



HCEO WORKING PAPER SERIES

Working Paper



HUMAN CAPITAL AND
ECONOMIC OPPORTUNITY
GLOBAL WORKING GROUP

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

www.hceconomics.org

The Development of Egalitarian Norm Enforcement in Childhood and Adolescence

Zvonimir Bašić^{1,2}, Armin Falk^{2,3}, and Fabian Kosse^{2,4}

¹Max Planck Institute for Research on Collective Goods, Bonn

²Institute on Behavior and Inequality (briq)

³University of Bonn

⁴University of Munich

March 2019

Abstract

The equal division of goods is a long-existing social norm present in societies around the world. In order to ensure that the egalitarian norm is followed, people engage in costly enforcement of norm-violating behavior. Despite its importance, little is known about the emergence of this enforcement and how it develops over time. Therefore, we take the most commonly-used third-party punishment game where a third party is added to a dictator game, adapt it for children and run an experiment with 9-18 year-old children and adolescents. We show that already at 9-10 years of age, a small but non-negligible proportion of subjects are costly enforcing the egalitarian norm. We find that this behavior then strongly develops in the following years: The proportion of egalitarian norm enforcers increases, becoming the most common behavioral type with 11-12 years of age, and the punishers' behavior fully develops until 13-14 years of age. Following those developmental changes, the enforcing behavior remains stable until adulthood. We find that some norm enforcers do not only punish selfish, but also generous deviations from the egalitarian norm. Looking at the dictators' behavior, we observe that they increase their transfer in the direction of the egalitarian norm primarily in the same period as we observe developmental changes on the punishers' side.

JEL Classification: C91, C93, D63, D90

Keywords: third-party punishment, norm enforcement, egalitarian norm, formation of preferences, children

Corresponding author: Zvonimir Bašić, Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Straße 10, 53113 Bonn, Germany, e-mail addresses: basic@coll.mpg.de

1 Introduction

Egalitarian norm is a long-existing and widely-pervasive organizing principle. It shapes our everyday behavior, e.g., when we split the bill in a restaurant or divide the profits from a business venture, and it functions as a crucial driver behind prosocial behavior (see Camerer, 2003; Engel, 2011), determining the way in which it is frequently modeled (Fehr and Schmidt, 1999; Falk and Fischbacher, 2006; Andreoni and Bernheim, 2009). To ensure that people act according to the norm, societies rely on enforcement mechanisms. In particular, if a person violates the egalitarian norm, unaffected third-parties commonly punish such behavior (e.g. Fehr and Fischbacher, 2003, 2004). The extent of punishment might vary, but the punishment itself is robust and widespread, observed in societies around the globe (Henrich et al., 2006; Bernhard et al., 2006). However, while many studies focus on egalitarian norm enforcement among adults, much less is known about its emergence within human development.

Despite a rich body of literature investigating the development of children's other-regarding preferences (e.g. Fehr et al., 2008, 2013; Almås et al., 2010; Falk and Kosse, 2016; Deckers et al., 2017; Brocas et al., 2017; Sutter et al., 2018), less focus is devoted to norm-enforcing behavior. Given its relevance, it holds strong interest to shed more light on the roots of egalitarian norm enforcement. In particular, as children become less selfish and more willing to share as they get older (Harbaugh et al., 2003; Benenson et al., 2007; Fehr et al., 2008, 2013; Bauer et al., 2014; Kosse et al., forthcoming), it is essential to understand how enforcement, which underpins such behavior, develops. Therefore, in this paper we investigate the emergence and development of egalitarian norm enforcement. More specifically, we take the most commonly-used third-party punishment game where a third-party is added to a regular dictator game (Fehr and Fischbacher, 2003, 2004; Henrich et al., 2006; Bernhard et al., 2006; Lewisch et al., 2011; Kriss et al., 2016), adapt it for children's understanding and run an experiment with 9-18 year-old children and adolescents.

We find that already at 9-10 years of age, a small but non-negligible proportion of subjects are costly enforcing the egalitarian norm. We show that this behavior then markedly develops in the following years: i) the proportion of norm enforcers increases as norm enforcers become the most prevalent behavioral type with 11-12 years of age, and in most of the comparisons across the older age cohorts and across robustness checks, the enforcers represent more than 50% of the subjects, and ii) the behavior of the punishers fully develops until 13-14 years of age. Following these developmental changes, the enforcing behavior remains stable from 13-14 years of age until adulthood. Furthermore, we find that some of the norm enforcers do not only punish selfish deviations from the norm, but are also willing to punish generous deviations. Finally, we show that the

dictators increase their giving in the direction of the egalitarian norm primarily in the same period as we observe development on the punishers' side.

