who do not want to redistribute anything, and *Egalitarians* who prefer redistribution to the fullest extent.

We observe a difference also in the distribution of fairness views between *Self-Report* and *Externality* – albeit this difference is much smaller than the one between *Nature* and the other treatments. The fact that cheating imposes negative consequences on an idle player leads to an even stronger shift in spectators' fairness views compared to costless cheating ($\chi^2(4) = 9.282$, $p = .054$ for *Self-Report* versus *Efficiency*). The potentially illegitimate claim of the high income now intercepts another person's income, making cheating not just an unethical, but also harmful action. This leads spectators to redistribute more than when stakeholders only cheated for their own benefit, without causing negative externalities on others.

Result 4. *The shadow of cheating produces a large and statistically significant shift in fairness views. While the fraction of Libertarians is similar across treatments, the share of Egalitarians becomes three to four times as large in the treatments where cheating is an opportunity. This implies that the shadow of cheating creates a polarization of fairness views, even more so in Externality than in Self-Report.*

To test the accuracy of our type classification, we compute the ex-post probability of any specific spectator to belong to a particular fairness

Figure 24: Fairness types and posterior probabilities



Notes: Each vertex of the triangle represents a fairness type and the bubbles in the corners report the relative frequency of spectators for whom we estimate a posterior probability higher than 90% of being of that particular type.

type. Figure 24 reports the simplex with the posterior probability for each spectator – see Conte and Hey (2013) for a similar exercise. Each vertex of the triangle represents one fairness type and each dot represents one spectator. The bubbles in the corners report the percentage of spectators who have a posterior probability higher than 90% of being that type. We can observe that the vast majority of the spectators – 79% in *Nature*, 83% in *Self-Report*, and 78% in *Externality* – are located in one of the three corners, hence suggesting that types are identified with great precision. The shift in fairness types from *Nature* to the other two treatments is illustrated on the horizontal axis at the bottom of all triangles in Figure 24 where we see the shift from *Choice Egalitarians* (in *Nature*) to *Egalitarians* (in *Self-Report* and *Externality*).[14]

THE ROLE OF BELIEFS. Next, we present some further analysis to investigate what might determine a spectator's fairness views. A straightforward candidate to drive one's fairness views is beliefs. It could be that spectators with different fairness views hold significantly different beliefs about the likelihood with which a stakeholder's high income might have been caused by cheating. For instance, one could imagine that *Libertarians* want to abstain from any kind of redistribution because they expect stakeholders to be (mostly) honest and therefore see no reason to take money away from them. Similarly, one could argue that *Egalitarians* favor extensive redistribution because they expect high income to be undeserved and (mostly) the result of cheating.

Figure 25, however, suggests that there is no correlation between beliefs and types.[15] This figure is based on data from two *Self-Report* and two *Externality* sessions in which we asked a total of 60 spectators in each treatment to guess (in an incentive compatible way) the fraction of stakeholders

[14] Actual and predicted redistribution choices are reported in the Appendix in Figures A-2 to A-4.

[15] See also Tables A-3 and A-4 in the Appendix for regressions. Both in Figure 25 and Tables A-3 and A-4 we define types based on posterior probabilities. Each spectator is assigned to a particular type if the posterior probability of being of that type is at least 0.5. The results are robust to more demanding cut-offs of 0.7 and 0.9, for instance.

who choose the risky option and how many of the latter report the high income. To avoid any priming or experimenter demand effect, we elicited beliefs only at the end of the session, after spectators had made all their distributive choices. For the sake of brevity, we elicited beliefs only for the income level of 300 tokens in the safe option.

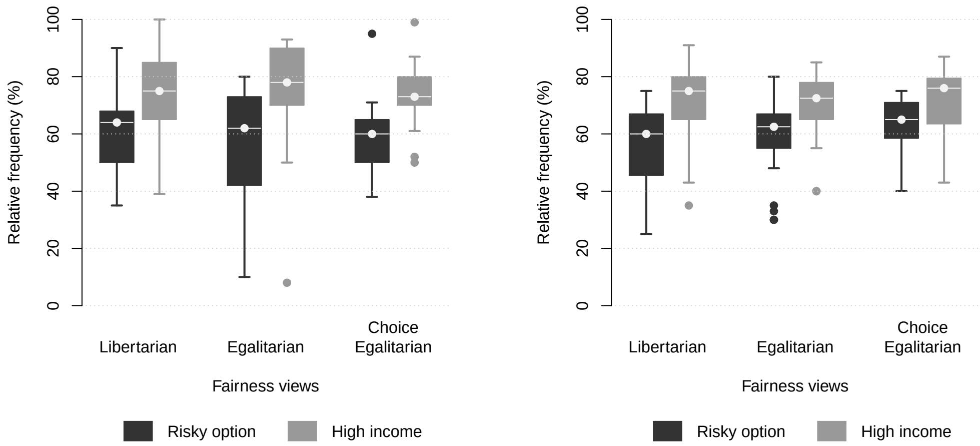
For this safe level, spectators expect 58% of stakeholders to choose the risky option in *Self-Report* and 59% in *Externality*. The expected fraction is quite close to the actual frequency (of 52% in *Self-Report* and 51% in *Externality*) with which stakeholders choose the risky option when the safe option pays 300 tokens (see right-hand side and bottom of Figure 21). Spectators expect on average that 74% of stakeholders who choose the risky option in *Self-Report* report the high income (70% in *Externality*), even though truthful reporting would yield a 50% chance for the high income. The expected fraction of reporting the high income (74%, respectively 70%) is again very close to the actual share of stakeholders reporting the high income (in case the safe option pays 300 tokens: 67% and 71% in *Self-Report* and *Externality*, respectively).

Interestingly, Figure 25 reveals that there are no differences in the beliefs of spectators with different fairness views and this is true for both treatments. We consider this a noteworthy finding. For instance, both *Libertarians* and *Egalitarians* expect three quarters of stakeholders who choose the risky option to report the high income. Evidently, both *Libertarians* and *Egalitarians* infer from this large fraction of high income that cheating is going on, but they nevertheless make opposite redistribution choices. Hence, fairness views are obviously not significantly driven by the beliefs about the risk-taking behavior and the expected honesty of stakeholders.[16]

[16] In order to address potential concerns about spectators' beliefs being elicited only after 20 redistribution choices, we also elicited beliefs among students not involved in the redistribution task. For that purpose, we invited 289 additional students from an Introductory Microeconomics course at the University of Cologne to predict (in an incentive compatible way) the stakeholders' behavior, conditional on the different levels of income from the safe option. For the safe amount of 300 tokens, they expected 66% of stakeholders to choose the risky option (where the actual relative frequency is 52%). Students estimated that on average 69% of stakeholders who choose the risky option would report the high income while, in fact, the actual number is 67%. Hence, these new

Figure 25: Spectators' beliefs about stakeholders' behavior

(a) *Self-Report* (b) *Externality*



Notes: The figure shows the distribution of spectators' beliefs about the percentage of stakeholders choosing the risky option (light gray) and about the percentage of stakeholders reporting the high income (black), by fairness views. The white line inside the boxes indicates the median of the distribution, the box represents the interquartile range, and the whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range. Dots indicate outliers, i.e., data points lying outside the whiskers.

Result 5. *Libertarians, Egalitarians, and Choice Egalitarians hold very similar beliefs about risk-taking and cheating among stakeholders (in the Self-Report and in the Externality treatment), but nevertheless make different redistribution choices.*

2.5 CONCLUSION

The growing gap between the haves and have-nots has revived a debate about what constitutes a fair level of redistribution to alleviate income inequalities

students who had not taken part in any of our treatments before were very capable of predicting the relative frequency of cheating among stakeholders, even when they were not asked to make redistribution choices themselves. This evidence suggests that our design did not induce any bias in spectators' beliefs.

within societies. This debate does not reject the possibility that income inequalities per se may actually work as an incentive for increasing efforts on the side of the less well-off. Accordingly, it is understood that there is a trade-off between efficiency and equality (Balafoutas et al., 2013). Whether a given society leans more towards incentives for efficiency – by largely refraining from redistribution from the rich to the poor in the hope of increasing effort levels – or favors more equality – by supporting more redistribution – depends largely on the perceived sources of income inequality (Konow, 2000; Alesina and Angeletos, 2005; Cappelen et al., 2013; Möllerström et al., 2015; Almås et al., 2016; Konow et al., 2016).

In this paper, we have introduced the possibility of cheating as a potentially important source of income inequality into the framework. We have studied how the shadow of cheating affects fairness views and the desired level of redistribution from the rich to the poor. A major motivation for our study was the observation that citizens in countries with larger income inequalities are also more likely to perceive their fellow citizens as corrupt and dishonest. Hence, cheating behavior can be suspected to be a source of income inequality. While it seems undisputed that income inequalities clearly caused by cheating should be eliminated, the situation becomes much less clear when it cannot be proven whether income inequalities have been caused by wealthy subjects cheating, or whether the wealthy acquired their income by honest means. Such ambiguous situations are hard to study in the field because too many other factors (such as institutional and legal frameworks, or the effectiveness of the legal system to detect cheating) come into play to isolate the effect the shadow of cheating casts on fairness views. For this reason, we have presented the first controlled laboratory experiment on how potential cheating as the source of income inequalities affects fairness views of impartial spectators who can redistribute money in pairs of rich and poor stakeholders. Our experimental treatment variation has allowed us to implement two otherwise identical conditions: one in which income inequalities cannot be caused by cheating, and another one in which cheating may well be the reason for income inequalities, but where spectators have no means to detect it. This implies they must take the shadow of cheating into consideration while making their redistribution decisions.

We have found a substantial shift in the distribution of fairness views when spectators know that cheating is possible. More precisely, under the shadow of cheating, we have observed a split of the spectators into two diametrically opposed subpopulations. On one side of the spectrum are *Libertarians* who abstain from any redistribution, no matter what might be the source of income inequalities. On the other side of the spectrum are *Egalitarians* who implement perfect equality.

The polarization of fairness views under the shadow of cheating is mainly driven by a strong shift from *Choice Egalitarians* – who do not redistribute between stakeholders when they have generated their income from different actions – to *Egalitarians*. Under the shadow of cheating, spectators face the conundrum of whether to take money away from a rich person who has either rightfully earned it or who may have purposefully acted dishonestly to profit from an unobservable situation. The strong increase in Egalitarian fairness views in such an environment may reflect the spectators' wish to reward stakeholders who chose the safe option and thus refrained from the temptation of falsely reporting their earnings. This tendency is even more pronounced in *Externality*, hence suggesting that a considerable portion of the population might feel especially strongly about supporting the welfare state when innocent parties suffer losses as a byproduct of someone else's unethical behavior.

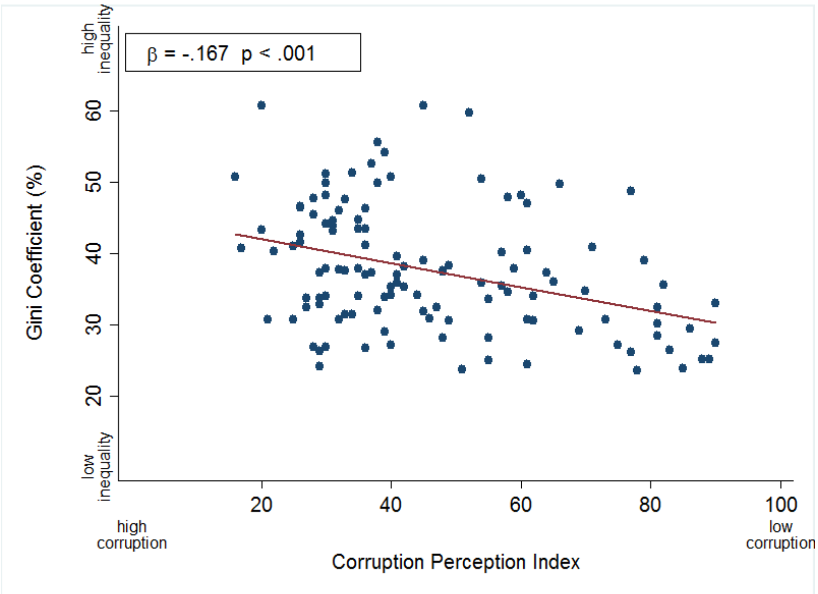
While the shadow of cheating has led to a large increase in the number of *Egalitarians*, we find it remarkable that the proportion of spectators with a *Libertarian* point of view has remained practically the same across all treatments. The fairness views of these spectators have not been significantly altered by suspected unethical behavior. This suggests that either dishonesty itself is not a good reason for these spectators to reduce inequality or they are concerned to wrongfully take money away from truly lucky stakeholders who have truthfully reported their income. We can rule out that *Libertarians* have different beliefs about the honesty of stakeholders than *Egalitarians* or *Choice Egalitarians* have. We consider the lack of differences in beliefs an important finding. Indeed, *Libertarians* deliberately refrain from redistributing despite knowing that some of the income disparities are caused by cheating on the part of the rich. Likewise, *Egalitarians* support an equal distribution of income although they acknowledge that some inequalities were not caused by dishonesty on the part of the rich.

Overall, the shadow of cheating has created a polarization of fairness views at opposite ends of the spectrum, having *Egalitarians* on the one end and *Libertarians* on the other end. This polarization has been most pronounced in the case where cheating had a negative externality on a powerless third party (in treatment *Externality*). We consider it a plausible assumption that negative externalities of cheating are rather the rule than the exception, for which reason we argue that the extent of the polarization of fairness views observed in *Self-Report* is most likely measuring a lower bound.

Overall, our findings suggest that the shadow of cheating could lead to increased social tensions and more disruptive changes in redistribution policies when political majorities swing back and forth between one camp (of *Egalitarians*) and the other camp (of *Libertarians*). Politicians might want to take this factor into account when setting the legal and institutional framework that is intended to prevent illicit behavior of citizens. In fact, failing to fight dishonesty will not only cause more illicit activities, but – according to our findings – it will also contribute to a polarization of fairness views and a demand for redistribution. The latter effect is likely an overlooked side-effect of failed attempts to fight corruption and illegal activities, which are often at the root of large income inequalities (Glaeser et al., 2003).

APPENDIX A: TABLES AND FIGURES

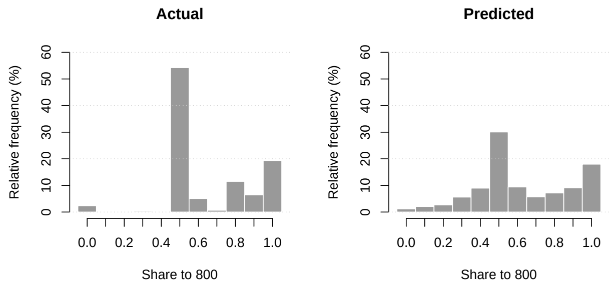
Figure A-1: Gini Index and Corruption Perception Index



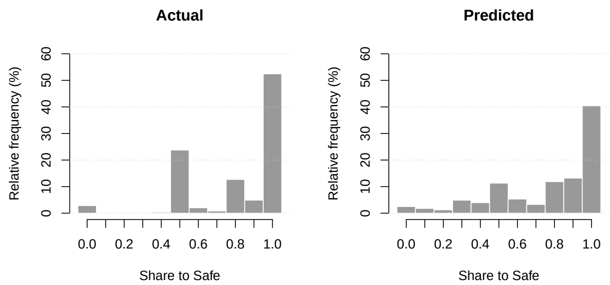
Notes: The graph reports data for 120 countries. The coefficient and significance level are obtained from an OLS regression. Data for the Gini coefficient are from the World Income Inequality Database (WIID release 3.4 `https://www.wider.unu.edu/data`). For each country, we considered the most recent year available. For the sake of homogeneity, we excluded countries with data only prior to year 2010 from the analysis. In case of multiple sources for the selected year, we computed the Gini coefficient as the average of all available sources. A coefficient of 0% indicates complete equality, a coefficient of 100% indicates complete inequality. The Corruption Perception Index is based on data from Transparency International (`https://www.transparency.org/`) and refers to year 2016 for all countries. An index of 0 indicates that a country is perceived as highly corrupt, while an index of 100 indicates that a country is perceived as very clean.

Figure A-2: Actual and predicted income redistribution - *Nature* treatment

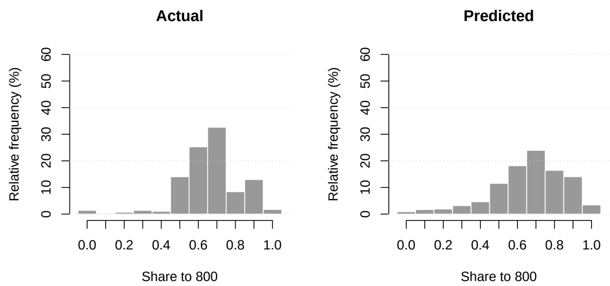
(a) *800-0*



(b) *Safe-0*



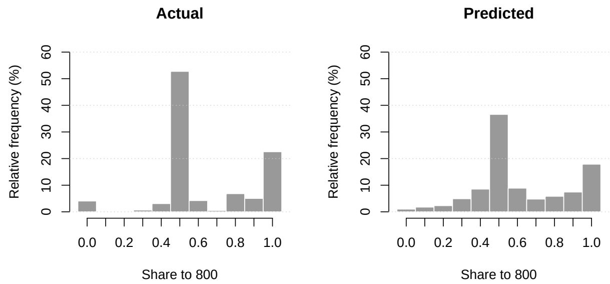
(c) *Safe-800*



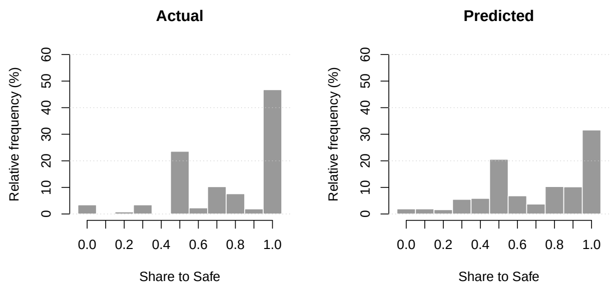
Notes: "Actual" refers to the choice made by the spectators. "Predicted" refers to simulated choices obtained using the discrete choice random utility model and the estimated parameters in Table A-2. For each spectator, we run 1000 simulations of the 20 choices he/she faced. In each simulation, we randomly draw a fairness view F^k and a β in accordance with the estimated parameters. Panel (a) shows actual and predicted choices in pairs of the type *800-0*; Panel (b) shows actual and predicted choices in pairs of the type *Safe-0*; and Panel (c) shows actual and predicted choices in pairs of the type *Safe-800*.

Figure A-3: Actual and predicted income redistribution - *Self-Report* treatment

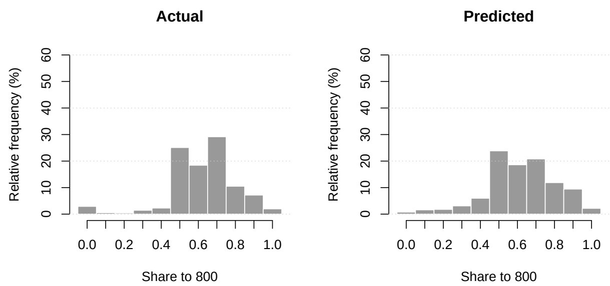
(a) *800-0*



(b) *Safe-0*



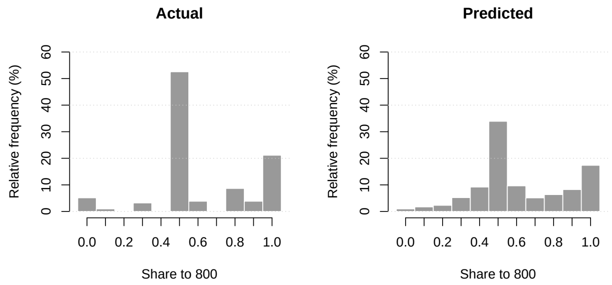
(c) *Safe-800*



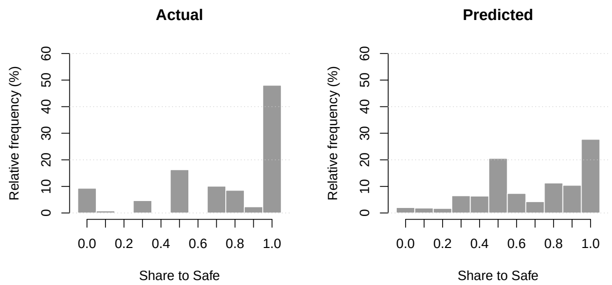
Notes: "Actual" refers to the choice made by the spectators. "Predicted" refers to simulated choices obtained using the discrete choice random utility model and the estimated parameters in Table A-2. For each spectator, we run 1000 simulations of the 20 choices he/she faced. In each simulation, we randomly draw a fairness view F^k and a β in accordance with the estimated parameters. Panel (a) shows actual and predicted choices in pairs of the type *800-0*; Panel (b) shows actual and predicted choices in pairs of the type *Safe-0*; and Panel (c) shows actual and predicted choices in pairs of the type *Safe-800*.

Figure A-4: Actual and predicted income redistribution - *Externality* treatment

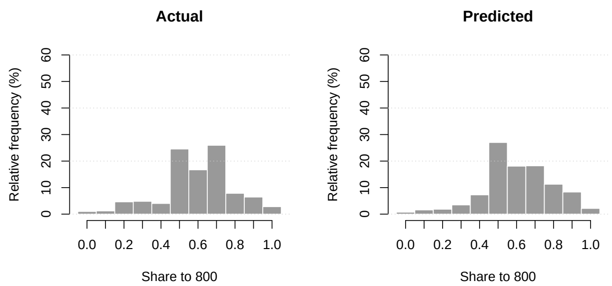
(a) *800-0*



(b) *Safe-0*



(c) *Safe-800*



Notes: "Actual" refers to the choice made by the spectators. "Predicted" refers to simulated choices obtained using the discrete choice random utility model and the estimated parameters in Table A-2. For each spectator, we run 1000 simulations of the 20 choices he/she faced. In each simulation, we randomly draw a fairness view F^k and a β in accordance with the estimated parameters. Panel (a) shows actual and predicted choices in pairs of the type *800-0*; Panel (b) shows actual and predicted choices in pairs of the type *Safe-0*; and Panel (c) shows actual and predicted choices in pairs of the type *Safe-800*.

Table A-1: Procedures and generation of decision sequence

	Nature			Self-Report and Externality		
Scenario	Stakeholder 1	Stakeholder 2	Safe level	Stakeholder 1	Stakeholder 2	Safe level
Based on pilot experiments						
1	Safe	800	300	Safe	800	400
2	800	0	200	Safe	0	300
3	0	0	200	Safe	800	200
4	0	0	100	800	0	300
5	800	0	100	800	0	200
6	0	0	100	Safe	Safe	400
7	Safe	800	300	0	800	200
8	Safe	800	100	800	800	100
9	0	Safe	100	800	800	200
10	Safe	Safe	400	800	Safe	100
11	800	Safe	500	Safe	800	400
12	0	0	100	800	Safe	400
13	0	Safe	500	Safe	Safe	500
14	0	0	300	Safe	800	500
15	800	0	100	Safe	800	400
16	Safe	0	400	800	800	200
Pre-defined by the experimenter						
17	Safe	0	S_i	Safe	0	S_i
18	800	Safe	S_i	800	Safe	S_i
19	0	800	S_i	0	800	S_i
Relevant for stakeholders' earnings						
20	*stakeholder 1*	*stakeholder 2*	S	*stakeholder 1*	*stakeholder 2*	S

Notes: Each stakeholder faced 20 scenarios. Scenarios 1 to 16 were based on a pilot experiment with 30 stakeholders per treatment and ran a few weeks prior to the proper experiment. Even though the sequences were pre-determined, all pairs were a possible outcome. Each scenario was generated by randomly drawing a pair (with reposition) and by randomly selecting a safe level for each chosen pair. The relevant outcomes for the selected pairs and safe level are reported in the table. The first 16 scenarios were treatment specific. Data from the pilot experiment on MTurk and the code to generate the sequence are available upon request from the authors. The outcomes (Safe, 800, 0) for the scenarios 17 to 19 were defined by experimenters and represent pairs with initial inequality. The safe level for these scenarios, S_i, was randomly drawn. An independent random draw was performed for each spectator. The randomly selected safe level was kept constant across scenarios 17 to 19. Finally, the last scenario was the payoff-relevant one. Each spectator was assigned to one pair of stakeholders.

Table A-2: Estimation of types

	Nature	*Self-Report*	*Externality*	Pooled	Pooled	Pooled	Pooled
				Nature & Self-Report	*Nature & Externality*	*Self-Report & Externality*	All
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
$\lambda^{Libertarians}$	0.463	0.403	0.399	0.418	0.419	0.401	0.411
	(0.071)	(0.047)	(0.069)	(0.039)	(0.049)	(0.039)	(0.034)
$\lambda^{Egalitarians}$	0.123	0.356	0.447	0.282	0.293	0.216	0.323
	(0.048)	(0.046)	(0.068)	(0.036)	(0.045)	(0.034)	(0.032)
$\lambda^{ChoiceEgalitarians}$	0.415	0.241	0.154	0.300	0.288	0.384	0.266
	(0.072)	(0.043)	(0.057)	(0.038)	(0.047)	(0.038)	(0.032)
ζ	4.635	5.297	5.026	5.110	5.019	5.092	5.082
	(0.117)	(0.059)	(0.062)	(0.051)	(0.059)	(0.040)	(0.038)
σ	3.161	3.351	2.937	3.103	2.978	3.045	3.049
	(0.127)	(0.070)	(0.077)	(0.058)	(0.065)	(0.047)	(0.044)
logLik	-1871.573	-5413.893	-3064.689	-7292.313	-4944.996	-8483.223	-10362.72
Degrees of freedom	4	4	4	4	4	4	4

Notes: The likelihood is maximized in R using the BFGS method with mle2 function (bbmle package). One population share and its standard error are calculated residually. Numerical integration is perfomed using 100 halton draws for each observation (Train, 2009). Models 1 to 3 report estimates separately by treatment: *Nature*, *Self-Report*, and *Externality.* Model 3 to 6 estimate pooled data from a pair of treatments, while Model 7 pools all the data.

Table A-3: Beliefs and fairness views – *Self-Report* treatment

Dep. var.:	Risky choices		High income	
Ex-post beliefs	Model 1	Model 2	Model 3	Model 4
Egalitarians *(d)*	-1.506	-0.266	-1.979	-0.424
	(5.109)	(5.728)	(5.127)	(5.705)
Choice egalitarians *(d)*	-0.245	0.216	-1.527	0.691
	(5.250)	(5.853)	(5.268)	(5.829)
Male *(d)*		-3.201		3.818
		(4.948)		(4.928)
Age (years)		-0.007		0.207
		(0.869)		(0.866)
Honesty-Humility score		-1.254		-0.052
		(0.926)		(0.922)
Center *(d)*		1.835		2.488
		(5.426)		(5.404)
Right *(d)*		4.764		1.064
		(8.249)		(8.215)
Inequality (survey)		-0.986		0.970
		(1.090)		(1.086)
Risk aversion		-1.996		-2.300
		(1.637)		(1.632)
Constant	58.941***	85.837***	75.122***	71.084***
	(3.336)	(25.859)	(3.348)	(25.754)
N.obs.	60	60	60	60
R^2	0.002	0.068	0.003	0.082

Notes: OLS regression. In Models 1 and 2, the dependent variable is the estimated frequency of risky choices; in Models 3 and 4, the dependent variable is the estimated fraction of subjects who report the high income from the risky investment. Symbols $***$ indicates significance at the 1% level, respectively. Dummy variables are indicated by *(d)*. *Risk aversion* takes values from 1 to 6, where 1 indicates risk aversion and 6 risk loving. See Table 22 for the explanation of the other regressors. Post estimation tests for Egalitarians = Choice Egalitarians: Model 1: $p = .824$; Model 2: $p = .938$; Model 3: $p = .937$; Model 4: $p = .857$.

Table A-4: Beliefs and fairness views – *Externality* treatment

Dep. var.:	Risky choices		High income	
Ex-post beliefs	Model 1	Model 2	Model 3	Model 4
Egalitarians *(d)*	1.964	-1.072	-1.007	3.945
	(3.899)	(4.362)	(3.955)	(4.553)
Choice egalitarians *(d)*	4.784	0.923	1.688	8.513
	(6.137)	(7.258)	(6.225)	(7.575)
Male *(d)*		-1.598		9.247**
		(4.081)		(4.259)
Age (years)		-1.000***		-0.065
		(0.329)		(0.343)
Honesty-Humility (score)		0.179		-0.396
		(0.766)		(0.799)
Center *(d)*		0.612		-0.283
		(4.724)		(4.931)
Right *(d)*		-3.328		-4.074
		(6.667)		(6.959)
Inequality (survey)		0.223		1.144
		(1.088)		(1.136)
Risk aversion		1.516		-0.359
		(1.456)		(1.520)
Constant	57.217***	75.430***	69.872***	66.304***
	(2.837)	(15.394)	(2.877)	(16.067)
N.obs.	60	60	60	60
R2	0.011	0.213	0.004	0.161

Notes: OLS regression. In Models 1 and 2, the dependent variable is the estimated frequency of risky choices; in Models 3 and 4, the dependent variable is the estimated fraction of subjects who report the high income from the risky investment. Symbols $***$ and $**$ indicate significance at the 1% and 5%, respectively. Dummy variables are indicated by *(d)*. *Risk aversion* takes values from 1 to 6, where 1 indicates risk aversion and 6 risk loving. See Table 22 for the explanation of the other regressors. Post estimation tests for Egalitarians = Choice Egalitarians: Model 1: $p = .636$; Model 2: $p = .759$; Model 3: $p = .656$; Model 4: $p = .502$.

APPENDIX B: ESTIMATION PROCEDURE

In this Appendix, we provide further details about the estimation of fairness views (based on Cappelen et al. 2007, 2013).

Given the random utility model in equation (2.4.1) and under the assumption that ε_{iy} is i.i.d. extreme value distributed and that $\log(\beta)$ is $\mathcal{N}(\zeta, \sigma^2)$, we can write the likelihood contribution of a spectator i conditional on fairness view k as follows:

$$L_{i,k}(\zeta, \sigma) = \int_0^\infty \left(\prod_{j=1}^{j_i} \frac{e^{V(y_{ij}; F^k, \beta, \cdot)}}{\sum_{s \in \mathcal{Y}_{ij}} e^{V(s; F^k, \beta, \cdot)}} \right) f(\beta; \zeta, \sigma) d\beta \tag{2..1}$$

where $f(\cdot)$ is the density function of the log normal distribution and y_{ij} is the allocation chosen by spectator i from the choice set $\mathcal{Y}_{ij} = \{0, 25, \ldots, X_{ij}\}$ that spectator i faces in the redistribution decision j.

To calculate the total likelihood contribution of spectator i, we take the weighted sum of the conditional likelihood $L_{i,k}$

$$L_i(\lambda^L, \lambda^E, \lambda^{CE}, \zeta, \sigma) = \sum_{k \in \{L,E,CE\}} \lambda^k L_{i,k} \tag{2..2}$$

where λ^k is the population share of spectators with fairness view $k \in \{L, E, CE\}$. k^L corresponds to *Libertarians* view, k^E corresponds to *Egalitarians* view, and k^{CE} corresponds to *Choice Egalitarian* view. Finally, the total log-likelihood is obtained by taking the sum of the log of the total likelihood contributions of each spectator.

Parameters are estimated by simulated maximum likelihood with 100 Halton draws for each observation (Train, 2009). One population share and its standard error are calculated residually. The estimation is performed in R using the BFGS method with mle2 function (bbmle package).

APPENDIX C: INSTRUCTIONS

Instructions for stakeholders (MTurk)

The study comprises two stages. Please find below the instructions for stage 1. Stage 2 of the study concerns the distribution of earnings from stage 1. Details of the second stage will be provided after the first stage is completed.

STAGE 1

If you complete the study, you will earn a **fixed amount of $0.60 plus a bonus** that depends on your choices. All earnings are expressed in tokens that will be converted into real money at the end of the study ($1=300 tokens).

The study will take about 10 minutes to complete (including the time for reading the instructions). You will receive a code to collect your payment via MTurk upon completion.

YOUR TASK

You will face **five decisions**. In each decision, you have to choose between two options: option A and option B (see Table B-1).

Table B-1

Decision	**Option A**	**Option B**
1	100 for sure	800 with prob 50% or 0 with prob 50%
2	200 for sure	800 with prob 50% or 0 with prob 50%
3	300 for sure	800 with prob 50% or 0 with prob 50%
4	400 for sure	800 with prob 50% or 0 with prob 50%
5	500 for sure	800 with prob 50% or 0 with prob 50%

OPTION A IS SAFE. The safe amount changes in each decision: it ranges from 100 tokens in decision 1 to 500 tokens in decision 5.

OPTION B IS RISKY. Option B is the same for all five decisions. If you select option B, you have a 50% probability of earning 0 tokens and a 50% probability of earning 800 tokens. [**Nature only:** If you choose option B for a given decision, the computer will resolve the lottery. The outcome will be reported in the end.] *[**Coin & Ext. only:***

If you choose option B for a given decision, after the last decision you have to flip a coin. If the coin lands face-up on Tails you get 0 tokens, if it lands face-up on Heads you get 800 tokens. Please notice that you have to report the outcome of the coin flip truthfully. You may also use justflipacoin.com to virtually flip a coin.]

[***Ext. only:***
You will be matched with another worker. The other worker will be informed about the rules and will have to answer control questions.
The bonus of this other worker will be either 0 or 800 tokens. The bonus is determined as follows:

- ***If you choose Option A (safe),*** *the bonus for the other worker is randomly determined by the computer (each outcome has the same probability of being drawn). That is, your choice will not affect the bonus of the other worker;*
- ***If you choose Option B (risky),*** *your bonus and the bonus of the other worker will depend on your coin flip:*
  - *If you report that the coin landed face-up on TAILS, you earn 0 tokens and the other worker earns 800 tokens;*
  - *If you report that the coin landed face-up on HEADS, you earn 800 tokens and the other worker earns 0 tokens*

]

At the end of the study, the computer will randomly select **one** decision that will be relevant for stage 2. Further details about stage 2 will be provided later.

If the instructions are clear, please enter your MTurk worker ID and proceed to the control questions for stage 1.

worker ID

$\Longrightarrow$ ——— *new screen* ——— $\Longleftarrow$

Suppose you chose Option A (safe) in decision 2. What is **your** outcome in this decision?

☑ The outcome is 200 for sure

☐ The outcome is 800 for sure

☐ The outcome can be either 0 or 800

☐ There is no bonus for sure

⟹ ——— *new screen* ——— ⟸

Suppose you chose Option B (risky) in decision 4. What is **your** outcome in this decision?

- ☐ The outcome is 400 for sure
- ☐ The outcome is 800 for sure
- ☑ *[**Coin & Ext. only:** You will toss a coin to determine]* [**Nature only:** A random draw of the computer will determine] the outcome that can be either 0 or 800
- ☐ There is no bonus for sure

⟹ ——— *new screen* ——— ⟸

Ext. only *Suppose you chose Option B (risky) in decision 4. You flip a coin and self-report HEADS. What is the outcome?*

- ☐ *You will receive 0 and the other worker 800*
- ☑ *You will receive 800 and the other worker 0*
- ☐ *Both, you and the other worker will receive 800*
- ☐ *Both, you and the other worker will receive 0*

⟹ ——— *new screen* ——— ⟸

Ext. only *Suppose you chose Option A (safe) in decision 4. What is the outcome?*

- ☐ *You will receive 0 and the other worker 400*
- ☐ *You will receive 400 and the other worker 0*
- ☐ *You will receive 800 and the other worker 0*
- ☑ *You will receive 400 and the bonus of the other worker will be determined by the computer*

⟹ ——— *text in case of wrong answer* ——— ⟸

⟹ ——— *new screen: sample screen for decision 1* ——— ⟸

Your answer was incorrect. If you fail next time, your HIT cannot be accepted.

DECISION 1 Please decide between option A and B.

- ○ Option A (safe): 100 tokens
- ○ Option B (risky): 0 or 800 tokens

Please remember that in option B the two outcomes (0 and 800 tokens) are equally likely. [**Nature only:** If you choose option B, the computer will resolve the lottery. The outcome will be reported in the end.]
*[**Ext. only:** If you choose option A you will receive that amount for sure while the other worker will determine his outcome independently.]* ***Coin and Ext.:*** *If you choose option B after the last decision you have to flip a coin. If it lands face-up on Tails you will get 0 tokens [**Ext. only:** and the other worker 800]. If it lands face-up on Heads you get 800 tokens [**Ext. only:** and the other worker 0]. Please notice that you have to report the outcome of the coin flip truthfully. You may also use justflipacoin.com to virtually flip a coin.*

$\Longrightarrow$ ——— *new screen: sample screen for result 1* ——— $\Longleftarrow$

DECISION 1 – COIN FLIP *[COIN AND EXT. ONLY]* You chose Option B (risky) in decision 1.
Please flip a coin and indicate the outcome. If it lands on Heads you receive 800 tokens, if it lands on Tails you will receive 0 tokens.
Please report your answer truthfully.

- ○ Tails: 0 tokens [***Ext. only:*** *for you and 800 tokens for the other worker*]
- ○ Heads: 800 tokens [***Ext. only:*** *for you and 0 tokens for the other worker*]

You may also use justflipacoin.com to virtually flip a coin.

$\Longrightarrow$ ——— *new screen: stage 2 and beliefs* ——— $\Longleftarrow$

STAGE 2

Thank you for completing stage 1 of the study. We will now explain stage 2. In stage 2 you will be randomly matched with another worker (partner, henceforth), who has completed the exact same study as you have. One of the 5 decisions will be randomly selected.

A third person will be informed about the assignment, the rules, your choice and your partner's choice in the selected decision. In case you or your partner chose option B (risky), the third person is also informed about the *[**Coin & Ext. only:** self-reported]* outcome of the [**Nature only:** random draw.] *[**Coin & Ext. only:** coin toss.]*

The third person will be given the opportunity to redistribute the total amount of tokens generated between you and your partner. The total amount redistributed to you and to your partner must be equal to the sum of tokens you two got in the selected decision. The third person can leae everything as it is, or he/she can give some of your tokens to your partner or vice-versa. The redistribution done by the third person will determine your bonus for the present assignment. You will receive your bonus within one week from the completion of the assignment.

Please answer the following questions for stage 2:

In decision 3 you selected **Option [A safe/B risky - and your *[Coin &Ext. only: self-reported]* outcome was [XX] tokens**]. Suppose your partner chose **Option A (safe)** for decision 3. A third person will now redistribute the sum of tokens, which equals [SUM], between you and your partner. How do you think the tokens will be redistributed?

(NOTE: The distributed tokens must sum up to [SUM] tokens.)

Amount of tokens you will receive: [blank]

Amount of tokens your partner will receive: [blank]

In decision 3 you selected **Option [A safe/B risky -and your *[Coin & Ext. only: self-reported]* outcome was [XX] tokens**]. Suppose your partner chose **Option B (risky) with the *[Coin & Ext. only: self-reported]* outcome of 0 tokens** for decision 3. A third person will now redistribute the sum of tokens, which equals [SUM], between you and your partner. How do you think the tokens will be redistributed?

(NOTE: The distributed tokens must sum up to [SUM] tokens.)

Amount of tokens you will receive: [blank]

Amount of tokens your partner will receive: [blank]

In decision 3 you selected **Option [A safe/B risky - and your *[Coin & Ext. only: self-reported]* outcome was [XX] tokens**]. Suppose your partner chose **Option B (risky) with the *[Coin & Ext. only: self-reported]* outcome of 800 tokens** for decision 3. A third person will now redistribute the sum of tokens, which equals [SUM], between you and your partner. How do you think the tokens will be redistributed?

(NOTE: The distributed tokens must sum up to [SUM] tokens.)

Amount of tokens you will receive: [blank]

Amount of tokens your partner will receive: [blank]

$\Longrightarrow$ ——— *new screen: validation code* ——— $\Longleftarrow$

VALIDATION CODE. **Please enter this code <*code here*> in the MTurk HIT to complete the study.**

IMPORTANT: you need to enter this code to collect your payments.

$\Longrightarrow$ ——— *new screen: last screen* ——— $\Longleftarrow$

Thank you for completing this study. Your answers were transmitted. You may close the browser, window or tab now.

Instructions for idle stakeholders (***Ext. treatment only***)

If you complete the study, you will earn a **fixed amount of $0.30 plus a bonus** which can be either 800 tokens or 0 tokens. All earnings are expressed in tokens that will be converted into real money at the end of the study ($1=300 tokens).

In this HIT there are two types of roles: worker 1 and worker 2. You have been assigned to the role of **worker 2**.

YOUR TASK:

Your task is to read worker 1's instructions. The instructions will give a detailed explanation of the task carried out by worker 1 and are also important for you as they explain how your bonus – either 0 tokens or 800 tokens – is determined. It is important that you read the instructions for worker 1 carefully, as you will be asked to answer a number of questions concerning these instructions.
The study will take about 10 minutes to complete (including the time for reading the instructions). You will receive a code to collect your payment via MTurk upon completion. You will only be able to receive the code if you answer all questions correctly.

⟹ ——— *new screen* ——— ⟸

Please read the instructions for worker 1 below and then proceed to the next page to answer the questions.

INSTRUCTIONS FOR WORKER 1

Worker 1 will face **five decisions**. In each decision, he has to choose between two options: option A and option B (see Table B-2).

OPTION A IS SAFE. The safe amount changes in each decision: it ranges from 100 tokens in decision 1 to 500 tokens in decision 5.

Table B-2

Decision	Option A	Option B
1	100 for sure	800 with prob 50% or 0 with prob 50%
2	200 for sure	800 with prob 50% or 0 with prob 50%
3	300 for sure	800 with prob 50% or 0 with prob 50%
4	400 for sure	800 with prob 50% or 0 with prob 50%
5	500 for sure	800 with prob 50% or 0 with prob 50%

OPTION B IS RISKY. Option B is the same for all five decisions. If worker 1 selects option B, he has a 50% probability of earning 0 tokens and a 50% probability of earning 800 tokens. If he chooses option B for a given decision, after the last decision he has to flip a coin. If the coin lands face-up on Tails he gets 0 tokens, if it lands face-up on Heads he gets 800 tokens. Please notice that he has to report the outcome of the coin flip truthfully. He may also use justflipacoin.com to virtually flip a coin.

Worker 1 will be matched with another worker - worker 2 (you). Worker 2 will be informed about the rules and will have to answer control questions.
The bonus of worker 2 will be either 0 or 800 tokens. The bonus is determined as follows:

- **If worker 1 chooses Option A (safe),** the bonus for worker 2 is randomly determined by the computer (each outcome has the same probability of being drawn). That is, his choice will not affect worker 2's bonus;
- **If worker 1 chooses Option B (risky),** his bonus and worker 2's bonus will depend on your coin flip:
  - If worker 1 reports that the coin landed face-up on TAILS, he earns 0 tokens and worker 2 earns 800 tokens;
  - If worker 1 reports that the coin landed face-up on HEADS, he earns 800 tokens and worker 2 earns 0 tokens

⟹ ——— *new screen* ——— ⟸

Suppose worker 1 chose Option A (safe) in decision 2. What is worker 1's outcome in this decision?

☑ The outcome is 200 for sure

☐ The outcome is 800 for sure

☐ The outcome can be either 0 or 800

☐ There is no bonus for sure

$\Longrightarrow$ ——— *new screen* ——— $\Longleftarrow$

Suppose worker 1 chose Option B (risky) in decision 4. What is worker 1's outcome in this decision?

☐ The outcome is 400 for sure

☐ The outcome is 800 for sure

☑ Worker 1 will toss a coin to determine the outcome that can be either 0 or 800

☐ There is no bonus for sure

$\Longrightarrow$ ——— *new screen* ——— $\Longleftarrow$

Suppose worker 1 chose Option B (risky) in decision 4. Worker 1 flips a coin and self-reports HEADS. What is the outcome?

☐ Worker 1 will receive 0 and worker 2 receives 800

☑ Worker 1 will receive 800 and worker 2 recieves 0

☐ Both, worker 1 and worker 2 will receive 800

☐ Both, worker 1 and worker 2 will receive 0

$\Longrightarrow$ ——— *new screen* ——— $\Longleftarrow$

Suppose worker 1 chose Option A (safe) in decision 4. What is the outcome?

☐ Worker 1 will receive 0 and worker 2 receives 400

☐ Worker 1 will receive 400 and worker 2 receives 0

☐ Worker 1 will receive 800 and worker 2 receives 0

☑ Worker 1 will receive 400 and worker 2's bonus will be determined by the computer

$\Longrightarrow$ ——— *text in case of wrong answer* ——— $\Longleftarrow$

$\Longrightarrow$ ——— *new screen* ——— $\Longleftarrow$

> Your answer was incorrect. If you fail next time, your HIT cannot be accepted.

Your bonus. Your bonus is ***XX tokens***.

$\Longrightarrow$ ——— *new screen: validation code* ——— $\Longleftarrow$

VALIDATION CODE. **Please enter this code <*code here*> in the MTurk HIT to complete the study.**

IMPORTANT: you need to enter this code to collect your payments.

$\Longrightarrow$ ——— *new screen: last screen* ——— $\Longleftarrow$

Thank you for completing this study. Your answers were transmitted. You may close the browser, window or tab now.

Instructions for Spectators[1]

Welcome. The purpose of this study is to investigate how people make decisions. From now until the end of the study, any communication with other participants is not allowed. If you have a question, please raise your hand and one of us will come to your desk to answer it. [**Nature & *Self-Report sessions without beliefs* only:** Upon completion of the study, you will receive a payment of €10, including €4 show-up fee.] [***Ext. & Self-Report sessions with beliefs only:*** *For showing-up on time, you will receive €4. This study comprises three parts and you can earn additional money during the study. Payments will be made upon completion of the study, anonymously, and in cash.*]

[***Ext. & Self-Report sessions with beliefs only:***

Instructions for Part 1

For Part 1, you will receive a fixed payment of €6.]

OVERVIEW. You will be presented with 20 decisions, one after the other. In each decision, your task is to decide how to redistribute the money between an ORANGE and a BLUE player. One of these decisions will have real monetary consequences for two individuals that we recruited via an international online marketplace to conduct an assignment. We will first explain in detail the task we gave to the individuals [***Ext. only:***, *ORANGE and BLUE,*] who participated in the online assignment. After that, we will provide you with further information about your task.

ONLINE ASSIGNMENT. A few days ago we recruited participants via an international online marketplace to conduct an assignment. They were offered a fixed participation compensation of $0.60. [***Ext. only:*** *We will now describe the assignment for ORANGE and BLUE. Each ORANGE and BLUE player was also matched with a GREEN player, whose task will be described later on.*]

The [***Ext. only:*** *ORANGE and BLUE player's*] assignment consisted of 5 decisions. In each decision, they had to choose between two options: option A and option B (see Table 1). All values in the assignment were expressed in tokens. Tokens are exchanged at the rate of $1=300 tokens. Please notice that the amount of money at stake is above the average amount for similar tasks in the same online marketplace.

[1] Translated from German. Original instructions are available upon request from the authors.

Table B-3: Online decisions

Decision	Option A	Option B
1	100 for sure	800 with prob. 50% or 0 with prob. 50%
2	200 for sure	800 with prob. 50% or 0 with prob. 50%
3	300 for sure	800 with prob. 50% or 0 with prob. 50%
4	400 for sure	800 with prob. 50% or 0 with prob. 50%
5	500 for sure	800 with prob. 50% or 0 with prob. 50%

Option A is safe. The safe amount changed in each decision: it ranged from 100 tokens in decision 1 to 500 tokens in decision 5.

Option B is risky. Option B was the same for all five decisions. If option B was selected, the participant had a 50% probability of earning 0 tokens and a 50% probability of earning 800 tokens. [**Nature only:** If a participant chose option B for a given decision, the computer resolved the lottery at the end of the assignment.] *[**Self-Report only:** If a participant chose option B, for a given decision, he/she was asked to flip a coin. If the coin landed face-up, on Tails, the outcome was 0 tokens; if it landed face-up on Heads, the outcome was 800 tokens. Participants were asked to report the outcome of the coin flip truthfully and were given a link to flip a coin virtually in case they did not have a coin with them (see sample screen shot in Figure B-1).]*

[***Ext. only:*** *Each ORANGE and BLUE player was matched with a GREEN player. The GREEN player was informed about the rules and had to answer the same control questions but did not make any decisions. The outcome for GREEN was either 0 or 800 and it was determined as follows:*

- *If ORANGE or BLUE chose **Option A (safe)** it did not affect GREEN's outcome. GREEN's outcome (0 or 800) was then randomly determined by the computer (each outcome has the same probability of being drawn).*
- *If ORANGE or BLUE chose **Option B (risky)** he affected GREEN's outcome:*
  - *If ORANGE or BLUE reported Tails, his outcome was 0 tokens and GREEN's outcome was 800 tokens.*
  - *If ORANGE or BLUE reported Heads, his outcome 800 tokens and GREEN' outcome was 0 tokens.*

]

Figure B-1: Sample screen shot from the online assignment **(Self-Report only)**

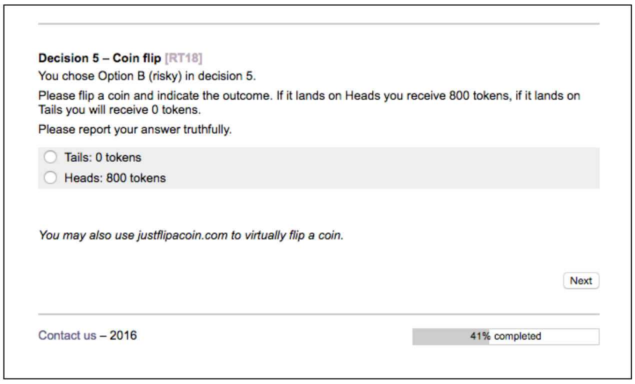
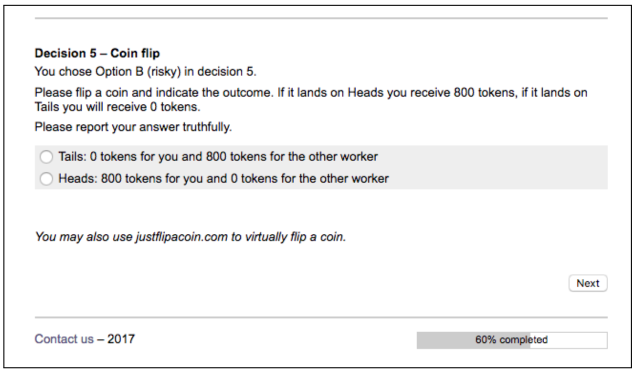


Figure B-2: Sample screen shot from the online assignment **(Ext. only)**



Participants were allowed to take part in the assignment only if had they correctly answered all control questions. After collecting all the data, we randomly formed pairs and selected at random one of the 5 decisions. After completing the assignment, participants were told that a third person would be informed about the rules and the outcome of the assignment, and would be given the opportunity to redistribute the earnings and thus determine how much they were paid for the assignment.

YOUR TASK. You are the third person and we now want you to choose whether to redistribute the tokens for the assignment between [**Nature &** ***Self-Report*** **only:** two people] [***Ext. only:*** *ORANGE and BLUE*]. Your decision is completely anonymous. The [***Ext. only:*** *ORANGE and BLUE*] people who participated in the online assignment will receive the payment that you choose for them within a few days, but will not receive any further information.

Figure B-3: Sample screen shot (*Self-Report* treatment)

Notes: In the *Nature* treatment, the sentence "ORANGE reported HEADS" was not displayed.

Figure B-4: Sample screen shot (*Ext.* treatment)

Figure B-3 shows a sample decision screen. In the upper part of the screen, you can see the initial situation for the ORANGE and BLUE player. For each player, you can see whether they chose option A (safe) or option B (risky). In each decision, you will be able to see the amount of tokens yielded by the safe option A. In this example, the safe level is 500 tokens. [***Ext. only:*** *You can also see who determined the outcome for GREEN 1 and GREEN 2.*]

In the example in Figure B-3, ORANGE chose option B and BLUE chose option A. Recall that the outcome of option B is determined [**Nature only:** by a random draw of the computer and both outcomes –0 and 800 tokens– have the same probability of being randomly selected.] *[**Self-Report & Ext. only:** by the toss of a coin. Participants [**Ext. only:** ORANGE and BLUE,] in the online assignment, were asked to toss a coin and self-report the outcome truthfully. If a participant reported TAILS the outcome of option B was 0 tokens, [**Ext. only:** and 800 tokens for GREEN]. If the participant reported HEADS the outcome of option B was 800 tokens [**Ext. only:** and thus GREEN's outcome was 0 tokens].]* In this example, the outcome for ORANGE was 800 tokens *[**Self-Report & Ext. only:** – as he reported HEADS]* – [***Ext. only:*** *and hence the outcome for GREEN 1 was 0. BLUE chose option A (safe) yielding an outcome of 500 tokens. Thus GREEN 2's outcome was determined by a random draw of the computer.*]

In the central part of the screen you can see the sum of the tokens by ORANGE and BLUE players. In the example, the sum of tokens is 1300. **Your task is to decide whether and how to redistribute the total amount of tokens between ORANGE and BLUE.** You can choose any positive amount in steps of 25 tokens, as long as you redistribute all tokens. In our example, the sum of what you give to ORANGE and BLUE must be exactly 1300 tokens.

You have to make 20 decisions and one decision will be relevant – that is, it will have actual monetary consequences – for two individuals who have completed the online assignment. You will not know in advance which decision is relevant for the earnings of other individuals. This means that you have to pay attention to every decision.

Before starting, please answer a few control questions.

⟹ ——— *new section* ——— ⟸

Control Questions

1. Suppose a participant in the online assignment chose Option A (safe) in decision 2 (see Table B-2). What is the outcome of this decision?

   ☑ The outcome is 200 for sure.

   ☐ The outcome is 800 for sure.

   ☐ The outcome can be either 0 or 800.

   ☐ The outcome is 0 for sure.

2. Suppose a participant in the online assignment chose Option B (risky) in decision 4. What is the outcome of this decision?

   ☐ The outcome is 400 for sure.

   ☐ The outcome is 800 for sure.

   ☑ *[**Self-Report & Ext. only:** The participant had to toss a coin to determine]* [**Nature only:** A random draw of the computer determined] the outcome that can be either 0 or 800.

   ☐ The outcome is 0 for sure.

3. You are the third person who has to choose how to redistribute the tokens from the assignment

   ☐ Your identity will be revealed to the participant in the online assignment.

   ☑ One of your decisions will have real monetary consequences for two participants in the online assignment.

   ☐ You have to make only one decision.

4. Suppose ORANGE chose option A in decision 3. BLUE, instead, chose option B and [**Nature only:** the computer selected at random the low amount.] *[**Self-Report & Ext. only:** self-reported TAILS.]* What is the total number of tokens earned in this situation? (e.g., the sum of the tokens by ORANGE and BLUE)

   ☐ The total number of tokens is 200.

- ☐ The total number of tokens is 800.
- ☐ The total number of tokens is 1100.
- ☑ The total number of tokens is 300.

5. Suppose the total number of tokens earned in a situation is 1600.
   - ☐ You can give 0 tokens to both ORANGE and BLUE.
   - ☑ The sum of the tokens you give to ORANGE and BLUE has to be exactly 1600.
   - ☐ The sum of the tokens you give to ORANGE and BLUE can be larger than 1600.
   - ☐ The sum of the tokens you give to ORANGE and BLUE can be smaller than 1600.