Our study contributes to several streams of literature. Primarily, it contributes to the literature on the development of other-regarding behavior, as we show that already at 9-10 years of age, some subjects are willing to costly punish egalitarian norm deviations, and that the enforcing behavior increases in the following years, fully developing until 13-14 years of age. Moreover, we replicate the most common finding in the mentioned literature, observing that children's selfish behavior decreases with age as older dictators give more on average than the younger ones (Harbaugh et al., 2003; Benenson et al., 2007; Fehr et al., 2008, 2013; Bauer et al., 2014; Kosse et al., forthcoming). Interestingly, we find that the decrease of selfishness primarily occurs in the same period when we observe development of behavior on the punishers' side, suggesting a potential connection between the two. This provides supporting evidence that punishment is one of the mechanisms contributing to the commonly-observed increase of prosocial behavior over age.

Most closely to our paper, our results connect to other third-party punishment studies with young children. Lergetporer et al. (2014) find that 7-11 year-old children do not enforce the norm of cooperation, while McAuliffe et al. (2015) and Jordan et al. (2014) find that 6-8 year-old children punish deviations from unfair allocations. Our findings are more in line with the latter, as our youngest cohort also exhibits norm-enforcing behavior; however, in contrast to McAuliffe et al. (2015) and Jordan et al. (2014), we find that the largest proportion of youngest participants still punish at the equal share, and do not enforce the norm. The difference between our findings could be explained by differences in design and through spiteful preferences or preferences towards advantageous inequality, which can still be highly present in this development period (Fehr et al., 2013; McAuliffe et al., 2014; see section 4 for discussion). Complementing the aforementioned studies with young children, our study connects the behavior from childhood to adulthood, as we record punishment from 9 to 18 years of age. We find that punishment behavior mainly develops by mid-adolescence. This supports the major contribution of Almås et al. (2010), who observe that children's understanding of fairness still strongly evolves in the mentioned period. Furthermore, it is in line with the findings of Gummerum and Chu (2014), who study the development of concerns for outcomes and intentions in second and third parties. While the authors do not observe a change in third-parties' concerns for outcomes or intentions in this period, they report an increase in punishment behavior when moving from pre-adolescence to mid-adolescence. Looking at our results with older adolescents, we observe that the punishment behavior does not change from mid-adolescence to adulthood, where our study connects to other third-party punishment studies with adults (Fehr and Fischbacher, 2003, 2004; Henrich et al.,

2006; Bernhard et al., 2006; Lewisch et al., 2011; Kriss et al., 2016). More specifically, the behavior of the oldest cohorts in our study is comparable to Bernhard et al. (2006), who similarly to our design, limit the punishers' action space to only three punishment options.¹

Finally, our results indicate that some subjects that do not punish on the egalitarian norm but enforce selfish transfers which fall short from it, also punish above-egalitarian transfers. The finding relates to other third-party punishment studies (Henrich et al., 2006; Gummerum and Chu, 2014) and complements the growing literature on antisocial punishment (Herrmann et al., 2008; Parks and Stone, 2010; Irwin and Horne, 2013). Importantly, it also connects to studies focusing on norm elicitation, where it has been shown that in contrast to the egalitarian transfer in a dictator game, a certain proportion of subjects do not approve of above-egalitarian transfers (Krupka and Weber, 2013; Kimbrough and Vostroknutov, 2016). Our findings suggest that this normative valuation stretches its roots to a much younger age.

The remainder of the paper is structured as follows. In section 2, we explain the experimental design and procedure. In section 3, we report the results of our study. In section 4, we discuss the potential mechanisms behind our findings, before we finally conclude the paper in section 5.

2 Experimental design and procedure

2.1 Experimental design

The experimental game we used comprises three players: a dictator, recipient and punisher (Figure 1). At the beginning of the game, the dictator is endowed with 10 tokens and has to divide them between himself and the recipient, who has no endowment. The punisher is endowed with 5 tokens. After the dictator makes his decision, the punisher can punish the dictator with up to 2 tokens. For each token that the punisher uses, the dictator's payoff decreases by 3 but is bounded by a minimum of 0.