Ext. only:

6. *Suppose a participant chosen option B (risky) and self-reported HEADS. What is the outcome of this decision?*
   - ☐ *The participant receives 0 and GREEN receives 800.*
   - ☐ *The participant and GREEN both receive 800.*
   - ☐ *The participant and GREEN both receive 0.*
   - ☑ *The participant receives 800 and GREEN receives 0.*

7. *Suppose a participant chose option A (safe) in decision 4. What is the outcome of this decision?*
   - ☑ *The participant receives 400 and GREEN's outcome will be randomly determined by the computer.*
   - ☐ *The participant receives 800 and GREEN 0.*
   - ☐ *The participant receives 400 and GREEN 0.*
   - ☐ *The participant receives 0 and GREEN 400.*

8. *Which of the following statements is not correct?*
   - ☐ *You will redistribute the sum of tokens between ORANGE and BLUE.*
   - ☑ *GREEN always flips a coin to decide his outcome.*

- ☐ *All participants were informed about the rules and had to answer control questions.*
- ☐ *GREEN's outcome is either 0 or 800.*

⟹ ——— *new section* ——— ⟸

Final Questionnaire

1. Gender
   - ☐ Male
   - ☐ Female

2. Age: ____

3. Field of study
   - ☐ Medicine
   - ☐ Physics, Biology, Mathematics
   - ☐ Computer science
   - ☐ Social sciences
   - ☐ Psychology
   - ☐ Other

4. Please indicate where you were born
   - ☐ Schleswig-Holstein
   - ☐ Mecklenburg-Vorpommern
   - ☐ Hamburg
   - ☐ Bremen
   - ☐ Niedersachsen
   - ☐ Hessen
   - ☐ Nordrhein-Westfalen
   - ☐ Rheinland-Pfalz

- ☐ Saarland
- ☐ Baden-Württemberg
- ☐ Bayern
- ☐ Brandenburg
- ☐ Berlin
- ☐ Sachsen
- ☐ Sachsen-Anhalt
- ☐ Thüringen
- ☐ Outside Germany

5. In political matters, people talk of *the left* and *the right.* How would you place your views on this scale, generally speaking?

	○	○	○	○	○	○	○	○	○	○	
Left	1	2	3	4	5	6	7	8	9	10	*Right*

6. We now want you to indicate to what extent you agree with the following statement. 1 means that you agree completely with the statement on the left, 10 means that you agree completely with the statement on the right, and the numbers in between indicate the extent to which you agree or disagree with the statements.

A society should aim to equalize incomes.	*A society should **not** aim to equalize incomes*

○	○	○	○	○	○	○	○	○	○
1	2	3	4	5	6	7	8	9	10

7. Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?

- ☐ Most people can be trusted.
- ☐ Need to be very careful.

In addition, subjects answered the 60-item version of the HEXACO Personality Inventory-Revised Test (http://hexaco.org/hexaco-inventory).

Instructions for beliefs and risk aversion in the lab[2]

Instructions for Part 2

In this part, we ask you to guess what people chose in the online assignment explained before.

YOUR TASK. Please consider the decision between Option A that yields 300 tokens for sure and Option B that yields 800 with a probability of 50% and 0 with a probability of 50%. You will have to answer the following two questions:

- **Question 1:** What is the percentage of participants in the online assignment who chose Option B (risky)?
- **Question 2:** Consider now the online participants who have chosen Option B: what is the percentage of participants who reported *Heads*? Please recall that *Heads* yielded 800 tokens for the participant and *Tails* 0 tokens for the participant.

YOUR PAYMENT. You can earn a substantial amount of money based on the accuracy of your guess, as reported in Table B-4. If your guess is correct, you can earn €22. If your guess deviates from the true value by 5 percentage points (plus or minus), you can earn €20.90. If your guess deviates by more than 21 percentage points, you can receive €2 for this part.

Table B-4: Your payment

deviation in percentage points	**payment**
exact number	€22.00
between 1 and 5	€20.90
between 6 and 10	€17.60
between 11 and 15	€12.10
between 16 and 20	€4.40
over 21	€2.00

[2] Translated from German. Original instructions are available upon request from the authors. This set of instructions was used only in two Self-Report and two Externalities sessions, for a total of 120 participants.

After everyone has answered both questions, six participants will be chosen at random for payment for this part. The selected participants will be paid for one of the two questions, selected at random. Since you do not know in advance who and which question will be chosen, it is important that you pay attention to both answers.

Instructions for Part 3

YOUR TASK. Now, please select one option out of six different options. The six different options are displayed in Figure B-5. You must select one and only one of these gambles.

Figure B-5: Options and payments

OPTION	Wenn GRÜN gewählt wird	Wenn ROT gewählt wird
I	28 Euro	28 Euro
II	36 Euro	24 Euro
III	44 Euro	20 Euro
IV	52 Euro	16 Euro
V	60 Euro	12 Euro
VI	70 Euro	2 Euro

OPTIONS AND EARNINGS. Each option has two possible colors (green and red), each with a 50% probabilitie of occurring. Your earnings for this part of the study will be determined by:

- Which of the six options you select; and
- Which of the two possible colors (green or red) occurs

For example, if you select Option 4 and green occurs, you earn €52. If red occurs, you earn €16.

At the end of this task, the computer will randomly select one participant for payment. The computer will then randomly draw one of the two colors (green or red) and the earnings for the selected participant will be determined. Please remember that for every option, each color has a 50% chance of occurring.

Instructions for beliefs in the classroom[3]

Instructions

Welcome. The purpose of this study is to investigate how people make decisions. *[**Experimenter only** If you have a question please raise your hand, after the instructions have been read and one of us will come to your desk to answer it. Your answers will be treated anonymously.]* More specifically, you will be asked to guess the results of a previous study. *[**Experimenter only** We will now explain both the previous task – an online assignment – and your task in detail.]* From now until the end of the study, any communication with other participants is not allowed.

ONLINE ASSIGNMENT. We recruited over 100 participants via an international online marketplace and asked them to make a series of decisions. Participants had to choose between:

- **Option A (safe)** yields a safe payment, with the amount specified on your decision sheet;
- **Option B (risky)** yields 800 tokens with a 50% probability and 0 tokens with a 50% probability. If a participant chose option B he/she was asked to flip a coin and self-report the result:
  - if the coin landed face-up on **Heads** the outcome was 800 tokens;
  - if the coin landed face-up on **Tails** the outcome was 0 tokens.

Participants were asked to report the outcome of the coin flip truthfully. Participants were aware that a self-reported coin toss would resolve the outcome for Option B before choosing between the two options. All earnings were expressed in tokens and exchanged at the rate of $1=300 tokens.

YOUR TASK. We ask you to guess what people did in the online assignment. You will have to answer the following two questions:

- **Question 1:** What is the percentage of participants who chose Option B (risky)?
- **Question 2:** Consider now the participants who have chosen Option B: what is the percentage of participants who reported *Heads*? Please recall that *Heads* yielded 800 tokens and *Tails* 0 tokens.

[3] Translated from German. Original instructions are available upon request from the authors. A total of 289 students participated in the classroom experiment.

YOUR PAYMENT. You can earn a substantial amount of money based on the accuracy of your guess, as reported in Table B-5. If your guess is correct you can earn €22.00. If your guess deviates from the true value by 5 percentage points (plus or minus), you can earn €20.90. If your guess deviates by more than 21 percentage points, you get €2.00.

Table B-5: Your payment

deviation in percentage points	**payment**
exact number	€22.00
between 1 and 5	€20.90
between 6 and 10	€17.60
between 11 and 15	€12.10
between 16 and 20	€4.40
over 21	€2.00

After everyone has answered both questions, one out of every 20 students will be chosen at random for payment. The selected students will be paid for one of the two questions, selected at random. Since you do not know in advance who and which question will be chosen, it is important that you pay attention to both answers. You can now make your decisions. Please read the information on the decision sheet carefully.

Decision sheet

Safe level for Option A = 100 tokens for sure

Participants in the online assignment had to make a decision between Option A and Option B.

Option A (safe)	**Option B (risky)**
100 tokens for sure	**800** tokens if Heads **0** tokens if Tails

Please answer the following questions

Question 1: What is the percentage of participants who chose Option B (risky)?

_____ %

Please write an integer number between 0 and 100

Question 2: Consider now the participants who have chosen Option B: what is the percentage of participants who reported *Heads*? Please recall that *Heads* yielded 800 tokens and *Tails* 0 tokens.

_____ %

Please write an integer number between 0 and 100

Gender:

☐ Male

☐ Female

Field of study:

☐ Economics

☐ Economics majoring in sociology

☐ Sociology

☐ Math

☐ Other

CHAPTER 3

ECONOMIC BEHAVIOR OF CHILDREN AND ADOLESCENTS – A FIRST SURVEY OF EXPERIMENTAL ECONOMICS RESULTS

Abstract

About 15 years ago, economic experiments with children and adolescents were considered as an extravagant niche of economic research. Since then, this type of research has exploded in scope and depth. It has become clear that studying the development of economic behavior and its determinants is important to understand economic behavior of adults and to provide a basis for potential policy interventions with respect to economic behavior in childhood and adolescence. Given the huge increase of papers, we provide the first overview of economic experiments with children and adolescents. We focus on the following aspects: rationality of choices, risk preferences, time preferences, social preferences, cooperation, and competitiveness. All of these aspects are analyzed with respect to the influence of age and gender, and we also consider the role of socio-economic status or interventions.

This chapter is joint work with Daniela Glätzle-Rützler and Matthias Sutter.

3.1 INTRODUCTION

In the late 1990ies, Bill Harbaugh, Kate Krause and co-authors pioneered what has become a very lively field of research since then, namely the experimental study of economic behavior by children and adolescents. Their early work has set examples of how to run experiments with children and adolescents and which topics can be studied with young experimental subjects, covering, among others, risk taking, social preferences, rationality of choices, or bargaining (Harbaugh and Krause, 2000; Harbaugh et al., 2001, 2002, 2003a,b, 2007). While in the early years after their seminal contributions the number of experimental papers with children was still fairly small and it was easy to keep track of the whole literature, in the past ten years the number of experimental papers has flourished tremendously, if not to say that it has exploded. For this reason, we think it is time to write a first survey of the main topics in this field of research in experimental economics.[1] Given the explosion of papers, this survey will not be able to mention all of them, and possibly we have overlooked some papers, but it will try to organize the literature along different dimensions of economic behavior, and present the general pattern of results that one can see from reading the literature. In doing so, we will address children's and adolescents' (i) rationality of choices, (ii) time preferences, (iii) risk preferences, (iv) social preferences, including allocation games, bargaining games and games of cooperation, and (v) competitiveness. We will put the main emphasis in each dimension on the influence of age (typically from 3 year-olds to 18 year-olds) and gender on economic decisions of children and adolescents. On top of that, we consider further determinants of economic behavior, such as socio-economic status of parents, or the social context of interaction (like in-group/out-group scenarios). The selection of the aforementioned determinants of economic preferences is based on the fact that these are the most common studied

[1] Please note that this survey focuses primarily on studies conducted in the field of experimental economics, aiming to give a detailed insight into economic research conducted with children. Literature in the field of psychology discussing experimental settings with children will not be the main focus (see e.g., Warneken, 2018 for a survey on psychology literature on cooperation in children).

predictors of children's economic behavior. Moreover, we will briefly address very recent studies that have run policy interventions to influence children's and adolescents' economic behavior. The latter type of studies rests on the knowledge of how economic behavior of children and adolescents looks and how it develops with age, for which reason this survey puts most weight on a descriptive analysis of children's and adolescents' economic behavior.

In the early years of experimental research with children and adolescents, editors and referees were often skeptical as to what could or should be learned from examining the economic behavior of children and adolescents.[2] Today, this type of research seems to be accepted as an established research field, also at the top journals, for a variety of reasons. First, studying behavior of children and adolescents can reveal whether economic behavior develops in characteristic patterns in the course of life. Similar to psychological research on the development of moral judgments, for instance, economic research is interested in whether fairness preferences, risk attitudes, impatience, rational choice behavior or competitive preferences develop in certain ways. Most behavioral models of social preferences, for instance (see, e.g., Fehr and Schmidt, 1999a; Bolton and Ockenfels, 2000; Charness and Rabin, 2002), have been based on experimental evidence from university students in their early 20ies. Research with children and adolescents can reveal whether such models (of non-standard preferences) also apply to pre-adulthood or whether the behavioral patterns of adults are the consequence of a directional development with age. Knowing more about such a potential development is a precondition for possible policy interventions that might try to promote particular types of behavior (such as patience with respect to attaining education, or avoiding conflicts through a mutual understanding of fairness and social norms). Second, from the viewpoint of economic theory it is interesting to study whether children and adolescents are sophisticated decision makers that make rational decisions and are capable of applying fundamental game theoretic concepts (such as backward induction or mixed

[2] In the early 2000s, the first author of this survey got editorial decision letters that called experiments with children exotic research that would not help the scientific community in economics to better understand how markets work, for which reason the editors recommended to look for outlets in psychology.

strategy play) in their behavior. This would mean that such fundamental concepts are useful also to describe (at least parts) of young children's and adolescents' behavior. Third, and related to the first reason, the study of economic behavior of children and teenagers has gained importance through the research program of researchers like James Heckman who have studied how non-cognitive skills influence subjects' academic attainment, social and economic success, in particular on labor markets, or their health (Heckman and Rubinstein, 2001; Heckman, 2006; Heckman et al., 2006; Kautz et al., 2014). As non-cognitive skills develop dynamically, early childhood programs and interventions aiming at improving non-cognitive skills have proven efficient and beneficial for lifetime outcomes (Heckman, 2006; Kautz et al., 2014). Hence, improving our knowledge of economic preferences as an important subset of non-cognitive skills contributes to this strand of literature in the tradition of Heckman.

Of course, experimental research with children and adolescents sometimes differs in procedural and design details from experimental research with adults. For example, with pre-school children, it is typically impossible to use money as incentives. Rather, children can earn tokens that may be exchanged for small presents (like stickers, candies or toys) in an experimental shop after an experiment. Salience of rewards is often ensured by showing children the presents before commencement of the experiment (see, e.g., Harbaugh and Krause, 2000). Experiments are usually conducted in a controlled setting in schools or day-care centers, minimizing self-selection effects (see, e.g., Harbaugh et al., 2003a, Sutter et al., 2013). To ensure understanding and full attention experiments with very young children are conducted in a one-on-one setting where an experimenter explains to a single child the rules of the experiment, rather than explaining everything in front of a whole group of participants (see, e.g., Fehr et al., 2008). In general, economic experiments with children and adolescents, despite the aforementioned differences to experiments with adults, have become more and more standardized over the past 10 years, and this standardization is important for making the research better comparable. For instance, it is customary nowadays to check for correct understanding by adding control questions about comprehension or to preserve anonymity by the use of sliding walls (and by paying with sealed envelopes or handing over presents in opaque bags). Careful attention is also paid to avoid spreading information about the experiment among subjects who have not yet participated in it (which is sometimes not easy in schools or kindergarten). Most importantly, the standardization of the conduct of experiments has become much more advanced, for example by extensive training of experimental helpers to use the same wording and sequence of explanations when explaining experimental rules to children. Recently, Schunk et al. (2017) or Hermes et al. (2018) have gone one important step further in this respect. They ensure comprehension with the support of animated visual- and audio-aids. More precisely, they use tablets and headphones to ensure identical delivery of instructions to children which is a further advance in standardization. Despite these improvements in methodology, the studies presented in this survey differ here and there in design or procedural details, which means, for example, that cooperation rates in a prisoner's dilemma depend on the exact parameters. For this reason we are not going into the details of the quantitative results of the papers discussed here, but rather we focus on qualitative patterns of behavior across different studies. There is a highly recommended companion paper by List et al. (2018) that asks how experiments with children can inform economics and that presents a state-of-the-art overview about the different methods – and its pros and cons – to run economic experiments with children and adolescents. The interested reader is referred to this paper for details. In our survey here we focus on the main results of experimental studies with children and adolescents.

In each section, we start the survey by presenting the main results with respect to the influence of age and gender on economic preferences of children and adolescents. In most sections, we will also refer to the relation of socio-economic status (SES) of parents to economic preferences of their offspring. Finally, in some sections we are also going to look at additional factors, like cognitive factors, in-group favoritism, or policy interventions, that are discussed in the literature as potentially affecting the economic behavior of children and adolescents.

By and large, the literature reveals the following pattern of the relationship of age and economic preferences.[3] In early childhood, children are relatively self-centered with respect to social preferences, impatient, and risk tolerant. Only when getting older, in particular in adolescence, subjects' social preferences shift towards egalitarian and more efficiency-oriented behavior, and subjects become relatively more patient and more risk averse than in early childhood. With increasing age, parochialism becomes more pronounced as well as subjects favor more often members of their own social group. With respect to gender, we note large differences in competitiveness and risk taking, but less clear-cut differences in other behavioral domains. With respect to social preferences, girls tend to make more altruistic, cooperative, and inequality averse choices while boys are more concerned with efficiency and tend to be more selfish. Concerning socio-economic status (SES) of parents, it seems that children from a low SES-background are often less patient, less pro-social and less competitive than children from higher SES-backgrounds.

In the following, we present more details on economic behavior in childhood and adolescence in separate sections for different behavioral categories and preferences.[4] We start our survey with a brief section on the rationality of children's economic choices. This section is motivated to address the (unwarranted) concern that economic choices of children and adolescents might be random and would therefore contain no systematic insights. Section 3.2 will show that this is not the case, thus lending support that we can draw meaningful conclusions when looking at different preferences and behavioral patterns. Section 3.3 is devoted to time preferences, and section 3.4 to risk preferences. The sections 3.5, 3.6, and 3.7 shed light on social preferences,

[3]When we talk about developments with age, we refer to cross-sectional evidence from cohorts of different age. Too few studies have a panel structure that could speak to a within-subject development of economic behavior with increasing age.

[4]In each section, we concentrate on papers whose main research question addresses that section's preference. When a paper uses that section's preference only as a control variable to investigate yet another preference, we do not discuss such a paper in that section. To illustrate the procedure with an example: Studies on children's competitiveness often include a measurement of risk attitudes as a control variable to explain competitiveness. We are not going to include such papers in the section on risk preferences.

once from individual allocations tasks (section 3.5), once with respect to bargaining games (section 3.6), and once from games of cooperation (section 3.7). Section 3.8 deals with competitiveness. Finally, section 3.9 concludes the survey with a short summary and a discussion of open questions and an outlook about promising further avenues for experimental economics research with children and adolescents.

3.2 RATIONALITY OF CHILDREN'S CHOICES

Harbaugh et al. (2001) have been the first to show that already at an early age children are able to make decisions according to basic requirements of rationality. In their experiment with 7 and 11 year-olds, children of both age groups have to choose among different bundles to check whether their choices obey the generalized axiom of revealed preferences (GARP). It turns out that even the younger children are doing better than chance. Yet, it is true that the number of preference violations decreases with age. For instance, only 25% of 7 year-olds, but 60% of 11 year-olds make choices that are consistent with utility maximization. Compared to an adult subject pool there is no increase in the rationality of choices between the ages of 11 and 21, showing that rational behavior is prevalent already during adolescence and comparable to the level of adults.

In addition to making rational choices, the ability to form reasonable beliefs and make correct inferences further plays an important role in economic decision-making. Barash et al. (2018) let children draw from an urn with different compositions of colored balls in order to study the updating of beliefs. Younger children (aged 6-8) make decisions based on the previous outcome, using heuristics to determine their next move. With increasing age children and young adolescents start to take the entire series of draws into consideration, but frequently fall prey to the gambler's fallacy. From age 15 onwards, adolescents increasingly play a Bayesian strategy, in line with behavior shown by adults. Children hence move closer to behaving rationally as they grow older (Barash et al., 2018). Similarly, Brocas and Carrillo (2018a) let children make choices (in a non-interactive version of the game

"Connect 4") and check whether they are able to think ahead a few moves. Similar to Barash et al. (2018), they find that older children are better in anticipating future moves and thus to reason in a more sophisticated way. Apesteguia et al. (2018) study imitation of successful choices of others in children (aged 8-10) and adults (university students). They let their subjects repeatedly (over 10 rounds) choose to draw a ball from six different urns with different payoffs. In the baseline treatment subjects are not able to observe other participants while in the observation treatment subjects observe the outcome of a draw of another subject. The analysis of the baseline treatment reveals that all subjects are able to learn across rounds. However, adults have a steeper learning curve than children. Results from the observation treatment indicate that children, unlike adults, are not able to take advantage of the additional information received by observing others, meaning that they do not engage in rational imitation.

Strategic sophistication is another fundamental requisite of economic decision-making in interactive contexts. Being able to anticipate an interaction partner's rationality and incentives is crucial for success in strategic interactions. Brocas and Carrillo (2018b) study two-person games with 4-7 year-olds where the games are characterized by different levels of iterative complexity, i.e., the number of iterations before reaching the equilibrium of the game. They find that older children in their sample are significantly more likely to reach the equilibrium, but it is reassuring to note that younger children can also play equilibrium strategies when the iterative dominance is not too demanding. While Brocas and Carrillo (2018b) notice an age trend in the degree of strategic sophistication for 4-7 year-olds, Czermak et al. (2016) find hardly any changes in strategic sophistication in 10-17 year-olds. They let adolescents play two-person normal form games with different degrees of iterated dominance. Only with respect to the likelihood of eliminating dominated strategies, they observe older adolescents to have a higher likelihood, but all age groups are equally likely to reach the efficient (non-equilibrium) outcome of the normal form games. Moreover, the estimation of strategic types reveals no age differences either, and the distribution of types is similar to adult university students (Sutter, Czermak, and Feri, 2013). Related to the concept of strategic sophistication is the ability to apply backward induction. Again, this ability seems to increase with age. Brosig-Koch et al. (2015) examine how children aged 6-15 play so-called race games in which two players can move sequentially in choosing numbers (in a predefined interval) until a pre-specified number is reached. These games can be solved by backward induction, and first movers have an advantage. They find that first graders of age 6 perform significantly worse than older children, but that the differences across age diminish as subjects reach middle adolescence.

Neither rational decision making nor strategic reasoning differs significantly by gender for most ages. In early childhood females play equilibrium at higher rates than their male counterparts (Brocas and Carrillo, 2018b), while boys are better able to do backwards reasoning than girls until early adolescence, which is when the gender gap closes (Brosig-Koch et al., 2015; Czermak et al., 2016).

Cognitive ability affects the extent of rational choice and strategic behavior displayed by adolescents. A better math grade positively correlates with higher strategic sophistication and more rational choices, ultimately leading to higher payoffs in the experimental games (Harbaugh et al., 2001; Brosig-Koch et al., 2015; Czermak et al., 2016). Similarly, children who are assessed by their teachers as suitable for “Gymnasium” (the higher track in the Austrian school system) are shown to have a steeper learning curve than those predicted not to reach the “Gymnasium” (Apesteguia et al., 2018).

SUMMARY RATIONALITY OF CHOICES: Already young children show rational behavior to a considerable extent, obeying the laws of transitivity, and making (often) correct inferences about the partner’s rationality, and applying strategic reasoning in choosing their strategy in interactive games. All of these skills develop and become more pronounced from childhood to adolescence, implying that adolescents’ behavior assimilates more and more towards behavior observed in adults. Hence, children and adolescents do not make decisions randomly but are able to take strategic considerations and basic principles of rational behavior into account when making economic decisions, thus gradually reaching the behavioral patterns observed in adult subject pools.

Table 31: Rationality of children's choices

Authors	Subject Pool			Experimental Task	Results			
	age	sample size	country		age	gender	SES	other
Harbaugh et al. (2001)	7 & 11	73	U.S.	choice sets	with age number of violations decrease			
Brosig-Koch et al. (2015)	6-15	120	GER	race game	with age performance increases	young females perform worse		
Czermak et al. (2016)	10-17	191	AUT	normal form games	with age elimination of dominated strategies	females better at detecting dominated strategies		
Barash et al. (2018)	6-18	334	U.S.	urn draw	younger children use heuristics, older children (wrong) inferences			
Brocas & Carillo (2018a)	4-5	72	U.S.	tasks to elicit anticipatory reasoning and logical reasoning	with age increase in strategic thinking			
Brocas & Carillo (2018b)	4-7	122	U.S.	matching, fighting, tower, shape game	with age increase in strategic play	females play equilibrium at higher rates		

3.3 TIME PREFERENCES

Time preferences are typically measured by letting subjects choose between a sooner, but smaller payment, and a larger, but later payment. Hence, they are a measure of how the present and the future (or the nearer and the more distant future) are traded off. Most studies find that children and adolescents become more patient as they grow older – i.e., they choose more often the larger, but later reward instead of a smaller, but sooner reward. In other words, older subjects are more likely to delay gratification to a later point in time. This pattern starts already at pre-school or kindergarten age, as Sutter et al. (2015) have found for 3-6 year-olds who had to choose between one small present today and two small presents tomorrow (see also Lemmon and Moore, 2007). Bettinger and Slonim (2007) also report that older children in their sample of 5-16 year-olds are more likely to wait for larger rewards in the future. They estimate that one additional year of age makes subjects about 2% more likely to be patient and choose the larger reward in the future. The method to elicit time preferences does not seem to matter, as Angerer et al. (2015b) show. They compare a simple choice list – where subjects choose between either a specific amount at an earlier point in time or a larger amount at a later date – with the elicitation method based on Andreoni and Sprenger (2012), called a convex budget set where subjects can allocate a specific amount between earlier and later points in time (and allocations to later points in time are more valuable). Both methods produce very similar results in a set of 7-11 year-old children, with older children being more patient with both methods. Using a similar age cohort of 7-10 year-olds, Deckers et al. (2015) also find older subjects to choose more often the larger, but later rewards. Only for teenagers, Sutter et al. (2013) fail to find a positive influence of age on the likelihood to delay gratification, but rather age is insignificant there for 10-18 year-olds.

While age is predominantly positively related to patience, the evidence with respect to gender is very mixed and all over the place. The earliest study of Bettinger and Slonim (2007) finds boys to be less patient than girls, and Castillo et al. (2011) report the same pattern. However, Golsteyn et al. (2014) and Deckers et al. (2015) show the opposite, namely that girls are less patient than boys. Other studies, like Lührmann et al. (2018) and Sutter et al. (2013) in their studies with teenagers or Sutter et al. (2015) with kindergarten children, indicate no significant gender differences, or only under very specific conditions (like in Sutter et al., 2013 where they show weak evidence of females being more patient only in a high stakes condition with no up-front delay).

Aside from the influence of age and gender, a few design parameters have expected effects. Children and adolescents react to larger stake sizes and to shorter waiting times (for the larger, but later reward) by making more patient choices (i.e., waiting more often). This means that prices and the duration of waiting influence behavior in a predictable way. Bettinger and Slonim (2007) find evidence for hyperbolic discounting of children and adolescents, meaning that if there is a positive upfront delay for the smaller, but sooner, reward it is more likely that subjects wait for the larger, but later reward (keeping the waiting time constant, of course).

Family background also matters. Typically, more patient parents have more patient children which speaks in favor of an intergenerational transmission of this preference (Kosse and Pfeiffer, 2012). Children from low SES families have been observed to make substantially more impatient choices compared to children from medium or high SES backgrounds (Deckers et al., 2017). Regarding ethnicity, Castillo et al. (2011) not only find black children to be more impatient, but this result to be especially pronounced for black boys. The discount rate of black boys is on average 14 percentage points larger than that of black girls or white boys.

Experimentally elicited time preferences have been found to be correlated to important field behavior, such as health or educational outcomes. A one standard deviation increase in the discount rate increases disciplinary referrals in schools by 14% in Castillo et al.'s (2011) sample of 9$^{\text{th}}$ graders. In a follow up, Castillo et al. (2018b) even find that time preferences are a good predictor of dropping out of high-school or finishing it. Given that disciplinary referrals in school or becoming a dropout are good indicators for later outcomes in educational attainment or labor market success (Segal, 2013), time preferences of adolescents are related to later labor market outcomes. Subjects who are more patient in time preference experiments are also more likely to save money from their available weekly allowance (Benjamin et al., 2013; Sutter et al., 2013; Lührmann et al., 2018). Concerning health related behavior, higher levels of impatience in children and adolescents are a significant indicator of spending more money on health endangering activities such as smoking or alcohol consumption (Sutter et al., 2013).

Given the importance of time preferences for field behavior, recent studies have started to investigate whether and how policy interventions might affect children's and adolescents' time preferences such that they might foster patience. Most notably, Alan and Ertac (2018a) have implemented an educational intervention promoting forward-looking behavior and patience in 9-10 year-old children in Turkish schools. The students have been exposed to a curricular intervention for several weeks, during which they have encountered various scenario techniques to imagine the trade-off between present and future. Treated students demand on average about 25% smaller rewards for a one week delay of gratification, compared to a control group. The effect is especially pronounced for previously present-biased students who reduce their demand by about 50%. Delay sensitivity also increases in 15 year-olds after an intervention on enhancing financial literacy in German schools (Lührmann et al., 2018).

SUMMARY TIME PREFERENCES: Patience increases typically with age, as older children and adolescents are typically more likely than younger ones to choose a larger, but later reward instead of smaller, but sooner reward. So far, the literature has not produced a clear-cut result on possible gender differences as results are all over the place, sometimes finding girls to be more patient, sometimes boys, and sometimes reporting no difference at all. Socio-economic status of parents is related to children's and adolescents' time preferences as a low SES-background is related to more impatient choices. Importantly, it has been shown that experimentally elicited time preferences are correlated with important field behavior, such as health or educational outcomes. Finally, patience seems to be a malleable skill as interventions can have a positive impact on more forward-looking behavior.

Table 32: Time preferences

Authors	Subject Pool			Experimental Task	Results			
	age	sample size	country		age	gender	SES	other
Bettinger & Slonim (2007)	5-16	191	U.S.	binary choice set	with age patience increases	females more patient		
Castillo et al. (2011)	13-14	878	U.S.	binary choice set		females more patient	black children less patient	impatience correlates with disciplinary referrals
Kosse & Pfeiffer (2012)	preschool age	213	GER	choice task				mother's and child's impatience are correlated
Golsteyn et al. (2014)	13		SWE	choice task		females less patient		patience positively correlated with higher grades, attaining university diploma, higher earnings
Sutter et al. (2013)	10-18	661	AUT	binary choice set		females more patient if high stakes and no up-front delay		impatient children more likely to spend money on smoking, alcohol, have more conduct referrals, less likely to save money
Angerer et al. (2015b)	7-11	561	IT	binary choice list & time investment-exercise	with age patience increases	females less patient in CL		both measures yield similar results
Deckers et al. (2015)	7-10	732	GER	piggy bank	with age patience increases	females less patient	low SES less patient	

Time preferences (cont.)

Authors	Subject Pool			Experimental Task	Results			
	age	sample size	country		age	gender	SES	other
Sutter et al. (2015)	3-6	336	AUT	allocate tokens to envelopes	with age patience increases			changing the default option increases patience
Deckers et al. (2017)	7-9	435	GER	piggy bank			low SES less patient	
Alan & Ertac (2018a)	9-10	1,921	TUR	binary choice set & convex time budget				students in intervention demanded 22-32% fewer presents for one week wait; treated students less likely to receive low "behavioral grade"
Castillo et al. (2018b)	13-15	878	U.S.	binary choice set		females more patient		more impatient children are less likley to graduate from high school
Lührmann et al. (2018)	13-15	914	GER	convex time budget				patience positively correlates with higher math grades and cognition scores

3.4 RISK PREFERENCES

Risk preferences are most of the time measured by letting subjects decide between a safe amount of money (or a non-monetary reward) and a lottery that pays either a higher or lower amount than the safe alternative. Sometimes, risk preferences are also measured by giving subjects a fixed endowment and letting them decide which part of it to invest into a lottery that has typically a positive expected value (Charness and Gneezy, 2010). Experimental studies with children and adolescents have elicited risk preferences for a wide age spectrum, ranging from kindergarten to the late teenage years. Harbaugh et al.'s (2002) seminal study let children and adolescents choose between a risky gamble and a safe outcome. They report that the probability of choosing the risky gamble decreases in adolescents compared to younger children (especially in the loss domain). The propensity to choose the gamble over the safe payoff increases with the probability of a win and decreases with a higher probability of a loss in their sample of 5-20 year-olds. Harbaugh et al. (2002) conclude that children's choices are consistent with the use of subjective probability weights which decrease as children get older, gradually reaching objective probability weighting in early adulthood. Deckers et al. (2015) also find that the willingness to seek risk is getting smaller with increasing age, covering an age range from 7-10 years. Yet, for adolescents, there is less evidence for an age effect. In fact, Sutter et al. (2013) find no age effects on risk taking in their set of 10-18 year-olds, nor do Eckel et al. (2012) for 15-17 year-olds and Munro and Tanaka (2014) for 12-18 year-olds. This suggests that changes in risk preferences might occur before the teenage years, with children becoming less risk seeking until they reach teenager age.

Regarding gender differences in risk preferences, there is strong evidence of girls being significantly more risk averse than boys (Levin and Hart, 2003; Borghans et al., 2009; Moreira et al., 2010; Booth and Nolen, 2012b; Cárdenas et al., 2012; Eckel et al., 2012; Sutter et al., 2013; Deckers et al., 2015; Glätzle-Rützler et al., 2015; Khachatryan et al., 2015; Alan et al., 2017; Castillo, 2017). This pattern reflects the common knowledge of adult women being, in general, more risk averse than adult men (Croson and Gneezy, 2009). The evidence for children and adolescents ranges from very early childhood, starting at 4 years of age, all the way through adolescence, and it also stems from many different regions of the world and cultures. One example for the latter type of work is presented by Cárdenas et al. (2012) who compare two samples of 9-12 year-old subjects in Columbia and Sweden. In their risk task, children can choose between a lottery that yields 0 or 10 points with equal probability, or choose a safe amount that varies between 2 and 7.5 points. In both countries, boys have a certainty equivalent of the lottery of about 4.5 points. Also, in both countries, girls are significantly more risk averse, but the gender differences are much more pronounced in Columbia (with girls' certainty equivalent around 3.2 points) than in Sweden (certainty equivalent around 3.8 points for girls). This evidence suggests that there might be an interaction of gender and culture in the willingness to take risks. Booth and Nolen (2012b) point towards another potential interaction effect by studying how the gender composition in school might affect gender differences in risk taking. They examine risk taking in single-sex and in co-education schools. Girls in co-education schools are 36% less likely to choose a risky lottery while there is no difference in the likelihood to take risks between boys and girls from single-sex schools. The authors argue that the environment significantly affects the propensity to take risks as female-only groups in the experiment induced more moderate risk taking in girls (irrespective of their school composition). However, Booth and Nolen (2012b) also discuss the possibility of self-selection effects into single-sex or co-education schools.

Several papers examine the transmission of risk preferences from parents to their children and observe similar risk taking behavior within parent-child-pairs. A mother's willingness to invest in a lottery correlates significantly with her child's risk preferences. Especially mothers who are more involved in the child's upbringing have a closely related risk tolerance to that of their daughters, as Alan et al. (2017) find in their sample of 7-9 year-olds. This connection of similar risk taking propensities in parents and children is already prevalent in early childhood, as parents' and children's number of risky choices are positively correlated, even though overall children from age 5-8 are more risk seeking than their parents (Levin and Hart, 2003).

Family background also enters via SES as a determinant of children's risk taking. Low SES children are generally more risk taking. Deckers et al. (2017) find that effect size is 23% of a standard deviation, compared to medium and high SES children. This difference by socio-economic status decreases with age and the effect diminishes around age 10 (Deckers et al., 2015). Low socio-economic status has an especially large effect on girls in Alan et al.'s (2017) sample, as girls in the lowest SES-quartile invested on average 14 percentage points more in a risky lottery. Castillo (2017) notes another important influence of family on risk taking. He shows that domestic violence in families affects children to be significantly more risk averse. A similar directional effect is observed by Eckel et al. (2012) who show that having low income peers (outside of the family) makes children more risk averse.

One other important factor that is often discussed in relation to risk preferences is cognitive abilities. Yet, here the literature has produced fairly divergent results. For instance, Benjamin et al. (2013) report high-school students with higher math grades to make more risk neutral choices. Eckel et al. (2012) and Sutter et al. (2013), however, do not find a correlation between math grades and risk taking. Alan et al. (2017) administer several tests of cognitive ability, but only one of them (inhibitory control) is associated with lower risk taking and then only in boys. Castillo (2017) fails to find any significant relationship of cognitive development in 5 and 8 year-olds and risky choices. Overall, the inconclusive pattern might be driven by design details and small differences, yet so far it seems unclear how cognitive abilities are related to risk taking of children and adolescents.

Experimentally elicited risk preferences have also been shown to relate to relevant field behavior. Castillo et al. (2018a) find that more risk averse adolescents are less likely to get disciplinary referrals in school and also less likely to drop out of high school.

SUMMARY RISK PREFERENCES: The overwhelming majority of studies find that girls are more risk averse (i.e., less risk tolerant) than boys, and this pattern persists across childhood and adolescence (and continues in adulthood). There appears to be an age trend as well, especially in childhood, as older children are less risk seeking or risk taking than younger children. This seems to be driven by subjective probability weights that change across age. Family background is important, as the risk preferences of parents are typically correlated to those of their offspring, but also as low socio-economic status of parents is associated with more (and sometimes excessive) risk taking.

Table 33: Risk preferences

Authors	Subject Pool			Experimental Task	Results			
	age	sample size	country		age	gender	SES	other
Harbaugh et al. (2002)	5-20	187	U.S.	safe option vs. gamble	with age less risk seeking			
Levin & Hart (2003)	5-7.8	102	U.S.	safe option vs. gamble		females more risk averse (only for losses)		parent's total number of risky choices positively related to child's
Borghans et al. (2009)	15-16	347	NL	Ellsberg two-color choice task		females more risk averse		
Moreira et al. (2010)	4-6	100	BRA	safe option vs. gamble		females more risk averse		
Booth & Nolen (2012b)	15	260	U.K.	safe option vs. gamble		females more risk averse		females from same-sex schools just as likely to enter lottery as males; being in an all-girl-group makes females less risk averse
Cárdenas et al. (2012)	9-12	1,200	SWE & COL	choice of six gambles and safe option		females more risk averse		competing and risk taking are correlated for Swedish children and Columbian males
Eckel et al. (2012)	14-15 & 16-17	490	U.S	choice out of six gambles		females more risk averse	low income peers reduce risk tolerance	
Tymula et al. (2012)	12-17	33	U.S.	safe option vs. gamble	adolescents more risk averse than adults			
Sutter et al. (2013)	10-18	661	AUT	Ellsberg two-color choice task		females more risk averse		

Risk preferences (cont.)

Authors	Subject Pool			Experimental Task	Results			
	age	sample size	country		age	gender	SES	other
Munro & Tanaka (2014)	12-18	412	UGA	Holt-Laury pair wise choice framework				71% of children have lower risk aversion than their parents
Alan et al. (2015)	6.8-8.9	746	TUR	Gneezy & Potters risk elicitation task	older males invest more	females more risk averse	low SES females invest more	child's risk taking correlates positively with mother's risk taking
Deckers et al. (2015)	7-10	732	GER	safe option vs. gamble	with age less risk seeking	females more risk averse	low SES more risk seeking	
Glätzle-Rützler et al. (2015)	11-18	755	AUT	investment in lottery	increase in investments between 10^{th} and 12^{th} graders	females invest less		no evidence of myopic loss aversion
Khachtaryan et al. (2015)	7-16	824	ARM	safe option vs. gamble	with age males become more risk seeking	females more risk averse		
Castillo (2017)	5 & 8	2,000	PER	choice of six gambles		females more risk averse		domestic violence correlates with risk averse decisions
Deckers et al. (2017)	7-9	435	GER	safe option vs. gamble			low SES more risk seeking	
Castillo et al. (2018a)	13-14	1,275	U.S.	choice between two lotteries				children who are more risk averse are less likely to receive disciplinary referrals, and more likely to complete high school

Broadly speaking, social preferences capture the different ways in which subjects consider own payoffs (or rewards) and others' payoffs. They capture standard preferences – that are often defined as subjects caring only for themselves and ignoring outcomes for others –, but also various forms of non-standard preferences which allow for positive (and in case of spite also negative) weights for others' outcomes in a subject's utility function. Social preferences play a role when subjects have to make allocation decisions in which they split up a pie among themselves and others, with others being powerless, but also when subjects interact in a strategic game with others, like in simple bargaining games or games of cooperation. We start the survey about children's and adolescents' social preferences by looking at allocation tasks that are a form of individual decision making, void of any strategic interaction. In the following sections 3.6 and 3.7 we will consider interactive games.

The most often used task to study social preferences of children and adolescents is the dictator game where a dictator is endowed with a fixed endowment and subsequently can distribute it between him- or herself and a powerless recipient. While it is commonly called a game, the dictator game is, in fact, an individual decision making task. A variant of the dictator game lets subjects choose between different allocations where the sum of money distributed in each allocation does not need to be constant across allocations.

The dictator game (and its variants) has been the most often used vehicle to study social preferences of children and adolescents. Concerning the influence of age, the evidence is pretty straightforward: the older subjects get, the more likely they are to transfer increasing parts of their endowment to the recipient. This is not to say that subjects become hyper-fair by offering more than 50% of their endowment, but older ones are less likely to be selfish by keeping the whole endowment for themselves and also give more often up to 50%. For instance, in Gummerum et al.'s (2010) study the modal offer of 3-4 year-olds is to keep everything for themselves, while 5-6 year-olds choose more often an equal split. Similar trends with very young children are found in List and Samek (2013), Ben-Ner et al. (2017) or Brocas et al. (2017), and for elementary school kids, aged 6 to around 10 or 11, it is also typically observed that older children are more generous towards the recipient (Harbaugh et al., 2003a; Bettinger and Slonim, 2006; Martinsson et al., 2011; Blake et al., 2015; Deckers et al., 2015; Chen et al., 2016; Maggian and Villeval, 2016; Brocas et al., 2017; Sutter et al., 2018).[5] This trend continues also into adolescence, where mean allocations to recipients increase with age and can reach levels of around 35% of the available endowment (Harbaugh et al., 2003a; Eckel et al., 2011; John and Thomsen, 2015), which is a high average compared to about 28% found in a meta-study of dictator games with adults (Engel, 2011).

A specific design invented by Fehr et al. (2008) allows to define different types of social preferences. The design consists of three “games” (again individual decision making tasks) in each of which a subject can choose among two options. One option is always an egalitarian outcome, while the alternatives differ in order to be able to classify the social preference type of a specific subject from the three choices made. While this classification has some limitations (see Bauer et al., 2014 for an explanation and extension), the three types for classification are the following: an egalitarian type who prefers the egalitarian options; a spiteful type who always minimizes the recipient’s payoff; and an altruistic type that maximize the recipient’s payoffs. The latter type is indistinguishable from an efficiency-maximizing type, however. Fehr et al. (2008) study 3-8 year-old children and find that egalitarian types, i.e., those with a strong aversion against inequality, become considerably more frequent from age 3 to age 8. Children at the age of 3-4 behave selfishly to a very large degree, whereas the majority of children aged 7-8 prefer egalitarian allocations that avoid both advantageous and disadvantageous inequality. More precisely, about 60% of 7-8 year-old children can be classified as having egalitarian preferences, while the

[5] Maggian and Villeval (2016) combine their dictator game with an option to lie about which allocation was randomly determined by a computer. They found that all children in their set of 7-17 year-olds have a strong aversion against lying (about 85% do not lie when it would potentially benefit them).

corresponding share for 3-4 year-olds is only 20% (see also Bauer et al., 2014 for a similar pattern of age effects). In a follow-up, Fehr et al. (2013) show that egalitarian types peak at around age 8. Looking at 9-17 year-olds, they find that efficiency seeking becomes much more prevalent – at the expense of egalitarianism – with increasing age. Prior to the latter finding, Almås et al. (2010), have already shown that efficiency seeking becomes the most important social preference motive in adolescence. Yet, they have added an interesting twist by letting their dictators divide a pie of money that has been generated through a real effort task, executed by both the dictator and the recipient. This allows examining whether social preferences – i.e., sharing behavior – depend upon the effort invested by dictators and recipients to generate the pie in the first hand. When efforts – and thus the contribution to the pie – differ, 10-11 year-olds typically do not condition their allocation choices on the differences in effort levels. However, adolescents around age 15 do so, and they are predominantly meritocrats who accept unequal earnings if they are due to unequal effort provision. Most adolescents hence deem it fair that those who have exerted less effort deserve to earn less (Almås et al., 2010, 2017).

When it comes to the examination of gender differences in allocation choices, the predominant finding is that girls are more generous in classical dictator games and more likely to be of an egalitarian type in the Fehr et al. (2008) design (Harbaugh et al., 2003a; Bettinger and Slonim, 2006; Houser and Schunk, 2009; Gummerum et al., 2010; Martinsson et al., 2011; Fehr et al., 2013; List and Samek, 2013; Angerer et al., 2015a; Deckers et al., 2015; Chen et al., 2016; Maggian and Villeval, 2016; Angerer et al., 2017). Boys are more likely to be efficiency seeking types that try to maximize the sum of payoffs when choosing between different allocations. If multiple options and recipients are available (as in designs based on Engelmann and Strobel, 2004, 2004), one can also see that girls can be classified more often as a maximin type that tries to maximize the minimum payoff in the set of people who are affected by a specific allocation (Sutter et al., 2018). Whereas such gender differences refer to the decision maker's gender, Houser and Schunk (2009) report also an influence of the gender of the recipient, as both genders share higher amounts when they are aware that the recipient is male.

Social preferences within families seem to be related between children and parents, although the relation is not always significant. Ben-Ner et al. (2017), for instance, find that the generosity of young children, aged 3-5, in a dictator game is related to parents' donation to a charity, but that the relation is significant only for firstborn children (who might be most strongly influenced by parents because parents could devote most time to their oldest offspring). Kosse et al. (2018) show that maternal pro-sociality and interaction patterns are able to predict pro-social behavior of 7-8 year-olds. The socio-economic status of parents also plays a role. Children from low SES-backgrounds are less altruistic and more spiteful (Bauer et al., 2014; Angerer et al., 2015a; Deckers et al., 2015, 2017; Kosse et al., 2018). For example, only 33% of low SES children prefer the egalitarian option of (1,1) over option (2,0) in the sharing game of Fehr et al. (2008), while almost half of high SES children select that option (Bauer et al., 2014).

Given the importance of SES for social preferences and given the relevance of social preferences as a non-cognitive skill that facilitates cooperation, Kosse et al. (2018) have run an intervention study to check whether social preferences are malleable. From control groups they see that there is a considerable gap in the social preferences of children from low SES-backgrounds and those from high SES-backgrounds. By implementing a mentoring program with a mentor who acts as a benevolent friend and spends time with a child from low SES-families, these treated children score about 25% of a standard deviation higher on the pro-sociality scale than children in the control group with low SES, and the intervention even closes the gap between treated low SES-children and those (untreated) from high SES-backgrounds. Hence, their study provides causal evidence that social preferences can be changed. One of the channels through which this works is that the intervention affects beliefs about pro-social behavior of others. Another intervention has been run by Cappelen et al. (2016) who have studied the effects of early education on social preferences by admitting either children to different preschool programs or by building up a parenting academy. These interventions on 3-4 year-olds lead to significantly higher levels of prosociality at age 7-8, showing that education programs can make children more prosocial. The literature has also examined other factors that determine children's and adolescents' social preferences. One of them refers to in-group favoritism, or in other words to the distinction between in-groups and out-groups. If a recipient belongs to the same social group it is typically associated with more generous behavior of the decision maker towards the recipient (Fehr et al., 2008, 2013). The same social group is usually defined as someone from the decision maker's class in school, from the same school or even only someone speaking the same language – whereas the outgroup is formed by someone from a different class, a different school or someone speaking a different language (Fehr et al., 2008, 2013). Even when children act as spectators and have no stakes in the decision, in-group favoritism is prevalent (Angerer et al., 2017). Especially boys' allocation choices are shaped by strong parochialism. While both genders discriminate against children of a different language group, about one quarter of 6-11 year-old boys decide strongly in favor of their own language group compared to only about one sixth of girls (Angerer et al., 2017).

Further factors that may affect sharing behavior of children and adolescents range from the information about a recipient's neediness or the prevailing social norms in the child's environment to a child's level of self-control. When the recipient is framed as "poor" (e.g., a child with no toys) or the donated money is given to a charitable cause, children become more generous. For example, Bettinger and Slonim (2006) observe higher donations to charities than when children share with their peers (38% versus 26%). Announcing decisions publicly to the classroom – and thus appealing to social image concerns – further increases the level of sharing in school children, especially for popular children (Chen et al., 2016). At the same time, however, a public announcement of decisions can reduce the amount sent in a dictator game if the situation is framed as a competition where the winner is the one who keeps the largest amount for him- or herself (Houser and Schunk, 2009). Furthermore, social norms of sharing play an important role for children and adolescents as making sharing norms salient induces higher rates of giving (Eckel et al., 2011; Blake et al., 2015). The level of self-control and IQ can positively influence the amount sent, while risk and time preferences affect donations positively, but in a non-linear way (Eckel et al., 2011; Angerer et al., 2015a; Blake et al., 2015; John and Thomsen, 2015; Chen et al., 2016).

SUMMARY SOCIAL PREFERENCES IN INDIVIDUAL DECISION MAKING TASKS: Very young children in kindergarten have mostly selfish tendencies in the dictator game and variants thereof, especially if it is costly to make a recipient better off. When entering school, children become more generous towards recipients, in particular needy ones, and they become in particular inequality averse in later childhood. The predominance of egalitarianism is not sustained for adolescents, however, as they are primarily motivated by efficiency and social welfare concerns. On top of that, adolescents become meritocratic, meaning that they make their allocation decisions dependent on subjects' levels of exerted effort. Girls are typically more generous and more inequality averse than boys who care more about efficiency. Children from low SES-backgrounds are often less pro-social and less generous, and there is a positive relation between parents' and children's social preferences. Other factors like in-group favoritism or self-control also play a role for social preferences.

Table 34: Social preferences – Individual decision making

Authors	Subject Pool			Experimental Task	Results			
	age	sample size	country		age	gender	SES	other
Harbaugh et al. (2003b)	7-18	310	U.S.	UG, DG	older children give more	females give more		
Bettinger & Slonim (2006)	6-14	572	U.S.	DG	older children give more	females give more		children give more to charity than peers
Fehr et al. (2008)	3-8	229	CH	DG, in/out-group	with age increase in egalitarian choices and parochialism	females less parochial		
Houser & Schunk (2009)	8-10	151	GER	DG		females give more		competition decreases fairness in males but not females
Almås et al. (2010)	11-19	486	NOR	DG, spectator	with age increase in meritocratic and efficiency-based choices	females less efficiency-based		
Gummerum et al. (2010)	3-5	77	U.K.	DG		males made zero offers more often		
Eckel et al. (2011)	14-17	490	U.S.	DG				high norm conformance give and expect more
Martinsson et al. (2011)	10-15	650	AUT & SWE	DG	older children less difference averse and more welfare concerned	females more difference averse		Swedish children less difference averse and more social-welfare oriented compared to Austrian children

abbreviations: dictator game (DG); prisoner's dilemma (PD)

Social preferences – Individual decision making (cont.)

Authors	Subject Pool			Experimental Task	Results			
	age	sample size	country		age	gender	SES	other
Fehr et al. (2013)	8-17	717	AUT	DG	older children more pro-social, less envious, weakly altruistic type, more parochial	females more egalitarian		
List & Samek (2013)	3-5	122	U.S.	DG		females more altruistic		
Bauer et al. (2014)	4-12	275	CZE	DG	older children more pro-social, more altruistic		low SES less altruistic and more spiteful	
Angerer et al. (2015a)	7-11	1,070	IT	DG	older children more altruistic	females give more		
Ben-Ner et al. (2015)	3-6	147	U.S.	DG	older children give more			parent and child giving not correlated
Blake et al. (2015)	6-13	433	IT	DG	older children give more			social norm on giving increases giving
Deckers et al. (2015)	7-10	732	GER	DG	older children more altruistic	females more altruistic	low SES do not become more altruistic with age	
John & Thomsen (2015)	10-16	895	GER	DG, PGG				academic track children give more in DG
Cappelen et al. (2016)	7-8	303	U.S.	DG				majority of children found inequality fair if there was initial inequality

abbreviations: dictator game (DG); prisoner's dilemma (PD)

Social preferences – Individual decision making (cont.)

Authors	Subject Pool			Experimental Task	Results			
	age	sample size	country		age	gender	SES	other
Chen et al. (2016)	6-12	231	IT	DG	older children give more			popularity promotes pro-social behavior when decisions are public
Maggian & Villeval (2016)	7-14	637	IT	DG, lying option	older children less selfish, more efficiency concerned	young females more egalitarian		
Almås et al. (2017)	14-15	524	NOR	DG, spectator			low SES more egalitarian	
Angerer et al. (2017)	6-11	824	IT	spectator		females discriminate less		
Brocas et al. (2017)	6-18	334	U.S.	DG, PD	increasing altruism until grade 4 then drops			
Deckers et al. (2017)	7-9	435	GER	DG			low SES less altruistic	
Kosse et al. (2018)	7-8	607	GER	DG			low SES less pro-social	mentoring program increases pro-sociality, altruism, and trust
Sutter et al. (2018)	8-17	883	AUT	DG, spectator	younger children more inequality averse, older children increasing efficiency concerns	females primary motive maximin, males primary motive efficiency		

abbreviations: dictator game (DG); prisoner's dilemma (PD)

3.6 SOCIAL PREFERENCES II: BARGAINING GAMES

Social preferences – like a concern for fairness or efficiency – are also important in interactive games where two parties bargain with each other (in a stylized way). The two most often used bargaining games with children and adolescents are the ultimatum game and the trust game. We deal with each of them consecutively.

3.6.1 ULTIMATUM GAME

In an ultimatum game, a proposer is equipped with a fixed endowment and can offer some of it to a responder. The responder can either accept the offer – in which case the proposed allocation is implemented – or reject it – in which case both the proposer and the responder receive nothing. Obviously, fairness concerns of proposers play an important role in this game, but also strategic considerations due to the responder's power to reject what he or she considers an unfair offer.[6] Considering the development of offers in the ultimatum game contingent on age, the literature does not provide a clear-cut pattern. Harbaugh et al. (2003a) find that offers increase significantly, but modestly, with age in their sample of 7-18 year-olds. They use a one-shot ultimatum game with incentives. When any of these two features is changed, results look differently. Murnighan and Saxon (1998) use a purely hypothetical scenario, and in such a setting younger children made larger offers in the ultimatum game than older children. Harbaugh et al. (2007) repeated an incentivized ultimatum game, and there they find no age effect on offers in their group of 8-18 year-olds. The pattern of behavior is qualitatively similar to adult behavior (Güth and Kocher, 2014). When endowed with ten tokens almost half of children and adolescents propose the egalitarian outcome of five tokens each, while 20% of proposals are lower than three tokens (Harbaugh et al., 2007). Across repetitions, Harbaugh et al. (2007) observe an interesting learning effect that is stronger for younger than

[6]The strategic considerations refer back to the ability of children and adolescents to understand strategic games and act sophisticatedly in such games. Section 3.2 has dealt with this aspect.

for older children. Receiving a rejection prompts especially young children in the role of proposer to increase their offer in the next round, indicating reinforcement learning and strategic behavior. Sutter (2007) also reports no age effects in his study of mini-ultimatum games with 7-15 year-olds. In these mini-ultimatum games, proposers always face only two allocations from which they can choose, and responders can then accept or reject the selected allocation. One of the available allocations is very unfair, as it yields 8 units of money for the proposer, but only 2 units for the responder – noted as allocation (8,2). Varying the alternative allocation – that can be (10,0), (8,2), (5,5) or (2,8) – it is possible to study the importance of intentions. For instance, offering (8,2) when the alternative would be (10,0) is a kind act, while it is not when the alternative is (5,5). It turns out that both proposals and rejection rates are practically the same for children (up to age 10) and adolescents (up to age 15), but that both children and teenagers base their rejection decisions relatively more than adults on actual outcomes (i.e., payoffs) rather than the proposer's intentions. This means that intentionality in a bargaining process is more important for adults than in pre-adulthood.

Rejection rates in the standard ultimatum game increase monotonically when offers get smaller, but they are also not contingent on age, as Harbaugh et al. (2007) show. An insignificant age effect on rejections is also reported in Castelli et al. (2010), although they note a trend that younger children seem to accept unfair offers slightly more often. The focus of their paper is, however, not on age (in their set of 5-10 year-olds), but on the effect of theory-of-mind on ultimatum game behavior. In this respect, they find that children who have developed theory-of-mind are more likely to accept unfair offers.

None of the above mentioned papers report any statistically significant differences in the size of offers or the acceptance rates between boys and girls (Harbaugh et al., 2003a, 2007; Sutter, 2007; Castelli et al., 2010), so gender does not seem to play a major role in ultimatum game behavior.

3.6.2 TRUST GAME

In a trust game, a trustor has some fixed endowment and can transfer a fraction (from 0% to 100%) to a trustee. The transferred amount is typically tripled, and then the trustee can send back any amount that he or she finds suitable (without any tripling of the return, though). The trustor's decision is usually interpreted as a measure of trust, whereas the trustee's decision is used to measure trustworthiness. Looking at how much children in the role of the trustor transfer to the trustee, it is evident that younger children transfer less than older children. Elementary school children (at roughly age 8-9) pass the smallest amount (Harbaugh et al., 2003b; Sutter and Kocher, 2007). Sutter and Kocher (2007) find a monotonic increase in transfers with increasing age all the way to adulthood, while Harbaugh et al. (2003b) only observe an increase until the 9th grade (around age 15), but a decrease in transfers for 12th graders. Both papers, nonetheless, confirm children and adolescents to be less trusting than adults.

Relative returns normalize the trustee's return to the trustor by the amount of the tripled transfer. These relative returns are reported to increase with age in Sutter and Kocher (2007). 8-12 year-olds have the lowest return rate ranging from 10% to 15%, while for adolescents it increases to roughly 30%. Harbaugh et al. (2003b), however, do not observe an increase in relative returns with increasing age. One potential design difference is that they use a strategy method for trustees by which the latter have to indicate their return for each possible level of the transfer, while Sutter and Kocher (2007) only ask for the return for the actual level of the trustor's transfer (so-called direct method). Both studies, however, note that returns depend positively on transfers, which indicates that reciprocity is a prevalent behavioral pattern already in childhood and adolescence. Both studies also agree in the finding that, given the actual return rates, the payoff maximizing strategy for children and adolescents is to send (close to) zero, while adults maximize expected payoffs by showing full trust and transferring their full endowment (Harbaugh et al., 2003b; Sutter and Kocher, 2007).

Concerning gender effects, both studies find hardly any differences, except for very narrow age brackets. Harbaugh et al. (2003b) observe higher trust levels of 8-9 year-old boys, and Sutter and Kocher (2007) higher trustworthiness of girls aged 8-9. Other than that, gender seems to be uncorrelated with trust game behavior.

Felfe et al. (2018) present an interesting natural field experiment about the effects of birthright citizenship on trust in German 15-16 year-olds. They study the behavior of native and immigrant adolescents, exploiting a law change in Germany in the year 2000 which automatically awards newborns within Germany with German citizenship. The hypothesis is that German citizenship for second-generation immigrants should lead to less discrimination between natives and immigrants. In fact, this is what they find, although the effect is significant only for boys. Those born immediately after the law change almost close the gap in the transfers to natives or immigrants, while for boys born immediately before the law change, there is a strong gap of about 20%, thus yielding much lower efficiency levels in interaction. This natural field experiment shows that behavior of adolescents is influenced by legal conditions of citizenship.

SUMMARY SOCIAL PREFERENCES IN BARGAINING GAMES: Fairness and efficiency concerns are important in bargaining games. When running ultimatum games with children and adolescents, they accept equal splits of the pie most often, and rejection rates increase in the spread between the proposer's and responder's share. Age effects are at best weak, since fairness concerns seem deep rooted and early developed. In trust games, transfers of trustors increase with age, which might coincide with the increasing importance of efficiency when children turn into adolescence, as shown in the previous section. There is some evidence that trustworthiness of trustees also increases with age, meaning that the extent of reciprocity might increase with age. In both games, there are hardly any gender effects on behavior.

Table 35: Social preferences - Bargaining games

Authors	Subject Pool			Experimental Task	Results			
	age	sample size	country		age	gender	SES	other
Murnighan & Saxon (1998)	5-15	240	U.S.	UG	younger children offer more and accept less	females give more		
Harbaugh et al. (2003b)	7-18	310	U.S.	UG, DG	older children give more	females give more		
Harbaugh et al. (2007)	8-18	256	U.S.	UG	older children make more consistent proposals			observing larger proposals by others leads to larger own proposals
Sutter (2007)	7-15	200	AUT	UG				even when proposer has no choice, 46% of children reject unfair offers
Castelli et al. (2010)	5-10	177	IT	UG	younger children accept unfair offers more often			theory of mind reduces acceptance of unfair offers
Harbaugh et al. (2002)	8-17	153	U.S.	TG	increase in trust until ninth grade, steep decrease for twelfth graders	third grade: females less trusting		
Sutter & Kocher (2007)	8-16	662	AUT	TG	older children make higher transfers	eight year-olds: females return more		
Felfe et al. (2018)	15-16	4,077	GER	TG				immigrant children (especially girls) discriminate against native children

abbreviations: ultimatum game (UG); trust game (TG)

3.7 SOCIAL PREFERENCES III: GAMES OF COOPERATION

Cooperation is almost always measured by either running a prisoner's dilemma game or the generalized version of it, a public goods game. These games of cooperation are characterized by a tension between individual incentives to defect (i.e., not cooperate) and a collective interest in cooperation as it maximizes social welfare, i.e., the size of the pie that can be generated in these games.

The majority of studies on cooperation of children and adolescents reveal that older children are more likely to cooperate, while younger children defect more often or contribute less in public goods games with a continuous action space (Fan, 2000; Harbaugh and Krause, 2000; Houser et al., 2012; Angerer et al., 2016; Brocas et al., 2017). A variant of a prisoner's dilemma game is used by Brocas et al. (2017) who let children from age 5 onwards play an alternating allocation task that is equivalent to a sequential, symmetric prisoner's dilemma game in which subjects have to trade off the short term gains from defection (i.e., selfishly picking the more rewarding option) and the long term gains from cooperation (going for the equal payoff for both players). They observe an age trend all the way through childhood and adolescence as cooperation increases with age. A similar pattern is observed in Angerer et al. (2016) who let 6-11 year-olds play a prisoner's dilemma game (where each subject has 5 tokens and each token sent to the other player is doubled in value). The 11 year-olds send about 25% more tokens to their partner than then 6 year-olds, and the increase is fairly linear across age. Some papers do not show a significant age trend, though. Lergetporer et al. (2014) and John and Thomsen (2015) observe on average the same cooperation levels across each age group for 7-11 (10-16) year-olds. Yet, both papers report a slight – but insignificant – tendency for cooperation rates to increase with age. Cipriani et al. (2013) find no significant age effect as well. However, their sample size (with 38 observations) is by far the smallest in the set of papers considered in this section, for which reason their null result might be taken with care.

Overall, girls and boys do not differ noticeably in their likelihood and extent of cooperation (Fan, 2000; Harbaugh and Krause, 2000; Cárdenas et al., 2014; Lergetporer et al., 2014; John and Thomsen, 2015; Brocas et al., 2017; Hermes et al., 2018). There is, nonetheless, a slight suggestive tendency for girls to cooperate more. For instance, Angerer et al. (2017) observe girls between the ages of 6-11 to be more cooperative than boys. Investigating gender differences in cooperation between Columbia and Sweden, results suggest Columbian girls to cooperate less than Swedish girls and Swedish girls to be more cooperative than Swedish boys. Children also tend to cooperate more with boys than with girls (Cárdenas et al., 2014).

Not much is known about the influence of family background on cooperation levels of children and adolescents. Cipriani et al. (2013) do not find any effect of their controls for socio-demographic background characteristics, but note the small sample size. They also fail to find a relation between parents' behavior and their children's behavior in the public goods game. Yet, family bonds do matter for children's level of cooperation, as Peters et al. (2004) show. They let children and parents play a public goods game. In one condition children are paired with their own parents, whereas in another they are paired with other children's parents. When the game is played only among family members, children contribute substantially more to the public good. Parents, however, do not condition their cooperation on whether or not they are paired with their own children or children of strangers. Hermes et al. (2018) study whether parents (and teachers) are able to predict the level of cooperation of their children through a questionnaire. They find that this is not the case.

Resembling the evidence from allocation tasks discussed previously, cooperation of children and adolescents depends on distinctions between in-groups and out-groups. Angerer et al. (2016) present an example of this effect by exploiting group identity in a bilingual city (where half of the inhabitants speak German and the other Italian and where schools are segregated by language). In their set of 6-11 year-olds, they find that children are least cooperative in a prisoner's dilemma game when matched with a child from the other language group (i.e., the out-group). The level of cooperation is higher when a child is matched with someone from the same language group, but a different school, and highest if the match is with someone from the same class (which implies the same language).

One way to increase cooperation in children may be through educational interventions. Fan (2000) examines whether special lectures that teach the value of cooperation can increase cooperation, but she fails to find an effect. Another intervention can be the introduction of a third, uninvolved party who has the ability to punish uncooperative behavior. This is usually called third party punishment. Lergetporer et al. (2014) find that such a costly punishment option for an uninvolved third party increases cooperation rates of 7-11 year-olds considerably by doubling them, in fact. The increase is due to two main reasons. The first, and straightforward factor is that the fear of getting punished lets subjects increase their likelihood of cooperation. Second, and less obvious, is the fact that players in the prisoner's dilemma become more optimistic about their partner's likelihood of cooperation when a third party with a punishment opportunity is present (who may punish the partner as well). Due to more optimistic expectations about the partner's likelihood of cooperation, players become more cooperative themselves – which proofs that already young children are conditional cooperators.

SUMMARY COOPERATION The level of cooperation is typically increasing with age, in particular in childhood. Younger children free-ride more often than older ones, while adolescents display more prosocial and reciprocal motives in public goods and prisoner's dilemma games. Gender effects are largely absent, and so far there is also little knowledge about the influence of family background. In-group favoritism promotes higher cooperation levels, as does the presence of third parties with an option to punish defectors.

Table 36: Social preferences – Cooperation

Authors	Subject Pool			Experimental Task	Results			
	age	sample size	country		age	gender	SES	other
Fan (2000)	6-11	196	TWN	PD	older children cooperate more			lecture on cooperation had a short term positive effect on cooperation levels
Harbaugh & Krause (2000)	6-12	208	U.S.	PGG	older children cooperate more but learn to free-ride			
Peters et al. (2004)	9-16	68	U.S.	PGG, in/out-group				parents give substantially more than children
Houser et al. (2012)	6-11	406	IT	common pool resource game	older children resist more in public condition			
Cipriani et al. (2013)	5-12	38	U.S.	PGG				parent's contribution does not affect that of child
Cárdenas et al. (2014)	9-12	800	COL & SWE	PD				
Lergetporer et al. (2014)	7-11	1,120	IT	PD, TPP				TPP more than doubled cooperation rates
John & Thomsen (2015)	10-16	895	GER	DG, PGG				academic track children give more in DG
Angerer et al. (2016)	6-11	828	IT	PD	older children cooperate more, more parochial	females cooperate more		cooperation highest for children in same school-class who speak same language, lowest for other language
Brocas et al. (2017)	6-18	334	U.S.	DG, PD	older children cooperate more			
Hermes et al. (2018)	6	129	GER	PGG				children cooperate conditionally

3.8 COMPETITIVENESS

Here we focus on competitiveness in the sense of a willingness to expose oneself to a competitive situation. We are not looking specifically at performance under competition, but rather at a preference to compete at all.[7] Most studies that investigate competitiveness follow the seminal design by Niederle and Vesterlund (2007) where participants perform a specific task in three different stages. In the first stage, they are paid a piece-rate, hence there is no competition. In the second stage, there is a tournament where only the winner gets paid (a higher piece rate than in the first stage). Finally, in the third stage, subjects are free to choose their compensation scheme by either selecting the piece rate or the tournament. It is the choice of the tournament in stage three that measures the willingness to compete.

Looking at the development of competitiveness across age, it seems to be the case that very early on, from age 3-6, children become more likely to compete in several studies (Khachatryan et al., 2015; Sutter and Glätzle-Rützler, 2015; Sutter et al., 2016; Khadjavi and Nicklisch, 2018). For example, Khadjavi and Nicklisch (2018) report in their study with 3-6 year-olds that the likelihood to compete increases by around 10 percentage points with each year. Beyond this early age, the literature does not report a clear-cut pattern in the willingness to compete (Sutter and Glätzle-Rützler, 2015), while with respect to performance in a given task children and adolescents become almost always better the older they get (Andersen et al., 2013; Sutter and Glätzle-Rützler, 2015; Sutter et al., 2016; Khadjavi and Nicklisch, 2018).

Most studies with children and adolescents document a strong gender gap in the willingness to compete (Booth and Nolen, 2012a; Andersen et al., 2013; Buser et al., 2014; Dreber et al., 2014; Sutter and Glätzle-Rützler, 2015; Almås et al., 2016; Sutter et al., 2016). Overall, girls are much less likely to choose a competitive payment scheme than boys. For example,

[7] Gneezy and Rustichini (2004) look at performance of 10-11 year-olds in a running task under competition. They observe that when running alongside another subject, boys improve their performance by a wide margin, while the performance of girls deteriorates. In their case, children had no choice, however, whether they wanted to compete or not.

in the study of Buser et al. (2014), 15 year-old girls have a 23 percentage point lower probability of choosing the tournament after controlling for performance and the associated likelihood of winning. In their subsample of 9-18 year-olds, Sutter and Glätzle-Rützler (2015) report a similar magnitude of the difference, since 40% of boys choose to compete in a math task, but only 19% of girls.[8] Again, a similar gap is found in Almås et al. (2016) where more than 50% of boys, aged 14-15, compete, while only 32% of girls choose to enter competition. Interestingly, the gender gap, if anything, is slightly getting larger in a panel study conducted by Sutter and Glätzle-Rützler (2015) who find that over a span of two years girls become even less likely to choose competition during adolescence while for boys there is hardly any change when they get two years older.

The literature is less clear about whether there is a specific age in which the gender differences in the willingness to compete set in – which would be important information for potential policy interventions. In her set of 3-5 year-olds in the U.S., Samek (2013) does not observe any gender differences in competitiveness. Sutter and Glätzle-Rützler (2015) – covering an age range from 3-18 year-olds in Austria – find a gender gap for 5 year-olds and the gap persists beyond that age. Going to less developed societies in Northeast India, Andersen et al. (2013) do not find a gender gap until the age of 12, but from then on the gap persists as well. This means that there might be an interaction between the cultural and economic background of a society and the gender gap in competitiveness. Andersen et al.'s (2013) study is an example for that. They study matrilineal and patriarchal societies. In societies based on strong patriarchal structures girls exhibit significantly lower competitive preferences compared to girls in matrilineal societies. This effect appears in adolescence and persists from then on. There is also some literature documenting that in less developed countries there may be no gender differences in competitiveness. Khachatryan et al. (2015) find no gender differences in the willingness to compete in 7-16 year-old Armenians.

[8]Most worryingly from an efficiency point of view is the observation that the gender gap is even more pronounced in the top-performing quartile of subjects where Sutter and Glätzle-Rützler (2015) and Sutter et al. (2016) find a gap of almost 40 percentage points between boys and girls.

Zhang (2011) investigates 11-15 year-old high school students in rural China and finds no gap between boys and girls when they belong to the majority group of Han Chinese. In ethnic minority groups, however, she observes the typical gender gap of boys being more willing to compete than girls. So, the evidence seems to suggest that in highly developed countries one can typically observe a gender gap in the willingness to compete, while in developing countries this is not the case or the gap emerges later or only for subgroups. Almås et al. (2016) provide a potential explanation for this cross-country pattern. They show in a sample of 14 to 15 year-old Norwegians that there exists a large and significant gender gap in the willingness to compete among adolescents who have parents with high education levels. However, there are no significant gender differences among adolescents with parents who have low levels of education, indicating that it is, perhaps, the level of parental education that matters. Generally speaking, these findings suggest that the overall education level – which presumably influences economic development – may explain why in less developed countries with lower overall education levels like India, Armenia or rural China gender differences in the willingness to compete seem not to exist among adolescents or at least set in at a later age compared to highly developed countries like Sweden, Norway or Austria.

The gender differences in competitiveness seem to be related to two other factors that differ across gender: beliefs about one's own (relative) performance and risk preferences. Boys are typically much more confident (i.e., often overconfident) that they win the tournament, for which reason it is reasonable for them in expectation to choose a competitive payment scheme. Evidence for such gender differences in expected performance abounds (Buser et al., 2014; Dreber et al., 2014; Sutter and Glätzle-Rützler, 2015; Sutter et al., 2016). There is also a gender difference in risk aversion (see also section 3.4) that contributes to the gender difference in competitiveness. Since the tournament payment scheme is risky (compared to the safe payment of a piece-rate scheme), more risk averse subjects are less likely to choose a competitive payment scheme. Given that girls are typically more risk averse than boys (Cárdenas et al., 2012; Buser et al., 2014; Dreber et al., 2014; Khachatryan et al., 2015; Sutter and Glätzle-Rützler, 2015), this difference also explains part of the gender gap in the willingness to compete. It is important, however, that controlling for both – beliefs about expected performance and risk aversion – there is usually still a significant gender gap left that is not explained by these two factors, and the unexplained gap is usually in the range of around 10 percentage points (Buser et al., 2014; Sutter and Glätzle-Rützler, 2015; Almås et al., 2016).

As already implied above, a low socio-economic status of a child's family has a negative influence on the likelihood to compete (Bartling et al., 2012; Almås et al., 2016). This effect is particularly strong for boys (Almås et al., 2016). Low SES can also be related to health issues, as Bartling et al. (2012) argue. Children who have had more health issues in recent months and come from a low SES background are about 10 percentage point less willing to self-select into competition. Khadjavi and Nicklisch (2018) reveal another facet of parental influence. Their ambitions for their offspring's achievements can positively increase preschoolers' likelihood to choose a competitive payment scheme (Khadjavi and Nicklisch, 2018). This means that education styles and parental attitudes and wishes for their child seem to affect the competitiveness of children and adolescents.

Experimentally elicited preferences for competition have been shown to predict important field behavior of adolescents. Buser et al. (2014) let 14-15 year-old Dutch secondary school students make experimental choices on their willingness to compete and then relate these choices to their selection of academic tracks. The more math- and science-intensive tracks are not only more prestigious, but they are also lead indicators of tertiary education and labor market success later on in life. Buser et al. (2014) find that the willingness to compete is a good predictor of choosing the more prestigious academic tracks, even when controlling for gender, academic abilities and a host of other relevant background variables.

Given the importance of a willingness to expose oneself to competition, Sutter et al. (2016) have explored whether girls can be encouraged to compete by introducing policy interventions like affirmative action programs. In their experiment, they study the effects of quota rules (where among a set of winners a minimum number must be female) and of preferential treatment (mimicking a rule that is applied in many countries, namely that in case of equal qualifications women need to be given preferential treatment in filling a position), and they find that both measures induce in particular girls with high abilities to choose the competitive payment scheme rather than the piece rate. By and large, both measures close the gender gap in the willingness to compete. Another way of closing it is presented in Alan and Ertac (2018b) who have ran an intervention on grit – by which elementary school students learn the role of effort in achievement and are encouraged to become more persistent in tasks. Treated children do not show any gender gap in the willingness to compete, mainly because it seems to increase girls' optimism about their future performance.

SUMMARY COMPETITIVENESS: There is a large gender difference in the level of competitiveness, with girls typically shying away from competition much more often than boys do. This effect is in many, particularly highly developed, countries prevalent already in childhood, and persists during adolescence. Parts of this gender gap can be explained by gender differences in (over)confidence and risk preferences. Competitiveness is lower in children from low SES-backgrounds, and girls' willingness to compete can be increased through affirmative action programs.

Table 37: Competitive preferences

Authors	Subject Pool			Experimental Task	Results			
	age	sample size	country		age	gender	SES	other
Gneezy & Rustichini (2004)	9-10	140	ISR	running task		males improve performance in second round		gender composition boy-boy: improved performance by large margin; girl-girl: worse performance
Zhang (2011)	11-15	544	CHN	math task	gender gap emerges in high school	females less competitive only for ethnic minorities		
Bartling et al. (2012)	5-6	223	GER	flipping toy frogs into pond			low SES children with recent medical condition less competitive	
Booth & Nolen (2012a)	15	260	U.K.	maze task		females less competitive		females in single-sex schools more competitive
Andersen et al. (2013)	7-15	318	IND	throwing tennis ball into a bucket	gender gap emerges at age 13	females from patriarchal society less competitive		older females in matrilineal societies more competitive
Samek (2013)	3-5	123	U.S.	toy fishing task				
Buser et al. (2014)	14-15	362	NL	math task		females less competitive		competitiveness can predict study track
Dreber et al. (2014)	15-19	216	SWE	math and verbal task		females less competitive in math task		

Competitive preferences (cont.)

Authors	Subject Pool			Experimental Task	Results			
	age	sample size	country		age	gender	SES	other
Khachtaryan et al. (2015)	7-16	824	ARM	running, skipping rope, math, verbal task	older children compete less in math task	no gender gap in tournament entry		
Sutter & Glätzle-Rützler (2015)	3-18	1,570	AUT	running, sorting, math task		females less competitive in all tasks		competitiveness persists over two years
Almås et al. (2016)	14-15	523	NOR	math task		females less competitive	low SES children less likely to choose competition (especially males)	
Sutter et al. (2016)	10-17	588	AUT	math task	older males more competitive	females less competitive		preferential treatment of females increases competitiveness
Alan & Ertac (2018b)	10	1,900	TUR	math task		females less competitive		grit and role model interventions increase competitiveness
Khadjavi & Nicklisch (2018)	3-6	84	GER	running task				if parents are highly ambitious even slow children opt for competition

3.9 CONCLUSION

Over the past 15 years, and with accelerating pace, experimental economics has discovered a strong interest in the economic decision making of children and adolescents.The experimental method has created plenty of opportunities to study the development of economic behavior in pre-adulthood. Age, gender, and other factors such as socio-economic background or in-group favoritism have been found to shape economic behavior of children and adolescents, with the latter often showing patterns of behavior that are similar to the evidence from adult subject pools. Acquiring knowledge on the development of behavior and the factors shaping it is a prerequisite for potential policy interventions that aim at promoting one type of behavior more than another. On the basis of the research described in this survey, a new wave of intervention studies has very recently got off the ground and it will provide many very important insights in the years to come. For instance, the intervention studies by Alan and Ertac (2018a) or Kosse et al. (2018) have shown ways to make young children more patient in their intertemporal choices and more prosocial and fairness-oriented in allocation tasks. This survey may prove useful for future intervention studies as it has intended to provide a diagnostic picture of what we know about the patterns of economic behavior before adulthood.

As such, this survey, as any other, has been selective. We have concentrated on the types of economic behavior that we consider most important and about which there has been most research done to date, namely risk, time, and social preferences, and competitive behavior. For each of these, subdivided into seven sections (including a section on rationality of children's choices), we have attempted to present an overview of the main determinants identified in the main papers on a specific dimension of behavior. The summaries at the end of each section have captured the main findings, in particular the influence of age and gender. Given the summaries in each section, however, we are not going to repeat the main findings here (please refer to the end of each section), but rather devote the conclusion of the paper to a few topics not covered here and to a brief outlook of what might come next.

While we have tried to focus on the topics that we deem most important – and where most of the experimental economics research with children and adolescents has been done – there are a few other fields that have not been covered or where there is hardly any research up to date. There is a small literature on honesty, respectively deception, in children and adolescents (e.g., Bucciol and Piovesan, 2011, Glätzle-Rützler and Lergetporer, 2015, Houser et al., 2016, Maggian and Villeval, 2016). This literature deals with the question how honesty as a social norm develops and what are the driving factors as to why subjects do not exploit informational asymmetries. Since informational asymmetries may trigger inefficiencies on markets, a better understanding of the conditions under which humans act honestly even under incentives for cheating may provide a better understanding of how social norms (here: of honesty) affect market outcomes. A topic that has not received attention in the experimental economics literature about children so far is the study of coordination games (yet, there are psychological papers on these games; see, e.g., Grueneisen et al., 2015a, Grueneisen et al., 2015b). Given the multiplicity of equilibria, these games are interesting to study with children and adolescents in order to see on which equilibria they might be able to coordinate on and which cues might help to coordinate efficiently. Likewise, information cascades might be interesting to study with children and adolescents because that might help us to understand better how fads (among youths) emerge and what is needed to break information cascades. A recent line of work has started to investigate other non-cognitive skills or personality traits and their relationship to economic preferences. For example, Bucciol et al. (2011) have investigated the role of temptation – and the ability to resist it – on productivity. Alan and Ertac (2018b) and Alan et al. (2016) have studied the role of grit for economic behavior and how curricular interventions can foster it. Deckers et al. (2017) draw an even larger picture by asking how socio-economic status of families shapes a child's personality. Understanding all of these relationships better will help us understand how policy interventions might have desired and how they might avoid undesired effects. So far, for instance, little is known about what happens to one type of economic behavior – say with respect to social preferences – when another type of economic behavior – for instance intertemporal choice in time preference experiments – is targeted in a specific intervention. While it is highly welcome to understand how interventions shape economic behavior in one particular domain – like in the intervention on time preferences in Alan and Ertac (2018a) or on social preferences in Kosse et al. (2018) – one important avenue for future research will be to examine also potentially unintended side-effects on other preference types. Of course, as always in science, the community will take step by step: first collect more knowledge about how specific interventions affect specific behavior – and we need certainly more evidence about what works and what does not – and then also consider interaction effects with other behavior or other personality traits. Besides these immediate steps, another challenge for the future will be to examine the long-run consequences of economic preferences and behavior of children and adolescents on lifetime outcomes, such as educational achievements, success on labor markets or a subject's health status. The relation of time preferences to these long-term outcomes is already fairly well understood, but the knowledge about long-run effects is more limited in the other domains discussed here. In sum, there is plenty of promising work ahead for the ever growing community that uses experimental economics as a tool to understand economic behavior of children and adolescents.

CHAPTER 4

COORDINATION GAMES PLAYED BY CHILDREN AND TEENAGERS – ON THE INFLUENCE OF AGE, GROUP SIZE AND INCENTIVES

Abstract

Efficient coordination is a great source of efficiency gains. We study in an experimental coordination game with 819 children and teenagers, aged 9 to 18 years, what determines efficient coordination. We find that, generally speaking, coordination gets less efficient with increasing age, but that smaller group sizes and larger incentives increase the likelihood of efficient coordination. Beliefs play an important role as well, as subjects coordinate more likely on the efficient equilibrium when they expect others to do so as well. All of our results are robust to controlling for individual risk and time preferences and for gender.

This chapter is joint work with Daniela Glätzle-Rützler and Matthias Sutter.

4.1 INTRODUCTION

An efficient organization of groups requires that group members are able to coordinate their actions successfully in order to achieve potentially large efficiency-gains from interaction. Examples abound across many different fields (Ochs, 1995; Camerer, 2003; Devetag and Ortmann, 2007). Think of hunting together or defending one's territory against potential invaders, coordinating on a ceasefire in military conflict, but also working together in modern companies with their manifold division of labor; or imagine companies coordinating on a new technological standard when competing platforms are available and customers benefit from a common standard. In such situations, typically multiple equilibria exist, such as several platforms that can serve as technical standards or all group members hunting a large prey, like a stag, rather than each hunting a small animal like a rabbit individually.

The ability to coordinate actions is a cornerstone for the flourishing of groups and it has important welfare implications. In many cases the multiple Nash equilibria in coordination games can be ranked according to their overall efficiency that is generated from the players' actions (Cooper et al., 1990; Van Huyck J. et al., 1990; Duffy and Feltovich, 2006; Feri et al., 2010; Cason et al., 2012; Cason et al., 2012; Brandts et al., 2016). Coordination failure creates inefficiencies, which can take on two forms: either players do not coordinate on the same action, but rather pick different actions and therefore mis-coordinate, or they coordinate on an equilibrium that is not the most efficient one (Brandts and Cooper, 2006; Weber, 2006).

Given the importance of the ability to coordinate actions when subjects have to interact with each other and multiple equilibria exist, it is interesting to note that there is hardly any knowledge about when humans learn to coordinate efficiently. With the exception of very few papers discussed below it remains an open question whether children and teenagers are able to coordinate on efficient outcomes and whether this ability depends in predictable ways on economic incentives and the number of interaction partners in the group.

In our paper, we study how 819 children and teenagers, aged 9 to 18 years, played a simple coordination game – a stag-hunt game Cooper et al. (1990) – in which there were only two actions available that could lead to two different equilibria; one equilibrium that is Pareto-efficient – which we will call the efficient equilibrium henceforth – and another equilibrium that is Pareto-dominated, but less risky than the Pareto-efficient one – we call this second equilibrium the inefficient one. Our primary interest lies in the influence of age on the ability to coordinate on the efficient equilibrium. Additionally, we examine two important dimensions that have been investigated with adults, but not yet with children: (1) whether economic incentives play a systematic role, i.e., whether children and teenagers are more likely to achieve the efficient equilibrium when it becomes more profitable or when playing it becomes less risky, and (2) whether coordination is more efficient when the group size is smaller. We want to examine whether already children and teenagers react systematically to changes in these parameters of a coordination game. In addition to studying the influence of external parameters of the game, we are able to link a subject's risk attitude and self-control as well as expectations about other subjects' behavior to the strategy chosen in the coordination game.

To the best of our knowledge there are no papers in experimental economics on how children play coordination games. However, a few developmental psychologist have investigated how young children play such games. Grueneisen et al. (2015b) have studied whether theory of mind is important for efficient coordination in a type of stag-hunt coordination game, played by 6 year-old children. Theory of mind is developed around the age of 6, and it allows subjects to put themselves into the shoes of another person. This ability is supposedly influential in coordination games because the strategic uncertainty about the interaction partner's behavior is important for one's own choices. Grueneisen et al. (2015b) find that children with higher theory of mind are better able to coordinate on more efficient outcomes. Their study does not consider a broader age range, however. This is done in Grueneisen et al. (2015a) where the authors let children aged 3 to 8 years play coordination games. In these games, there is a focal point that can be used to coordinate the players' actions. Yet, all possible equilibria of the game are equally efficient, i.e. there is no ranking of equilibria possible. Grueneisen et al. (2015a) report that 5-year-old children and 8-year-old children are able to coordinate on the focal point with a likelihood that is significantly greater than chance, but 3-year-old children are not yet able to do so. Our paper differs from Grueneisen et al. (2015a) in several aspects. First, our coordination games have two Pareto-ranked equilibria which allows us to study whether children are able to coordinate on the efficient outcome. Second, we do not implement any focal point, thus making efficient coordination considerably harder. Third, contrary to both Grueneisen et al. (2015b) and Grueneisen et al. (2015a) we do not only consider pairs of subjects playing a coordination game, but we also study behavior in larger groups, because many coordination games are played by multiple players. Our treatment variation with respect to group size (pairs of subjects vs. groups of five) allows investigating whether larger groups can also succeed in efficient coordination. Fourth, we consider a broader age range than the previous study by Grueneisen et al. (2015a) by considering children and teenagers aged 9 to 18 years. This age spectrum has not yet been considered in the literature on coordination games, and it covers the area in life where subjects get increasingly involved in economic activities that are also related to coordination problems (e.g., when downloading or exchanging music from different platforms or when choosing a social media platform).

From the experimental economics literature on the behavior of children in economic games we can form some expectations about the potential influence of age on behavior in coordination games. Reasoning in economic games seems to develop below the age of 10 and is rather stable in the teenage years, as Brosig-Koch et al. (2015) have shown in an experiment on a backward-induction task (that has a unique equilibrium). Similarly, Czermak et al. (2016) have found in a series of normal form games (also with unique equilibria) that strategic thinking (i.e., the ability to take into account the interaction partner's choice set when making one's own decision) is rather stable in the teenage years. Yet, in some of the games strategic thinking seems to improve slightly with increasing age. A priori, it is unclear what the implications for coordination games would be. In coordination games, strategic uncertainty is given because a player needs to form an expectation about the other players' actions in order to best respond. This does not necessarily imply that more sophisticated strategic thinking would lead to more efficient coordination. The direction of influence will depend on expectations, and they might depend on age. More pessimistic expectations might lead to less efficient coordination with increasing age, if larger strategic uncertainty leads to more pessimistic expectations as subjects get older.

Besides investigating age effects, we also examine the influence of incentives and of group size. With respect to incentives, the prediction is straightforward (see Schmidt et al. (2003) or Brandts and Cooper (2006) for coordination games with adults). The efficient equilibrium in our stag-hunt game should become more likely when it becomes more profitable – by increasing the payoffs in the efficient equilibrium – or when it is less risky – by increasing the payoffs in case of failure to coordinate on the efficient equilibrium. With respect to the effects of group size, experimental coordination games show typically more efficient play in smaller groups than in larger groups, and such patterns have been found consistently with adult participants (see Weber (2006) and his Table 2 for a summary of group size effects, but also the survey of Devetag and Ortmann (2007)). We vary group sizes by letting subjects play once in pairs and once in groups of five, expecting more efficient coordination in smaller groups.

Our experiment has the following major results: (1) Pooling across group sizes and different incentives, we find hardly any significant influence of age. When looking into more detail, we see that in groups of five coordination becomes systematically less efficient with age. This is largely driven by older subjects having more pessimistic expectations about the other players' choices in larger groups, meaning that strategic uncertainty gets more of a concern for older subjects. In groups of two subjects, there is no clear cut age effect in the aggregate, while for different incentives there are opposing age trends. (2) As expected, incentives play a systematic role. The efficient equilibrium strategy is chosen more often as it gets more rewarding and less risky. (3) Coordination is much easier and more efficient in groups of two subjects than in groups with five subjects. A larger group size entails more strategic uncertainty, which has a negative effect on efficiency.

The rest of the paper is organized as follows. In section 2 we present the experimental design and our experimental treatments. In section 3 we present the results, starting with the highest level of aggregation and continuing with a more detailed description of results by investigating behavior in the different treatments. Section 4 concludes the paper.

4.2 EXPERIMENTAL DESIGN AND PROCEDURE

4.2.1 EXPERIMENTAL PROCEDURE

We conducted our study in seven different schools in Tyrol (Austria), including 28 different classes that comprised fourth grade (with 9-10 years old children), sixth grade (11-12 years), eight grade (13-14 years), tenth grade (15-16 years) and twelfth grade (17-18 years).[1] In total, we had 819 participants, with a slight surplus of girls (see Table 41), as one of the schools was a girls-only school. Our study was approved by the internal review board at the University of Innsbruck and the state board of education in Tyrol. Further consent was given by the headmasters of each school. Prior to the experiment, parents received an information leaflet explaining the nature of the study, without revealing details of the experiment, though. Parents could opt their child out of the experiment, but none did. Participation was, of course, voluntary for children, but all of them agreed to participate. The experiment was run during regular school hours. Given that all children of the randomly selected classes in each school participated, there was no self-selection effect into experimental participation.

The experiment was computerized (using zTree, Fischbacher (2007)) and run with a portable lab. Instructions were handed out in written form, but also explained verbally by taking great care that children and teenagers could understand the rules of the coordination game and in particular how

[1] This study was part of a larger experimental series over the course of 2 years in which we also elicited (at different dates) risk and time preferences or competitive preferences (see Sutter et al. (2013) and Sutter and Glätzle-Rützler (2015))

Table 41: Number of participants and gender composition

School grade (age of students)	Number of subjects per treatment "9/1"	"9/5"	"13/1"	Total (female)
4 (9/10 years)	42	39	41	122 (61)
6 (11/12 years)	88	59	59	208 (118)
8 (13/14 years)	79	61	83	223 (126)
10 (15/16 years)	62	36	44	142 (88)
12 (17/18 years)	49	32	45	126 (67)
Total (female)	320 (194)	227 (105)	272 (161)	819 (460)

Notes: Experimental participants earned points in the experiment that were exchanged into Euro at the following age-dependent rates: €0.3 per point for 9-10-year-olds, €0.5 for 11-12-year-olds, €0.7 for 13-14-year-olds, €1.4 for 15-16-year-olds and €2.2 for 17-18-year-olds.

to interpret the payoff matrix. Before proceeding to the experiment, we had two control questions that were checked by the experimenter. If a child had picked a wrong answer the experimenter carefully explained the instructions once more.

The experiment was incentivized, and each point earned in the experiment was exchanged into Euro at a pre-specified exchange rate which was announced in the beginning. This exchange rate differed by school grades, and was made proportional to the average weekly allowance across the whole age range. In the legend to Table 41 we present the exchange rate for each age group.

4.2.2 THE EFFECTS OF INCENTIVES – BETWEEN-SUBJECTS TREATMENTS

The experimental design had both a between-subjects and a within-subjects variation. The between-subjects variation concerned the payoff matrix of the coordination game. Here we had three different treatments which are shown in Table 42. In each treatment, subjects could choose between action "One" and action "Two". The combination of one's own action and the minimum action chosen by the other players in one's group determined the payoffs. This general structure yielded two equilibria of the game, one where all players in a group play "Two" – which is the Pareto-efficient equilibrium – and one in which all players play "One" – which is the inefficient, but less risky equilibrium. The choice of "One" always yielded 8 points for the active player.

Table 42: Payoff matrices in the three treatments

Treatments	"9/1"		"9/5"		"13/1"	
	Minimum action of other group members					
Own action	Two	One	Two	One	Two	One
Two	9, 9	1, 8	9, 9	5, 8	13, 13	1, 8
One	8, 1	8, 8	8, 5	8, 8	8,1	8, 8

Notes: The Pareto-efficient equilibrium (Two, Two) is depicted in the top left corner of each treatment, the inefficient equilibrium (One, One) in the bottom right corner. First numbers in a cell present the row player's payoff. All players saw themselves in the role of the row player. The options "One" and "Two" in the columns of each treatment indicate the minimum choice of all other players in the group (excluding the row player), and second numbers in each cell show the payoffs of the second player only in the games with a group size of two subjects.

The three treatments shown in Table 42 differed with respect to the possible payoffs from choosing strategy "Two". The treatment labels indicate the two possibilities in case of picking "Two". In treatment "9/1", presented on the left hand side of Table 42, a player earned 9 points if the minimum choice in the group was "Two", but earned only 1 point if someone in the group chose "One". Of course, the efficient equilibrium (of all players earning 9 points) requires all group members to choose "Two", but this is risky, because if only one group member chooses "One", then the payoff drops to 1 point for all players who chose "Two".

The other two treatments, shown in the middle and on the right hand side of Table 42, vary either the payoff in the efficient equilibrium or the payoff in case of mis-coordination (i.e., the payoff for the row player in the upper right cell of the matrix). In treatment "9/5" it becomes less risky to play "Two" because of a payoff of 5 – rather than 1 as in "9/1" – if another group member chooses "One". Finally, in treatment "13/1" the payoff in the efficient equilibrium increases from 9 - as in "9/1" - to 13, keeping everything else constant.

Given that playing "Two" becomes less risky in "9/5" and more profitable in "13/1" we expect to observe more choices of "Two" in these two treatments than in "9/1". It is unclear what to expect from a comparison of "9/5" and "13/1", however. Both treatments yield an expected payoff of 7 from playing "Two" under the assumption that it is equally likely that the minimum chosen by the other players is "One" or "Two". If increased payoffs in the efficient equilibrium (13 points instead of 9) weigh in larger than the increased payoffs in case of mis-coordination (5 points instead of 1), then we should observe more choices of "Two" in "13/1", otherwise the reverse should be true.

4.2.3 THE EFFECTS OF GROUP SIZE – WITHIN-SUBJECTS VARIATION

While the effects of incentives were examined in a between-subjects variation, we examined the possible effects of different group sizes by letting each subject play two games (with the same incentives each) with different group sizes. The "2-persons groups" had a two members, meaning that the row player's payoff in any of the treatments shown in Table 42 depended on the own choice and the choice of the only other player in the group. The

“5-persons groups” had five group members. Given the row player’s choice, the payoff for the row player from choosing “Two” was then determined by the minimum choice of the other four players in the group. In other words, choosing “Two” paid off for the row player only when all four other group member chose “Two” as well.

Both games were explained in the beginning of the experiment (see experimental instructions in the appendix), yet we balanced the order in which the two games were introduced in order to control for potential order effects. After we had worked through the instructions and participants had answered two control questions, we let them make their decisions in the order as explained in the instructions. At the end of the experiment, it was randomly determined which game (first or second) was payoff relevant.

Subjects had no information about the other group members other than that they were attending the same school grade, and would therefore be of the same age. It was additionally explained that the matched subjects did not attend the same school to exclude any potential influence of personal relationships in the decision making process. Given the latter matching procedure, all subjects were only paid after all data had been collected. This was done within 1-2 weeks after executing the experiments. The money was handed over in sealed envelopes with an anonymous ID-code on it.

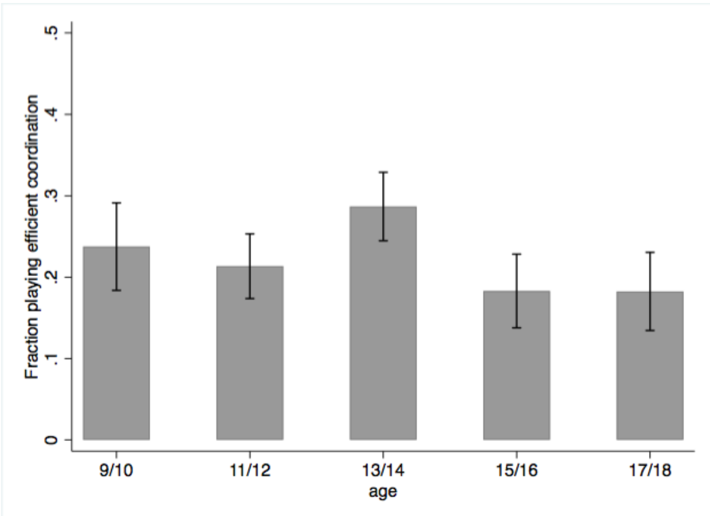
4.3 EXPERIMENTAL RESULTS

We start the results section by presenting first a descriptive analysis of our data. After that, we present a regression analysis that includes further controls, such as personal characteristics and risk and time preferences and a subject’s expectation about other group members’ behavior. Recall that the effects of group size were examined in a within-subject design in which the group size was randomly ordered (first two persons and then five persons, or vice versa). We did not find any order effects (see Table A-1 in the appendix) for which reason we pool the two different orders.

4.3.1 DESCRIPTIVE ANALYSIS

We begin with a highly aggregated analysis by first looking at the main effects of age, incentives, and group sizes. Figure 41 presents the relationship between age (on the horizontal axis) and the likelihood of choosing strategy "Two" (on the vertical axis). It shows for each age group the relative frequency with which subjects across all treatments and across both group sizes chose strategy "Two". The overall relative frequency is 22.7%. Across the five different age groups, we notice an up and down, with a slight tendency of less efficient choices as subjects get older. In the next subsection we are going to examine this potential age trend in more detail.

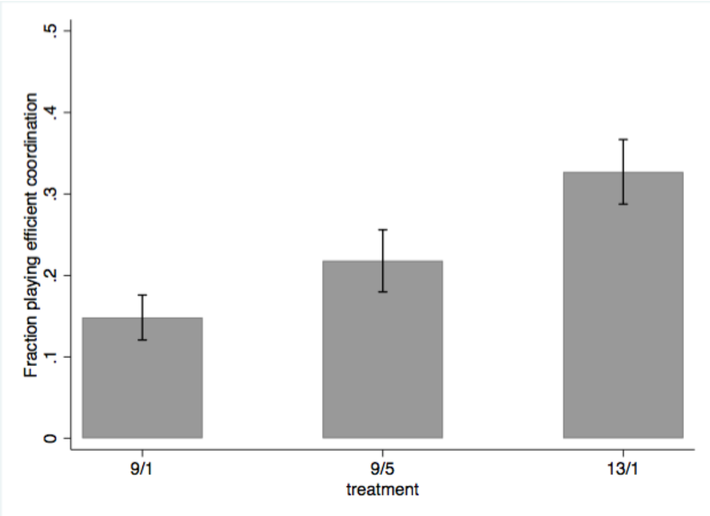
Figure 41: Efficient Coordination by Age



Notes: Relative frequency of choosing the efficient strategy "Two", contingent on age.

Figure 42 shows a very clear treatment effect of incentives on playing the efficient strategy. Here we pool over all age groups and over both group sizes. As predicted, the relative frequency of choosing strategy "Two" is lowest in treatment "9/1" with only 15%. Both of the other treatments have a considerably larger fraction, with 22% in "9/5" and 33% in "13/1", indicating that higher payoffs in the efficient equilibrium (of "13/1") are even more conducive to efficient play than reducing the risk of failing to coordinate on the efficient equilibrium (in “9/5”).

Figure 42: Efficient Coordination and Incentives

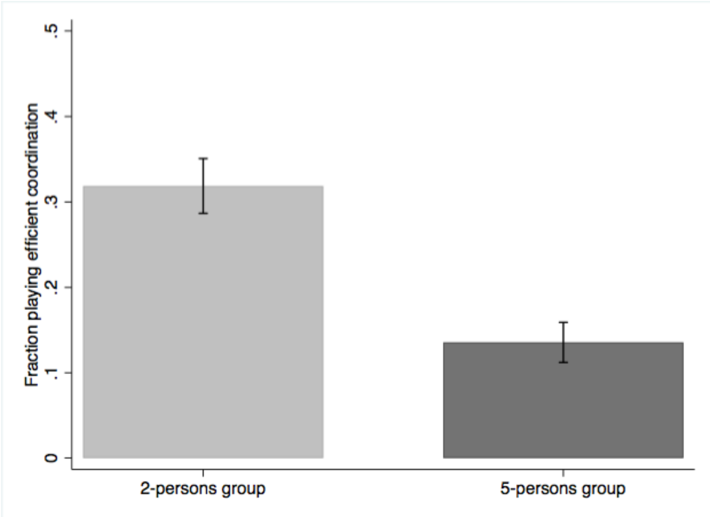


Notes: Relative frequency of choosing strategy “Two”, contingent on incentives (between-subjects treatments)

Figure 43 presents results for the potential group size effect, studied through our within-subject variation. Here we pool across all age groups and across the three between-subjects treatments. As expected, we see a much higher relative frequency of choosing the efficient strategy “Two” in groups with two players (32%) than in groups with five players (14%). Obviously, children and teenagers react to the increased strategic uncertainty in larger groups by choosing the safe strategy “One” much more often, i.e., in about six out of seven cases in groups with five players.

In a next step, we consider several interaction effects of our main factors of interest. 44, we show how age and group size interact with each other. The figure shows the relative frequency of choosing strategy “Two”, separately for each age group in the 2-persons groups and the 5-persons groups. The larger degree of strategic uncertainty in 5-persons groups induces fewer attempts to coordinate on the efficient equilibrium. While for 2-persons groups there

Figure 43: Efficient Coordination and Group Size



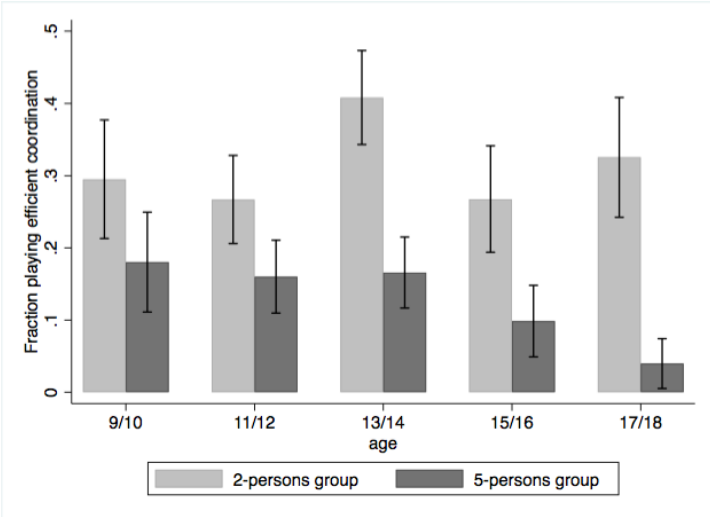
Notes: Relative frequency of choosing strategy "Two", contingent on group size (within-subjects variation).

is no significant age trend, for 5-persons groups we observe a downward trend ($p<0.01$, Cuzick's Wilcoxon-type test for trend), indicating that older subjects choose the efficient strategy less often in 5-persons groups than younger subjects do.

The downward trend in choosing "Two" in the 5-persons groups is nicely mirrored in subjects' expectations about the other group members' behavior, as we show in Figure 45. There we display the relative frequency with which subjects expect the strategy "Two" as the minimum choice of their other group members. This relative frequency is clearly higher in 2-person groups (light bars) than in 5-person groups (dark bars), reflecting more pessimistic expectations when the group size – and thus the extent of strategic uncertainty – gets larger. We observe also a clear downward trend with age in Figure 45, which might be a driving force for subjects to choose less often the efficient strategy "Two" when they get older. The regression analysis in section 3.2 will get back to this relationship.

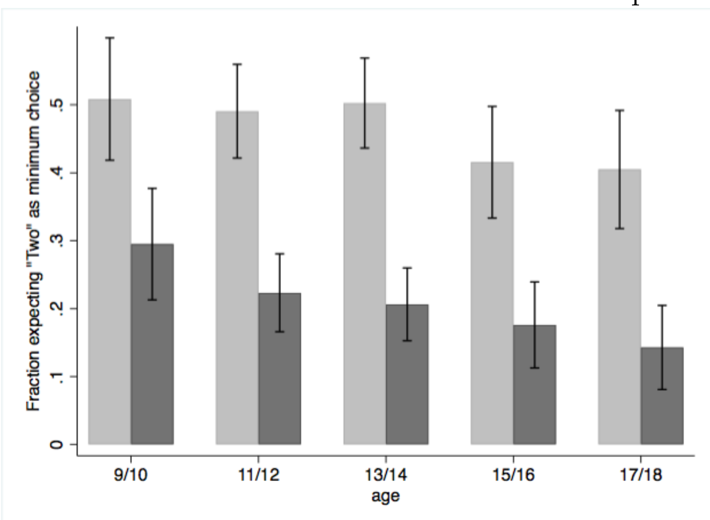
Figures 46 and 47 add the treatments on incentives to the picture drawn in Figure 44, and present the data separately for 2-persons groups (Figure

Figure 44: Efficient Coordination and Age and Group Size



Notes: Relative frequency of choosing strategy "Two", contingent on interaction of age and group size.

Figure 45: Beliefs - Relative Frequency of expecting "Two" as the Minimum Choice in a Group

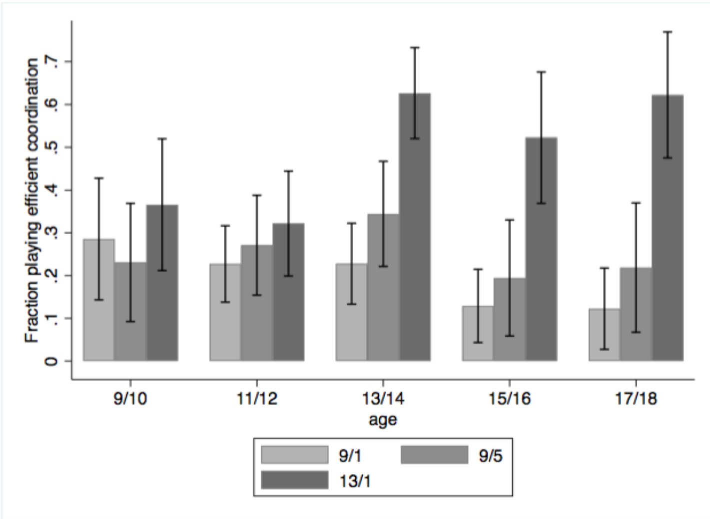


Notes: Subjects' expectations about others' behavior. Relative frequency with which subjects expect "Two" to be the minimum choice in their group, contingent on group size.

46) and 5-persons groups (Figure 47). In Figure 46 we see from the light bars that the choice of the efficient strategy “Two” gets less likely with age in treatment “9/1”, from the gray bars in the middle of each age group that it is basically stable in treatment “9/5”, and from the dark bars that it increases strongly in treatment “13/1” from around 30% in the younger age groups to around 60% in the older age groups. Hence, there is a diverging effect of age here. When efficient coordination is very risky in treatment “9/1” it becomes less likely with age ($p<0.05$, Cuzick's Wilcoxon-type test for trend), when it is very profitable in treatment “13/1” it becomes more likely with age ($p<0.01$, Cuzick's Wilcoxon-type test for trend).

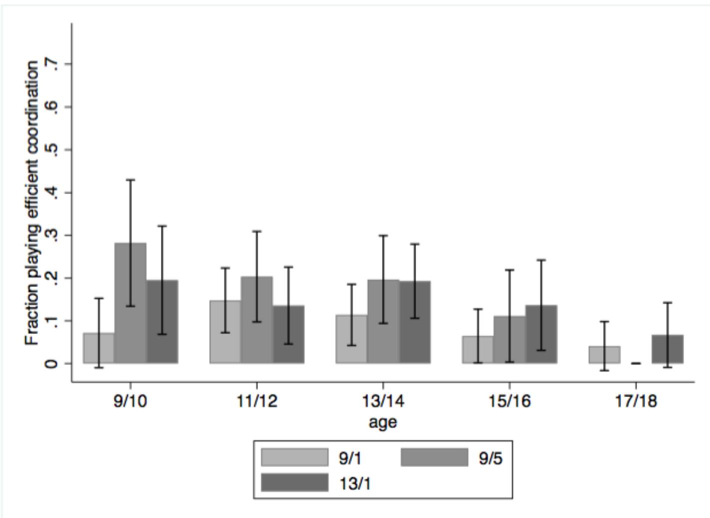
In Figure 47 we show the data for the 5-persons groups, contingent on age and incentives. Here we notice an overall pattern of a decline in efficient coordination with increasing age, and this trend is similar in all treatments, hence independent of the prevailing incentives. This means that the larger strategic uncertainty in 5-persons groups leads to fewer attempts of achieving the efficient equilibrium, and that the increased incentives in “13/1” or the lower risk in “9/5” do not have an impact on this downward trend in larger groups.

Figure 46: Efficient Coordination in 2-Persons Groups



Notes: Relative frequency of efficient strategy "Two" in 2-persons groups, contingent on age and incentives (between-subjects treatments).

Figure 47: Efficient Coordination in 5-Persons Groups



Notes: Relative frequency of efficient strategy "Two" in 5-persons groups, contingent on age and incentives (between-subjects treatments).

4.3.2 REGRESSION ANALYSIS

In the following, we present an econometric analysis of coordination behavior. We use probit regressions to estimate the likelihood of subjects choosing the efficient strategy "Two".[2] Recall that each subject played both in the 2-persons and 5-persons groups (in random order), but that the incentive effects were studied in a between-subject design. Hence we present in Table 43 separate estimations for treatments "9/1", "9/5", and "13/1". In each estimation we cluster at the level of the individual subject because each subject made two choices, one for each group size.

The first model for each treatment includes as independent variables age (measured in years and months), a dummy variable for 5-persons groups, an interaction term for age and large group, and a subject's expectation about the minimum choice of the other group member(s). We code the variable "Expected efficiency" as 1 if a subject expects the minimum choice to be "Two" in the group, and as 0 if the expectation is that at least one other group member chose "One".

In the second model for each treatment we add gender (with a female dummy) and data on (most) subjects' risk and time preferences. Risk attitudes were elicited utilizing the framework of Ellsberg's two-color choice task (Ellsberg, 1961) (see Figure A-1 in the appendix for further information). Subjects could choose either a lottery that paid €10 or zero with equal probability, or a safe amount of money that increased in increments of €0.50 from €0.50 to €10. As a measure of a subject's risk aversion r we use

$$r = 1 - \frac{CE}{10}$$

where CE denotes the certainty equivalent of the risky prospect. The variable risk aversion can range from 0 (very risk loving) to 1 (very risk averse), with 0.5 as risk neutral.[3]

[2] OLS estimations yield basically the same results as those reported in the main text from the probit regressions.

[3] For roughly one quarter of the sample we examined the effects of stake sizes by increasing the lottery prize from €4 for 10-year-olds in steps of €2 up to €12 for 18-year-olds. We did not find any stake size effect and hence control for this stake size variation by normalizing the CE to €10.

A subject's impatience was measured using time preferences decision sheets, giving students the option of choosing a fixed amount at an earlier point in time or waiting for an increased payoff at a later point in time (see Figure A-2 in the appendix).[4] The variable "Impatience" measures the frequency with which a subject chose the earlier (and smaller) payoff. Larger values indicate higher impatience.

Table 43 presents the results of our estimations. Below the table we present a series of post-estimation Wald-tests to test for joint significance effects. Overall, the following main patterns emerge from our data (and confirm our descriptive analysis). When subjects make choices in 2-persons groups, there is a negative age trend for choosing the efficient strategy "Two" in treatment "9/1", but a positive age trend in treatment "13/1". Hence, only when incentives to aim for the efficient equilibrium are relatively high, then older subjects are more likely to choose this strategy. When playing this strategy is very risky (as in treatment "9/1"), however, older subjects shy away from it more often.[5] In 5-persons groups, we find a negative and significant age effect for treatment "9/5". We observe no significant age effect in treatments "9/1" and "13/1", even though the likelihood of choosing "Two" goes down, by and large, with age (see Figure 47).

[4]Each subject faced 8 different decision sheets with 20 decisions each. Only one of the choices was paid out in the end. The decision sheets varied in a 2x2x2 factorial design the size of the fixed payoff at the earlier point in time (either €10.10 or €4.05), the waiting time between the earlier payoff and the later payoff date (3 weeks or 1 year) and the up-front delay (no up-front delay or 3 weeks up-front delay).

[5]The explanatory variable "age" in the regression shows the age effect for the reference category 2-persons groups. The post-estimation Wald-tests beneath the table repeat this age effect (for completeness) and show the age effect for 5-persons groups.

Table 43: Probit estimation of determinants of efficient coordination. Dependent variable: likelihood to choose strategy "Two"

Dep. var.: Efficient coordination	(1)	(2)	(3)	(4)	(5)	(6)
	9/1	9/1	9/5	9/5	13/1	13/1
Age	-0.0707**	-0.0637*	0.0174	0.0377	0.106***	0.0799**
	(0.036)	(0.038)	(0.037)	(0.039)	(0.030)	(0.035)
Group size 5	-0.731	-0.538	2.003***	2.510***	1.239**	1.049
	(0.643)	(0.703)	(0.716)	(0.794)	(0.592)	(0.788)
Age*group size 5	0.0291	0.0155	-0.146***	-0.183***	-0.136***	-0.130**
	(0.044)	(0.048)	(0.052)	(0.057)	(0.042)	(0.054)
Expected efficiency	1.132***	1.179***	1.119***	1.054***	1.302***	1.416***
	(0.140)	(0.151)	(0.147)	(0.165)	(0.134)	(0.163)
Female		-0.119		-0.119		-0.158
		(0.158)		(0.173)		(0.148)
Risk aversion		-0.102		-0.746**		-0.433
		(0.311)		(0.371)		(0.381)
Impatience		-0.00176		-0.00147		-0.000566
		(0.002)		(0.002)		(0.002)
Intercept	-0.369	-0.224	-1.536***	-1.218*	-2.289***	-1.581***
	(0.512)	(0.556)	(0.539)	(0.640)	(0.446)	(0.544)
N	640	584	454	380	544	452
Wald tests (p-values)						
H0: No age effect for ...						
...group size 2 ($\beta age = 0$)	0.0474	0.0940	0.6411	0.3335	0.0005	0.0237
...group size 5 ($\beta age + \beta age * groupsize5 = 0$)	0.2634	0.2053	0.0012	0.0020	0.4267	0.2925
H0: No large group effect for ...						
...10-year-olds ($\beta groupsize5 + \beta age * groupsize5 * 10 = 0$)	0.0471	0.1141	0.0168	0.0085	0.5420	0.3499
...12-year-olds ($\beta groupsize5 + \beta age * groupsize5 * 12 = 0$)	0.0130	0.0346	0.1132	0.0787	0.0068	0.0067
...14-year-olds ($\beta groupsize5 + \beta age * groupsize5 * 14 = 0$)	0.0063	0.0099	0.7751	0.7623	0.0000	0.0000
...16-year-olds ($\beta groupsize5 + \beta age * groupsize5 * 16 = 0$)	0.0621	0.0503	0.0878	0.0429	0.0000	0.0000
...18-year-olds ($\beta groupsize5 + \beta age * groupsize5 * 18 = 0$)	0.3137	0.2328	0.0247	0.0076	0.0000	0.0000

Robust standard errors in parentheses. Clustered on the level of individual subjects.

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

With respect to group size, we find that efficient coordination becomes less likely in 5-persons groups, as one can see from the Wald-tests beneath Table 43. Only from these tests it becomes clear that the positive coefficient for the main effect (from the dummy "5-persons group") is reversed to a negative effect when the significantly negative interaction term of age and large group is taken into account properly.[6]

In each single estimation, we note that expectations play a significant role and in the anticipated direction. If subjects expect the other group member(s) to play the efficient strategy "Two", they are more likely to do so themselves. Not a single estimation shows any gender effect, meaning that efficient coordination behavior of children and teenagers does not seem to depend on gender. Risk aversion has always a negative sign, but is only significant in treatment "9/5". Impatience is also always negative, but never significant.

4.4 CONCLUSION

We have studied the evolvement of coordination behavior of children and teenagers in a large experiment with 819 participants, aged nine to 18 years. We have focused on the influence of age, group size, and incentives on the likelihood to play the efficient equilibrium strategy in an experimental stag-hunt game. While the effects of group size and monetary incentives have been investigated previously in coordination games with adults (Van Huyck J. et al., 1990; Brandts and Cooper, 2006; Weber, 2006), the relationship of age and the ability to coordinate efficiently has not received proper attention so far. Given that coordination problems abound also in the pre-adulthood age (e.g. selecting on which social media to be active), we consider it as

[6] More precisely, the joint effect of 5-persons group and the interaction term of age and 5-persons group is significantly negative – thus showing that coordination is less efficient in the larger groups – in all age groups of treatment "9/1", except the 18 years old teenagers; in treatment "9/5" for all age groups except for 14 years old teenagers, and in treatment "13/1" for all age groups except those 10 years old. In treatment "9/5" we observe a reverse group size effect – that coordination is more efficient in the larger groups – but only for 10-years-olds.

important to understand the development of efficient coordination in the (pre)teenage years.

Overall we have found the following main results: age does not have an unambiguous effect, as it shows interaction effects with group size and monetary incentives. In small groups, older subjects are more likely to coordinate on the efficient equilibrium when the payoffs in this equilibrium are highest (in "13/1"), but older subjects are less likely to choose this strategy when it is very risky to do so (in "9/1"). This pattern of results indicates that older subjects react more sensitively to monetary incentives than younger subjects do. In other words, incentives lead to stronger behavioral changes when our subjects are older. The pattern observed for older subjects has the side-effect that, in the aggregate, it looks as if age had no effect at all in small groups.

In large groups, older subjects are less likely to choose the efficient equilibrium strategy in treatment "9/5". In the other treatments, there is no significant age effect, even though the average likelihood of choosing the efficient strategy tends to go down with age. The lack of a clear age effect on coordination behavior is reminiscent of previous results in other domains of economic preferences. Risk and time preferences (Sutter et al., 2013) or preferences for competition (Sutter and Glätzle-Rützler, 2015) do not seem to be affected significantly by age during the teenage years. Similarly, the ability to use backward induction (Brosig-Koch et al., 2015) or the strategic sophistication in normal-form games (Czermak et al., 2016) does not change considerably in the teenager years either.

Our third main finding concerns the relationship of monetary incentives and efficient coordination. The larger these incentives, the more likely it is that subjects choose the strategy that can lead to the efficient equilibrium. Here we have found that increasing the profits in the efficient equilibrium (in "13/1") seems to work better than providing a cushion (in "9/5") for subjects who choose the efficient strategy, but where groups fail to achieve the efficient equilibrium. Overall, our results about the effects of incentives seem to be in line with what has been found in experiments with adults (Brandts and Cooper, 2006; or Devetag and Ortmann, 2007), indicating that already children and teenagers react in a predictable way to monetary incentives in coordination games.

APPENDIX A: TABLES AND FIGURES

Figure A-1: Sample choice list for the measure of risk attitudes

Risk					
[1]	draw from bag A	O	or	O	0.50 euro for sure
[2]	draw from bag A	O	or	O	1 euro for sure
[3]	draw from bag A	O	or	O	1.50 euro for sure
[4]	draw from bag A	O	or	O	2 euro for sure
[5]	draw from bag A	O	or	O	2.50 euro for sure
[6]	draw from bag A	O	or	O	3 euro for sure
[7]	draw from bag A	O	or	O	3.50 euro for sure
[8]	draw from bag A	O	or	O	4 euro for sure
[9]	draw from bag A	O	or	O	4.50 euro for sure
[10]	draw from bag A	O	or	O	5 euro for sure
[11]	draw from bag A	O	or	O	5.50 euro for sure
[12]	draw from bag A	O	or	O	6 euro for sure
[13]	draw from bag A	O	or	O	6.50 euro for sure
[14]	draw from bag A	O	or	O	7 euro for sure
[15]	draw from bag A	O	or	O	7.50 euro for sure
[16]	draw from bag A	O	or	O	8 euro for sure
[17]	draw from bag A	O	or	O	8.50 euro for sure
[18]	draw from bag A	O	or	O	9 euro for sure
[19]	draw from bag A	O	or	O	9.50 euro for sure
[20]	draw from bag A	O	or	O	10 euro for sure

Notes: Subjects could either draw from bag A or choose fixed payment. Bag A contained 10 white and 10 orange balls – if right color was picked from bag payment of €10.

Figure A-2: Sample choice list for the measure of impatience

Time preferences – Decision sheet *(the order of decision sheets was random)*

	amount **today**		or		amount in **3 weeks**
[1]	10.10 euro today	O	**or**	O	10.10 euro in 3 weeks
[2]	10.10 euro today	O	**or**	O	10.30 euro in 3 weeks
[3]	10.10 euro today	O	**or**	O	10.50 euro in 3 weeks
[4]	10.10 euro today	O	**or**	O	10.70 euro in 3 weeks
[5]	10.10 euro today	O	**or**	O	10.90 euro in 3 weeks
[6]	10.10 euro today	O	**or**	O	11.10 euro in 3 weeks
[7]	10.10 euro today	O	**or**	O	11.30 euro in 3 weeks
[8]	10.10 euro today	O	**or**	O	11.50 euro in 3 weeks
[9]	10.10 euro today	O	**or**	O	11.70 euro in 3 weeks
[10]	10.10 euro today	O	**or**	O	11.90 euro in 3 weeks
[11]	10.10 euro today	O	**or**	O	12.10 euro in 3 weeks
[12]	10.10 euro today	O	**or**	O	12.30 euro in 3 weeks
[13]	10.10 euro today	O	**or**	O	12.50 euro in 3 weeks
[14]	10.10 euro today	O	**or**	O	12.70 euro in 3 weeks
[15]	10.10 euro today	O	**or**	O	12.90 euro in 3 weeks
[16]	10.10 euro today	O	**or**	O	13.10 euro in 3 weeks
[17]	10.10 euro today	O	**or**	O	13.30 euro in 3 weeks
[18]	10.10 euro today	O	**or**	O	13.50 euro in 3 weeks
[19]	10.10 euro today	O	**or**	O	13.70 euro in 3 weeks
[20]	10.10 euro today	O	**or**	O	13.90 euro in 3 weeks

Notes: Impatience calculated by summing up all the instances where a subject chose the earlier (and smaller) payoff across all 8 different choice lists High value – impatient, Low value – patient.

Table A-1: Probit estimation of determinants of efficient coordination. Dependent variable: likelihood to choose strategy "Two".

Dep. var.: Efficient coordination	(1)	(2)	(3)	(4)	(5)	(6)
	9/1	9/1	9/5	9/5	13/1	13/1
Age	-0.0687*	-0.0625	0.0169	0.0368	0.106***	0.0794**
Group size 5	-0.712	-0.524	2.008***	2.511***	1.244**	1.060
Age*group size 5	0.0279	0.0147	-0.146***	-0.183***	-0.137***	-0.131**
Expected efficiency	1.123***	1.173***	1.118***	1.050***	1.300***	1.415***
Order	-0.160	-0.134	-0.0282	-0.0217	-0.0484	-0.0613
Female		-0.120		-0.117		-0.159
Risk aversion		-0.111		-0.742**		-0.435
Impatience		-0.0133		-0.00847		-0.00406
Intercept	-0.321	0.590	-1.517***	-0.740	-2.258***	-1.307
N	640	584	454	380	544	452

Robust standard errors in parentheses. Clustered on the level of individual subjects.

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

APPENDIX B: INSTRUCTIONS

Experimental instructions (for "9/1" treatment)[7]

Welcome to our game. Before we start, we will explain the rules of our game. From now on, please don't talk to your neighbor and listen carefully. We will frequently stop during our explanation and allow you to ask questions. Therefore, please raise your hand and one of us will come to you to answer your question. You can earn money in this game. You will receive your payments in an envelope marked with your ID within 2 weeks.
Your payment is denoted in points. At the end of the session we will convert these points into Euros. We apply the following conversion rate:

1 point = 0.3 €(30 cents)[8].

You will receive a conversion table which will tell you how much your earned points are worth. *(Carefully explain the conversion table with a few examples.)*
Everybody ok so far? *Leave time for questions and answer them privately.*
For this game we will build random groups. You will play one part in a 2-persons group and one part in a 5-persons group. Your partner(s) are at a different school but attend the same grade level as you do. The screen will tell you whether you're playing in a 2- or 5-persons group.[9]
Everybody ok so far? *Leave time for questions and answer them privately.*

In this game you have to pick a number (either 1 or 2). Each group member will decide independently of each other. Your payment will depend on your decision and on the decision(s) made by your partner(s). Your payment can be inferred from a table.

For example:

[7] The other two treatments are analogous by changing the respective payoffs.

[8] Conversion rates were age dependent: €0.3 per point for 9-10-year-olds, €0.5 for 11-12-year-olds, €0.7 for 13-14-year-olds, €1.4 for 15-16-year-olds and €2.2 for 17-18-year-olds.

[9] Here we present the order in which the 2-persons game was explained first and the 5-persons game second. This was balanced in other classes where we explained the 5-persons game first and 2-persons game second.

		Your number	
		TWO	ONE
Your number	TWO	9	1
	ONE	-	8

You have 1 partner!

Playing with one partner:
If you choose 1, you will receive 8 points for sure. If you choose 2, your payment will depend on your partner's choice. If your partner chooses 2 as well, you will receive 9 points. If your partner chooses 1, then you will receive 1 point.
Everybody ok so far? *Leave time for questions and answer them privately.*

		Lowest number of all 4 partners	
		TWO	ONE
Your number	TWO	9	1
	ONE	-	8

You have 4 partners!

Playing with four partners:
If you choose 1, you will receive 8 points for sure. If you choose 2, your payment will depend on your partners' choices. Specifically, the minimum number (lowest number) chosen by your four partners will count. If all four partners choose 2 (and you choose 2 as well), then you will receive 9 points. If one or more of your partners choose 1 (and you choose 2), then you will receive 1 point.
Everybody ok so far? *Leave time for questions and answer them privately.*
After completing these two parts, please answer an estimation-question which will appear on your computer screen. If you answer the estimation-question correctly, you will receive one additional point. Before we start, you will have to answer two control questions, which we will check before we start playing the game.

Everybody ok so far? *Leave time for questions and answer them privately.* If nobody has any further questions, please start the experiment.

CHAPTER 5

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 Anna Untertrifaller.

5.1 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). Non-cognitive skills have furthermore been shown to be malleable in children (Almlund et al., 2011; Kautz et al., 2014; Alan et al., 2016). In a recent study Alan et al. (2016) stage a successful school intervention on grit for 10 to 11 year-old children, aimed at improving children's beliefs about the malleability of their own ability. 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.

Larbi and Christian (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 (Larbi and Christian, 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). 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.

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) shows that the willingness to challenge oneself is as well important 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.

We 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 Larbi and Christian (2016) who show a higher likelihood of procrastination in less diligent adult subjects.

Concerning the influencing factors of diligence, we find support for more diligent children to be more likely to choose the challenging task over the easy option. Children were presented 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 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., 2014; Sutter et al., 2013; Alan and

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

Ertac, 2018a). 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.

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. Parents who report to procrastinate more frequently have children who procrastinate more often in the experiment. We find no evidence of parental diligence to be correlated with the children's diligent behavior. Higher education of parents positively affects the child's diligence.

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 5.3 discusses the results, section 5.4 focuses on the influence of family background, and section 5.5 concludes.

[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.

5.2 DESIGN

We ran an experimental study with three to six 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 51). Children were paid in tokens which could be exchanged one-to-one for small presents like balloons, hair clips, key chains, or "minion"-themed stationary 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 51: 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, 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

[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. Participation of kindergartens was voluntary. 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.

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 5.2.3). 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]

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

[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 didn't answer all control questions correctly from our sample.

5.2.2 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]

5.2.3 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 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

[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).

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 "cubicle" and again briefly reminded of the instructions. The remaining procedure closely followed the *today* condition, where children worked on

[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 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.

the task for as long as they chose to and were then paid out according to the number of beads collected.[9]

5.2.4 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.

5.3 RESULTS

The result section first gives an overview of the decisions made in each task (see Table 52). Section 5.3.2 provides a detailed analysis for treatment results, of the effect on time scheduling on diligence. Section 5.3.3 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.

5.3.1 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 51. 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 52: Summary statistics

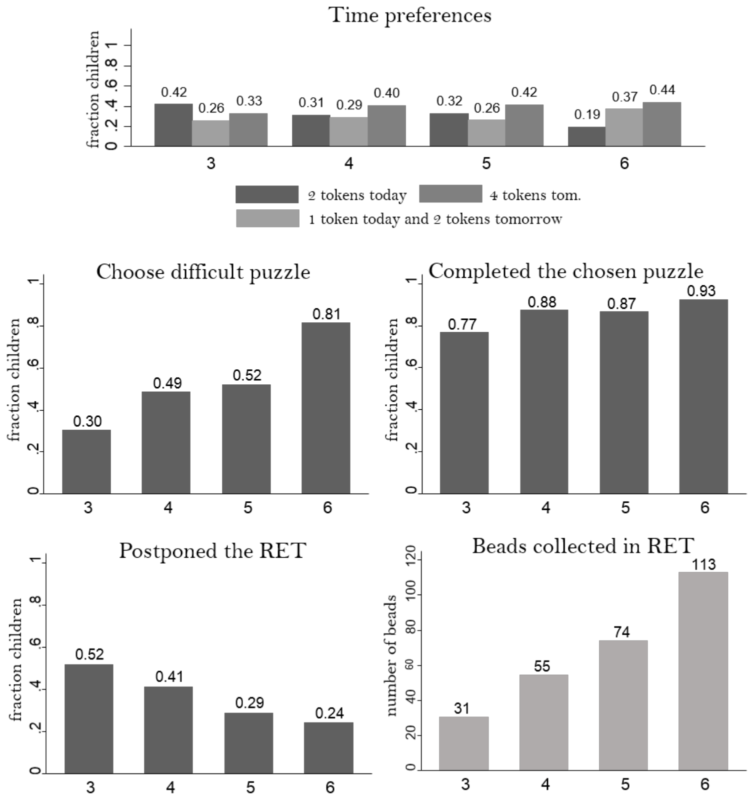
Variable	**Mean**	**N**
PATIENCE		
2 tokens today	0.30	429
1 token today, 2 tokens tomorrow	0.30	429
4 tokens tomorrow	0.40	429
WILLINGNESS TO CHALLENGE ONESELF		
chose difficult puzzle	0.56	429
completed the chosen puzzle	0.87	429
DILIGENCE		
postponed the task	0.34	241
beads collected in ability check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

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.

[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.

Figure 51: Development by age



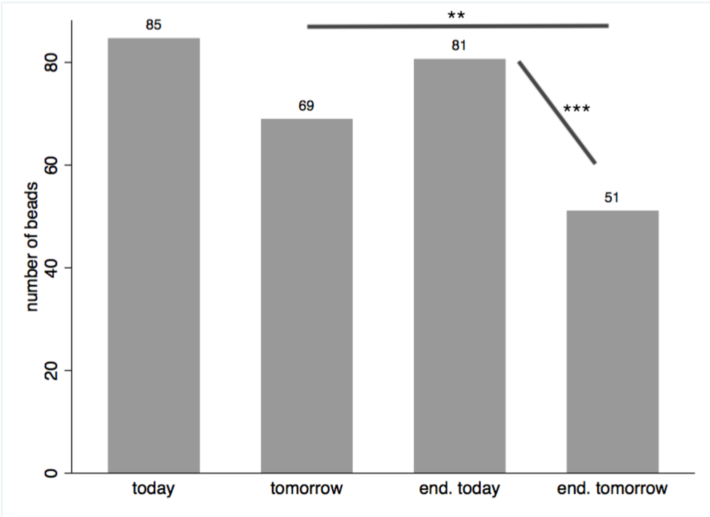
5.3.2 IMPACT OF TIME SCHEDULING ON DILIGENCE

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 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

[12]Treatments were randomized at class level.

next day (*endogenous tomorrow*), while 158 children preferred to work on the task today (*endogenous today*).

Figure 52: 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* vs *endogenous* treatment does not yield any differences in diligence (p=.622, Mann-Whitney-U-test). 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 52). 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=0.700, 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.

Focusing only on the children who procrastinate, we examine the driving forces behind procrastination. Due to the lower number of observations we pool the data by age groups into the younger (3 and 4 year-olds) and older cohort (5 and 6 year-olds) to receive 47 and 36 observations, respectively.[13] Age plays a relevant role in procrastination behavior. Younger children in our sample are more likely to procrastinate the task to the next day (see Figure 53a). 44% in the younger cohort and 27% of the older cohort chose to postpone collecting beads (p=.006, Mann-Whitney-U-test).

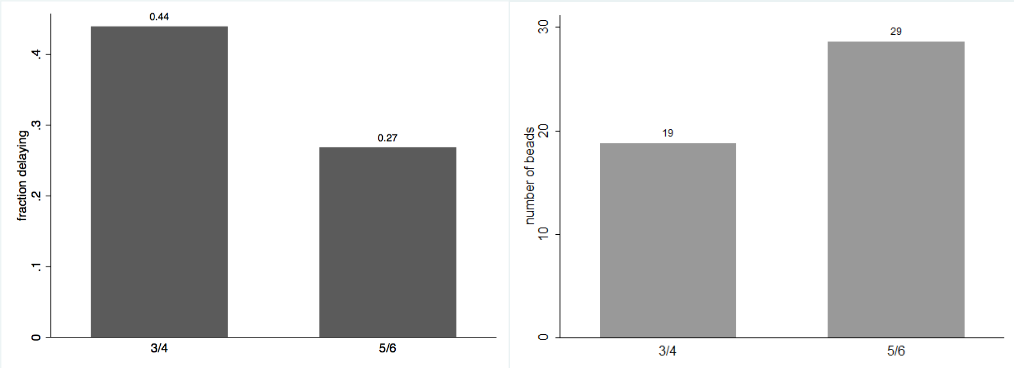
Result 1. *Younger children are significantly more likely to procrastinate the real effort task to the next day.*

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. Figure 53b 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). Children who self-selected into *endogenous tomorrow* performed

[13]The number of children who procrastinated the task by increasing age was 15, 32, 24, and 12.

Figure 53: Procrastination

(a) Procrastination by age (b) Difference in number of beads



Notes: Figure 3a 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. Figure 3b reports the difference in collected beads between the endogenous today and endogenous tomorrow treatment by age group.

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 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.*

5.3.3 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 53) 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 53, 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 53). 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.[14] 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 53. 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

[14] The child was required to complete the puzzle by herself without help or supervision. 88% of children completed the puzzle.

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.[15] [16]

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 53 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.

[15] 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.

[16] 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.

Table 53: 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

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

5.4 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[17]. As highlighted in Table 54, 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

[17] As submission of the questionnaire was voluntary we acknowledge that the data analysis in this section may be influenced by a selection bias of parents regarding parental behavior. Some questionnaires did not contain full information on all questions, accounting for the difference in number of observations.

in the RET task between the samples including parent data and excluding 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 54: 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

5.4.1 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[18], we observe that parents with higher education displayed more patience in the incentivized intertemporal

[18]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).

5.4.2 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 55 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.[19]

[19] 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 55: 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 55). 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.*

5.4.3 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[20] 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

[20]Note: 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 56: 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

5.5 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. As effort provision in the self-selected today treatment is higher, older children in our sample seem more proficient in knowing when they will exert maximal effort and thus select the immediate option, earning more tokens.

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 parent's 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.

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, one might identify those who are more prone to procrastination and offer them advice and support in their endeavor to work more diligently.

APPENDIX A: TABLES AND FIGURES

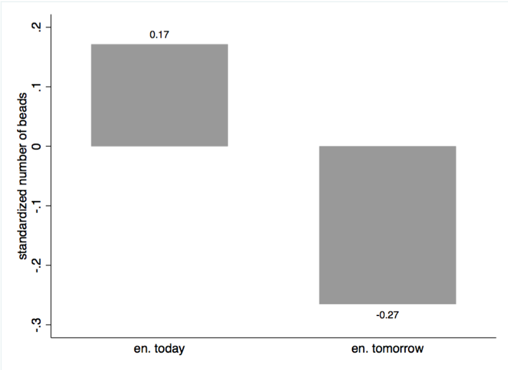
Figure A-1: Puzzle task to challenge oneself

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

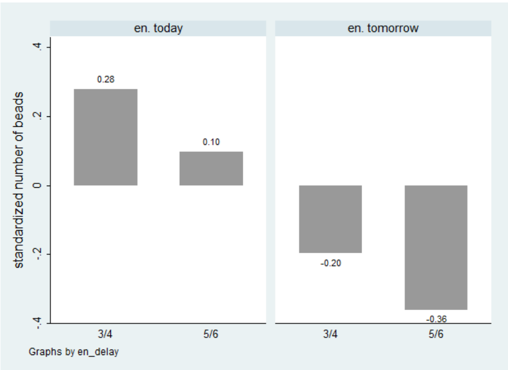
Notes: Children were seated in separate "cubicles" while working on the RET. The RET consisted of collecting the yellow beads out of the green bowl putting them into the yellow bowl.

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 by 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

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

APPENDIX B: 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: In case 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. [21] 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.

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

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

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 just as exciting 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's outcome.

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 whether to do the task right away or to postpone 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.

REFERENCES

Abeler, J., D. Nosenzo, and C. Raymond (2018). Preferences for Truth-Telling. *Econometrica forthcoming*.

Akbaş, M., D. Ariely, and S. Yuksel (2016). When Is Inequality Fair? An Experiment on the Effect of Procedural Justice and Agency. Discussion papers, Available at SSRN: https://ssrn.com/abstract=2474368.

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

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

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

Alan, S. and S. Ertac (2018b). Mitigating the gender gap in the willingness to compete: Evidence from a randomized field experiment. *Journal of the European Economic Association, forthcoming*.

Alesina, A. and G.-M. Angeletos (2005). Fairness and Redistribution. *American Economic Review 95*(4), 960–980.

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

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

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

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

Almås, I., A. W. Cappelen, and B. Tungodden (2016). Cutthroat Capitalism Versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians? Discussion Paper No. 18/2016, NHH Dept. of Economics.

Alstadsæter, A., N. Johannesen, and G. Zucman (2017). Tax Evasion and Inequality. Technical report.

Andersen, S., S. Ertac, U. Gneezy, J. A. List, and S. Maximiano (2013). Gender, competitiveness, and socialization at a young age: Evidence from a matrilineal and a patriarchal society. *The Review of Economics and Statistics 95*(4), 1438–1443.

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

Angerer, S., E. G. Dutcher, D. Glätzle-Rützler, P. Lergetporer, and M. Sutter (2017). Gender differences in discrimination emerge early in life: Evidence from primary school children in a bilingual city. *Economics Letters 152*, 15–18.

Angerer, S., D. Glätzle-Rützler, P. Lergetporer, and M. Sutter (2015a). Donations, risk attitudes and time preferences: A study on altruism in primary school children. *Journal of Economic Behavior & Organization 115*, 67–74.

Angerer, S., D. Glätzle-Rützler, P. Lergetporer, and M. Sutter (2015b). How to measure time preferences in children: a comparison of two methods. *Journal of the Economic Science Association 1*(2), 158–169.

Angerer, S., D. Glätzle-Rützler, P. Lergetporer, and M. Sutter (2016). Cooperation and discrimination within and across language borders: Evidence from children in a bilingual city. *European Economic Review 90*, 254–264.

Apesteguia, J., S. Huck, J. Oechssler, E. Weidenholzer, and S. Weidenholzer (2018). Imitation of peers in children and adults. *Games 9*(1).

Ashby, J. S. and I. Schoon (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.

Ashton, M. C. and K. Lee (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment 91*(4), 340–345.

Balafoutas, L., M. G. Kocher, L. Putterman, and M. Sutter (2013). Equality, Equity and Incentives: An Experiment . *European Economic Review 60*, 32–51.

Barash, J., I. Brocas, J. D. Carrillo, and N. Kodaverdian (2018). Heuristic to Bayesian: The evolution of reasoning from childhood to adulthood. *Journal of Economic Behavior & Organization*.

Bartling, B., E. Fehr, and D. Schunk (2012). Health effects on children's willingness to compete. *Experimental Economics 15*(1), 58–70.

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

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

Benjamin, D. J., S. A. Brown, and J. M. Shapiro (2013). Who is "Behavioral"? Cognitive ability and anomalous preferences. *Journal of the European Economic Association 11*(6), 1231–1255.

Bettinger, E. and R. Slonim (2006). Using experimental economics to measure the effects of a natural educational experiment on altruism. *Journal of Public Economics 90*(8), 1625 – 1648.

Bettinger, E. and R. Slonim (2007). Patience among children. *Journal of Public Economics 91*(1), 343 – 363.

Blake, P. R., M. Piovesan, N. Montinari, F. Warneken, and F. Gino (2015). Prosocial norms in the classroom: The role of self-regulation in following norms of giving. *Journal of Economic Behavior & Organization 115*, 18 – 29.

Bolton, G. E., J. Brandts, and A. Ockenfels (2005). Fair Procedures: Evidence from Games Involving Lotteries. *Economic Journal 115*(506), 1054–1076.

Bolton, G. E. and A. Ockenfels (2000). ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review 90*(1), 166–193.

Booth, A. and P. Nolen (2012a). Choosing to compete: How different are girls and boys? *Journal of Economic Behavior & Organization 81*(2), 542 – 555.

Booth, A. L. and P. Nolen (2012b). Gender differences in risk behaviour: Does nurture matter? *The Economic Journal 122*(558), F56–F78.

Borghans, L., J. J. Heckman, B. H. H. Golsteyn, and H. Meijers (2009). Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association 7*(2-3), 649–658.

Brandts, J. and D. J. Cooper (2006). A Change Would Do You Good - An Experimental Study on How to Overcome Coordination Failure in Organizations. *American Economic Review 96*(3), 669–693.

Brandts, J., J. Cooper D., E. Fatas, and S. Qi (2016). Stand by me – Experiments on help and commitment in coordination games. *Management Science 62*.

Breyer, B. and D. Danner (2015). Grit Scale for Perseverance and Passion for Long-Term Goals. *Zusammenstellung sozialwissenschaftlicher Items und Skalen*.

Brocas, I. and J. D. Carrillo (2018a). The determinants of strategic thinking in preschool children. *PLoS ONE 13*(5).

Brocas, I. and J. D. Carrillo (2018b). Iterative dominance in young children: Experimental evidence in simple two-person games. *Journal of Economic Behavior & Organization, forthcoming*.

Brocas, I., J. D. Carrillo, and N. Kodaverdian (2017). Altruism and strategic giving in children and adolescents. Technical report, CEPR Discussion Paper No. DP12288.

Brosig-Koch, J., T. Heinrich, and C. Helbach (2015). Exploring the capability to reason backwards: An experimental study with children, adolescents, and young adults. *European Economic Review 74*, 286 – 302.

Bucciol, A., D. Houser, and M. Piovesan (2011). Temptation and productivity: A field experiment with children. *Journal of Economic Behavior & Organization 78*(1), 126 – 136.

Bucciol, A. and M. Piovesan (2011). Luck or cheating? A field experiment on honesty with children. *Journal of Economic Psychology 32*(1), 73 – 78.

Buhrmester, M., T. Kwang, and S. D. Gosling (2011). Amazon's Mechanical Turk. *Perspectives on Psychological Science 6*(1), 3–5.

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

Buser, T., M. Niederle, and H. Oosterbeek (2014). Gender, competitiveness, and career choices . *The Quarterly Journal of Economics 129*(3), 1409–1447.

Camerer, C. F. (2003). Behavioral game theory: Plausible formal models that predict accurately. *Behavioral and Brain Sciences 26*(2), 157–158.

Cappelen, A., S. Fest, E. O. Sørensen, and B. Tungodden (2016). Choice and personal responsibility: What is a morally relevant choice? Discussion Paper 27/2014, NHH Dept. of Economics.

Cappelen, A. W., C. Cappelen, and B. Tungodden (2017). False Positives and False Negatives in Distributive Choices. Discussion paper, Mimeo.

Cappelen, A. W., A. D. Hole, E. O. Sørensen, and B. Tungodden (2007). The Pluralism of Fairness Ideals: An Experimental Approach. *American Economic Review 97*(3), 818–827.

Cappelen, A. W., J. Konow, E. O. Sørensen, and B. Tungodden (2013). Just Luck: An Experimental Study of Risk-Taking and Fairness. *American Economic Review 103*(4), 1398–1413.

Cappelen, A. W., J. A. List, A. Samek, and B. Tungodden (2016). The Effect of Early Education on Social Preferences. Working Paper 22898, National Bureau of Economic Research.

Cárdenas, J.-C., A. Dreber, E. von Essen, and E. Ranehill (2012). Gender differences in competitiveness and risk taking: Comparing children in Colombia and Sweden. *Journal of Economic Behavior & Organization 83*(1), 11 – 23. Gender Differences in Risk Aversion and Competition.

Cárdenas, J.-C., A. Dreber, E. von Essen, and E. Ranehill (2014). Gender and cooperation in children: Experiments in Colombia and Sweden. *PLOS ONE 9*(3), 1–7.

Cason, T. N., A. C. Savikhin, and R. M. Sheremeta (2012). Behavioral spillovers in coordination games. *European Economic Review 56*(2), 233 – 245.

Cason, T. N., R. M. Sheremeta, and J. Zhang (2012). Communication and efficiency in competitive coordination games. *Games and Economic Behavior 76*(1), 26 – 43.

Cassar, L. and A. H. Klein (2016). A Matter of Perspective: How Experience Shapes Preferences for Redistribution. Discussion paper, Mimeo.

Castelli, I., D. Massaro, A. G. Sanfey, and A. Marchetti (2010). Fairness and intentionality in children's decision-making. *International Review of Economics 57*(3), 269–288.

Castillo, M. (2017). Negative childhood experiences and risk aversion: evidence from children exposed to domestic violence. Technical report, ICES.

Castillo, M., P. J. Ferraro, J. L. Jordan, and R. Petrie (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.

Castillo, M., J. L. Jordan, and R. Petrie (2018a). Children's rationality, risk attitudes and field behavior. *European Economic Review 102*, 62 – 81.

Castillo, M., J. L. Jordan, and R. Petrie (2018b). Discount rates of children and high school graduation. *The Economic Journal*.

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

Charness, G. and M. Rabin (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics 117*(3), 817–869.

Chen, J., D. Houser, N. Montinari, and M. Piovesan (2016). Beware of popular kids bearing gifts: A framed field experiment. *Journal of Economic Behavior & Organization 132*, 104 – 120.

Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner (2014). Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility. *American Economic Review 104*(5), 141–47.

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

Conrads, J., B. Irlenbusch, R. Rilke, and G. Walkowitz (2013). Lying and Team Incentives. *Journal of Economic Psychology 34*(C), 1–7.

Conte, A. and J. D. Hey (2013). Assessing Multiple Prior Models of Behaviour under Ambiguity. *Journal of Risk and Uncertainty 46*(2), 113–132.

Cooper, R. W., D. V. DeJong, R. Forsythe, and T. W. Ross (1990). Selection Criteria in Coordination Games: Some Experimental Results. *The American Economic Review 80*(1), 218–233.

Corak, M. (2013). Income Inequality, Equality of Opportunity, and Intergenerational Mobility. *Journal of Economic Perspectives 27*(3), 79–102.

Croson, R. and U. Gneezy (2009). Gender Differences in Preferences. *Journal of Economic Literature 47*(2), 448–74.

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

Deckers, T., A. Falk, F. Kosse, P. R. Pinger, and H. Schildberg-Hörisch (2017). Socio-economic status and inequalities in children's IQ and economic preferences. Technical report, IZA Discussion Paper No. 11158.

Deckers, T., A. Falk, F. Kosse, and H. Schildberg-Hörisch (2015). How does socio-economic status shape a child's personality? Technical report, IZA Discussion Paper.

Devetag, G. and A. Ortmann (2007). When and why? A critical survey on coordination failure in the laboratory. *Experimental Economics 10*(3), 331–344.

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

Dreber, A., E. von Essen, and E. Ranehill (2014). Gender and competition in adolescence: task matters. *Experimental Economics 17*(1), 154–172.

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

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

Duffy, J. and N. Feltovich (2006). Words, Deeds, and Lies: Strategic Behaviour in Games with Multiple Signals. *The Review of Economic Studies 73*(3), 669–688.

Eckel, C., P. J. Grossman, C. A. Johnson, A. C. M. de Oliveira, C. Rojas, and R. Wilson (2011). Social norms of sharing in high school: Teen giving in the dictator game. *Journal of Economic Behavior & Organization 80*(3), 603 – 612.

Eckel, C. C. and P. J. Grossman (2008). Forecasting Risk Attitudes: An Experimental Study using Actual and Forcasted Gamble Choices. *Journal of Economic Behavior & Organization 68*(1), 1–17.

Eckel, C. C., P. J. Grossman, C. A. Johnson, A. C. M. de Oliveira, C. Rojas, and R. K. Wilson (2012). School environment and risk preferences: Experimental evidence. *Journal of Risk and Uncertainty 45*(3), 265–292.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics 75*(4), 643–669.

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

Engelmann, D. and M. Strobel (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American Economic Review 94*(4), 857–869.

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

Falk, A. and J. J. Heckman (2009). Lab Experiments Are a Major Source of Knowledge in the Social Sciences. *Science 326*(5952), 535–538.

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

Fan, C.-P. (2000). Teaching children cooperation – An application of experimental game theory. *Journal of Economic Behavior & Organization 41*(3), 191 – 209.

Fehr, D. (2018). Is increasing inequality harmful? experimental evidence. *Games and Economic Behavior 107*, 123 – 134.

Fehr, E., H. Bernhard, and B. Rockenbach (2008). Egalitarianism in young children. *Nature 454*, 1079–1083.

Fehr, E., D. Glätzle-Rützler, and M. Sutter (2013). The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence. *European Economic Review 64*, 369 – 383.

Fehr, E. and K. M. Schmidt (1999a). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics 114*(3), 817–868.

Fehr, E. and K. M. Schmidt (1999b). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics 114*(3), pp. 817–868.

Felfe, C., M. Kocher, H. Rainer, J. Saurer, and T. Siedler (2018). More Opportunity, More Cooperation? The Behavioral Effects of Birthright Citizenship on Immigrant Youth. Economics Series 340, Institute for Advanced Studies.

Feri, F., B. Irlenbusch, and M. Sutter (2010). Efficiency gains from team-based coordination – Large-scale experimental evidence. *American Economic Review 100*(4), 1892–1912.

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

Fischbacher, U. and F. Föllmi-Heusi (2013). Lies in Disguise: An Experimental Study on Cheating. *Journal of the European Economic Association 11*(3), 525–547.

Gächter, S. and J. F. Schulz (2016). Intrinsic Honesty and the Prevalence of Rule Violations across Societies. *Nature 531*(7595), 496–499.

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

Glaeser, E., J. Scheinkman, and A. Shleifer (2003). The Injustice of Inequality. *Journal of Monetary Economics 50*(1), 199–222.

Glätzle-Rützler, D. and P. Lergetporer (2015). Lying and age: An experimental study. *Journal of Economic Psychology 46*, 12 – 25.

Glätzle-Rützler, D., M. Sutter, and A. Zeileis (2015). No myopic loss aversion in adolescents? – An experimental note. *Journal of Economic Behavior & Organization 111*, 169 – 176.

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

Gneezy, U., A. Kajackaite, and J. Sobel (2018). Lying aversion and the size of the lie. *American Economic Review 108*(2), 419–53.

Gneezy, U. and A. Rustichini (2004). Gender and Competition at a Young Age. *American Economic Review 94*(2), 377–381.

Golsteyn, B., H. Grönqvist, and L. Lindahl (2014). 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.

Grolleau, G., M. G. Kocher, and A. Sutan (2016). Cheating and Loss Aversion: Do People Cheat More to Avoid a Loss? *Management Science 62*(12), 3428–3438.

Grueneisen, S., E. Wyman, and M. Tomasello (2015a). Children use salience to solve coordination problems. *Developmental Science 18*(3), 495–501.

Grueneisen, S., E. Wyman, and M. Tomasello (2015b). "I know you don't know I know" Children use second-order false-belief reasoning for peer coordination. *Child Development 86*(1), 287–293.

Gummerum, M., Y. Hanoch, M. Keller, K. Parsons, and A. Hummel (2010). Preschoolers' allocations in the dictator game: The role of moral emotions. *Journal of Economic Psychology 31*(1), 25 – 34.

Güth, W. and M. G. Kocher (2014). More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. *Journal of Economic Behavior & Organization 108*, 396 – 409.

Harbaugh, W., K. Krause, and S. Liday (2003a). Bargaining by children. Working Paper 2002-4, University of Oregon Economics.

Harbaugh, W., K. Krause, and S. J. Liday (2003b). Trust in children. In *Trust and Reciprocity: Interdisciplinary Lessons for Experimental Research.*

Harbaugh, W. T. and K. Krause (2000). Children's altruism in public good and dictator experiments. *Economic Inquiry 38*(1), 95–109.

Harbaugh, W. T., K. Krause, and T. R. Berry (2001). GARP for kids: On the development of rational choice behavior. *The American Economic Review 91*(5), 1539–1545.

Harbaugh, W. T., K. Krause, and L. Vesterlund (2002). Risk attitudes of children and adults: Choices over small and large probability gains and losses. *Experimental Economics 5*(1), 53–84.

Harbaugh, W. T., K. Krause, and L. Vesterlund (2007). Learning to bargain. *Journal of Economic Psychology 28*(1), 127 – 142.

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

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science 312*(5782), 1900–1902.

Heckman, J. J. and Y. Rubinstein (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review 91*(2), 145–149.

Heckman, J. J. and J. A. Smith (1995). Assessing the case for social experiments. *Journal of Economic Perspectives 9*(2), 85–110.

Heckman, J. J., J. Stixrud, and S. Urzua (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics 24*(3), 411–482.

Hermes, H., F. Hett, M. Mechtel, F. Schmidt, D. Schunk, and V. Wagner (2018). Do children cooperate conditionally? Adapting the strategy method for first-graders. Technical report, SSRN.

Horton, J. J., D. G. Rand, and R. J. Zeckhauser (2011). The Online Laboratory: Conducting Experiments in a Real Labor Market. *Experimental Economics 14*(3), 399–425.

Houser, D., J. A. List, M. Piovesan, A. Samek, and J. Winter (2016). Dishonesty: From parents to children. *European Economic Review 82*, 242 – 254.

Houser, D., N. Montinari, and M. Piovesan (2012). Private and public decisions in social dilemmas: Evidence from children's behavior. *PLOS ONE 7*(8), 1–6.

Houser, D. and D. Schunk (2009). Social environments with competitive pressure: Gender effects in the decisions of German schoolchildren. *Journal of Economic Psychology 30*(4), 634 – 641.

John, K. and S. L. Thomsen (2015). School-track environment or endowment: What determines different other-regarding behavior across peer groups? *Games and Economic Behavior 94*, 122 – 141.

Kajackaite, A. (2016). Lying About Luck versus Lying About Performance. Discussion paper, Available at SSRN: https://ssrn.com/abstract=2851149.

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

Khachatryan, K., A. Dreber, E. von Essen, and E. Ranehill (2015). Gender and preferences at a young age: Evidence from Armenia. *Journal of Economic Behavior & Organization 118*, 318 – 332. Economic Experiments in Developing Countries.

Khadjavi, M. and A. Nicklisch (2018). Parent's ambitions and children's competitiveness. *Journal of Economic Psychology 67*, 87 – 102.

Klimm, F. (2018). Suspicious success - cheating, inequality acceptance and political preferences. Technical report, LMU.

Konow, J. (2000). Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions. *American Economic Review 90*(4), 1072–1091.

Konow, J. (2009). Is Fairness in the Eye of the Beholder? An Impartial Spectator Analysis of Justice. *Social Choice and Welfare 33*(1), 101–127.

Konow, J., T. Saijo, and K. Akai (2016). Equity versus Equality. Discussion paper.

Kosse, F., T. Deckers, H. Schildberg-Hörisch, and A. Falk (2018). The formation of prosociality: Causal evidence on the role of social environment. *Journal of Political Economy, forthcoming*.

Kosse, F. and F. Pfeiffer (2012). Impatience among preschool children and their mothers. *Economics Letters 115*(3), 493 – 495.

Kuziemko, I., M. I. Norton, E. Saez, and S. Stantcheva (2015). How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments. *American Economic Review 105*(4), 1478–1508.

Larbi, A. and F.-R. Christian (2016). Know when to Fold 'em: The Grit Factor. Working Papers 899, Barcelona Graduate School of Economics.

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

Lefgren, L. J., D. P. Sims, and O. B. Stoddard (2016). Effort, Luck, and Voting for Redistribution. *Journal of Public Economics 143*, 89–97.

Leiner, D. J. (2014). SoSci Survey (Version 2.5.00-i) [Computer software].

Lemmon, K. and C. Moore (2007). The development of prudence in the face of varying future rewards. *Developmental Science 10*(4), 502–511.

Lergetporer, P., S. Angerer, D. Glätzle-Rützler, and M. Sutter (2014). Third-party punishment increases cooperation in children through (misaligned) expectations and conditional cooperation. *Proceedings of the National Academy of Sciences 111*(19), 6916–6921.

Levin, I. P. and S. S. Hart (2003). Risk preferences in young children: Early evidence of individual differences in reaction to potential gains and losses. *Journal of Behavioral Decision Making 16*(5), 397–413.

Levitt, S. D. and J. A. List (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives 21*(2), 153–174.

Levitt, S. D. and J. A. List (2009). Field experiments in economics: The past, the present, and the future. *European Economic Review 53*(1), 1 – 18.

List, J. A., R. Petrie, and A. Samek (2018). How experiments with children can inform economics. Technical report, Work in progress.

List, J. A. and A. C. Samek (2013). Exploring the origins of charitable acts: Evidence from an artefactual field experiment with young children. *Economics Letters 118*(3), 431 – 434.

Litman, L., J. Robinson, and T. Abberbock (2016). TurkPrime.com: A Versatile Crowdsourcing data Acquisition Platform for the Behavioral Sciences. *Behavior Research Methods*, 1–10.

Litman, L., J. Robinson, and C. Rosenzweig (2015). The Relationship between Motivation, Monetary Compensation, and Data Quality among US- and India-based Workers on Mechanical Turk. *Behavior Research Methods 47*(2), 519–528.

Lührmann, M., M. Serra-Garcia, and J. Winter (2018). The impact of financial education on adolescents' intertemporal choices. *American Economic Journal: Economic Policy 10*(3), 309–32.

Maggian, V. and M. C. Villeval (2016). Social preferences and lying aversion in children. *Experimental Economics 19*(3), 663–685.

Martinsson, P., K. Nordblom, D. Rüzler, and M. Sutter (2011). Social preferences during childhood and the role of gender and age – An experiment in Austria and Sweden. *Economics Letters 110*(3), 248–251.

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

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

Möllerström, J., B.-A. Reme, and E. O. Sørensen (2015). Luck, Choice and Responsibility – An Experimental Study of Fairness Views. *Journal of Public Economics 131*, 33–40.

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

Moreira, B., R. Matsushita, and S. D. Silva (2010). Risk seeking behavior of preschool children in a gambling task. *Journal of Economic Psychology 31*(5), 794–801.

Munro, A. and Y. Tanaka (2014). Risky rotten kids: An experiment on risk attitudes amongst adolescents in rural Uganda. GRIPS Discussion Papers 14-01, National Graduate Institute for Policy Studies.

Murnighan, J. K. and M. S. Saxon (1998). Ultimatum bargaining by children and adults. *Journal of Economic Psychology 19*(4), 415 – 445.

Niederle, M. and L. Vesterlund (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics 122*(3), 1067–1101.

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

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

Ochs, J. (1995). Coordination problems. *Handbook of Experimental Economics*.

Peters, H. E., A. S. Ünür, J. Clark, and W. D. Schulze (2004). Free-riding and the provision of public goods in the family: A laboratory experiment. *International Economic Review 45*(1), 283–299.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Harvard University Press.

Rey-Biel, P., R. M. Sheremeta, and N. Uler (2011). (Bad) Luck or (Lack of) Effort?: Comparing Social Sharing Norms between US and Europe. Discussion paper, Availalable at http://pareto.uab.es/prey/Neslihansub.pdf.

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

Samek, A. C. (2013). Is there a gender gap in preschoolers' competitiveness? An experiment in the U.S. *Journal of Economic Behavior & Organization 92*, 22 – 31.

Schmidt, D., R. Shupp, J. M. Walker, and E. Ostrom (2003). Playing safe in coordination games: The roles of risk dominance, payoff dominance, and history of play. *Games and Economic Behavior 42*(2), 281–299.

Schunk, D., E. Berger, H. Mueller, K. Winkel, and E. Fehr (2017). The KIDS-WIN-Study: Design of a working memory training intervention in primary schools. Technical report.

Segal, C. (2013). Misbehavior, education, and labor market outcomes. *Journal of the European Economic Association 11*(4), 743–779.

Shepard, E. I., T. J. Clifton, and D. Kruse (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. (2007). Outcomes versus intentions: On the nature of fair behavior and its development with age. *Journal of Economic Psychology 28*(1), 69 – 78.

Sutter, M., S. Czermak, and F. Feri (2013). Strategic sophistication of individuals and teams. Experimental evidence. *European Economic Review 64*, 395 – 410.

Sutter, M., F. Feri, D. Glätzle-Rützler, M. G. Kocher, P. Martinsson, and K. Nordblom (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. and D. Glätzle-Rützler (2015). Gender differences in the willingness to compete emerge early in life and persist. *Management Science 61*(10), 2339–2354.

Sutter, M., D. Glätzle-Rützler, L. Balafoutas, and S. Czermak (2016). Cancelling out early age gender differences in competition: an analysis of policy interventions. *Experimental Economics 19*(2), 412–432.

Sutter, M. and M. G. Kocher (2007). Trust and trustworthiness across different age groups. *Games and Economic Behavior 59*(2), 364 – 382.

Sutter, M., M. G. Kocher, D. Glätzle-Rützler, and S. T. Trautmann (2013). Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. *American Economic Review 103*(1), 510–31.

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

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

Tinghög, G., D. Andersson, and D. Västfjäll (2017). Are Individuals Luck Egalitarians? An Experiment on the Influence of Brute and Option Luck on Social Preferences. *Frontiers in Psychology 8*(460).

Train, K. (2009). *Discrete Choice Methods with Simulation.* Cambridge University Press.

Van Huyck J., B., C. Battalio R., and O. Beil R. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review 80*.

Warneken, F. (2018). How children solve the two challenges of cooperation. *Annual Review of Psychology 69*(1), 205–229.

Weber, R. A. (2006). Managing Growth to Achieve Efficient Coordination in Large Groups. *American Economic Review 96*(1), 114–126.

Weisel, O. and S. Shalvi (2015). The Collaborative Roots of Corruption. *Proceedings of the National Academy of Sciences 112*(34), 10651–10656.

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

Zhang, J. (2011). Do girls in China compete just as much as boys? Evidence from an experiment that predicts educational choice. Technical report, Citeseer.

Erklärung

nach §6 der Promotionsordnung vom 16. Februar 2008

Ich erkläre hiermit, dass ich die vorgelegte Arbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Aussagen, Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise unentgeltlich geholfen:

Weitere Personen neben den in der Einleitung der Arbeit aufgeführten Koautorinnen und Koautoren waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem lnhalt der vorgelegten Dissertation stehen. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt. Ich versichere, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe

Unterschrift:

CURRICULUM VITAE

PERSONAL DETAILS

Name	Claudia Elisabeth Zoller
Date of birth	January 23, 1990
Place of birth	Rum, Austria

EDUCATION

2017 – 2018	Research fellow, Max-Planck-Institute for Research on Collective Goods, Bonn
2015 – 2018	Ph.D. student, University of Cologne
2015 – 2017	Research fellow, University of Cologne
2014	Diploma in International Business Administration and Economics, University of Innsbruck
2008	Matura, Adolf-Pichler-Platz Gymnasium, Innsbruck

Cologne, November 19, 2018