In the game we use, norm enforcers are primarily seen as subjects who do not punish on the transfer of 5, but punish selfish deviations that fall short from the equal split (see Fehr and Fischbacher, 2004; Henrich et al., 2006; Bernhard et al., 2006). Hence, to investigate norm enforcement, we apply the same definition. However, this definition only encaptures those that disapprove of selfish behavior, and does not take into account that someone might punish overly-generous behavior that exceeds the norm.² Therefore,

¹While many studies implement the third-party punishment game where a punisher is added to a dictator game (Fehr and Fischbacher, 2003, 2004; Henrich et al., 2006; Lewisch et al., 2011; Kriss et al., 2016), we limit the punisher's action space to three possible actions as in Bernhard et al. (2006).

²Fehr and Fischbacher (2004) take into account the punishment of overly-generous behavior in their

in addition to norm enforcement of selfish deviations, we also investigate and report norm enforcement of generous deviations, where a norm enforcer is a subjects that does not punish on the transfer of 5, but punishes transfers that are larger than 5.

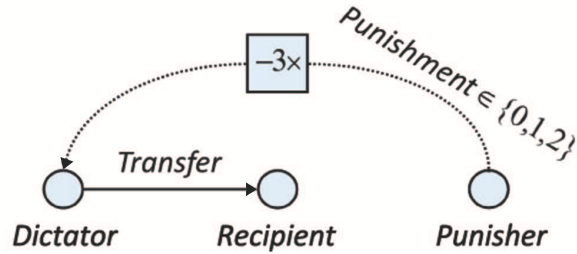


Figure 1: Dictator game with a third-party punisher. The dictator is endowed with 10 tokens, which he divides between himself and the recipient, who has no endowment. The punisher is endowed with 5 tokens. After the dictator’s decision, the punisher can punish the dictator with up to 2 of his tokens. For each token that the punisher uses, the dictator’s payoff decreases by 3 but is bounded by a minimum of 0.

To gain a comprehensive picture of the punishment behavior, we implemented the strategy method. Each punisher had to indicate how much he would punish conditional on every potential choice of the dictator.³

2.2 Procedure and subject pool

The study was conducted in Našice, a small city in eastern Croatia with an approximate population of 16,000. The participants were students in 3rd, 5th and 7th grade of elementary school and 1st and 3rd grade of high school. The age of the children was 9-10, 11-12, 13-14, 15-16 and 17-18, respectively.⁴ Altogether, 635 students from two elementary schools and one high school participated in the study. The experiment was conducted during regular school hours. Sessions were organized for each class and each student was assigned one role in the game, which was the same for everybody in a class. Our approach allowed us to minimize any selection bias when comparing

analysis. Moreover, the existence of above-egalitarian punishment can be seen in several other studies (Henrich et al., 2006; Bernhard et al., 2006; Gummerum and Chu, 2014).

³Note that the strategy method was shown not to influence results of third-party punishment (Jordan et al., 2015), and was successfully used with 7 year-old children (Lergetporer et al., 2014). We discuss the details of our strategy-method implementation in the subsection 2.2.

⁴Our sample contains few exceptions where a child is older or younger than his classmates, e.g., if a child fails a grade or starts school earlier. In our main comparison, we follow the usual approach in the literature and bin the children based on their age in their corresponding age cohorts (9-10, 11-12, 13-14, 15-16 and 17-18 years of age). Classifying the children strictly according to their grade does not change our results.

high school and elementary school subjects. In particular, the participating schools were the only schools in the local area. Moreover, at the time of the experiment, the participating high school had the largest number of students among high schools in Croatia, and it offered a variety of study programs to its students. This allowed us to balance the sample across the different types of programs within each high school age cohort (Table 1).⁵ Three participating classes of students were from three-year high school programs, and students from different classes were assigned a different role in the game. Three-year programs are considered the least difficult to finish and have the lowest enrollment requirements. Furthermore, three participating classes of students were from four-year non-gymnasium programs and three were from four-year gymnasium programs. The latter are considered the most difficult programs and have the highest enrollment requirements. For both gymnasium and non-gymnasium four-year programs, students from different classes were also assigned different roles in the game. For elementary schools, three classes were participating from each grade level and school, and children across these classes were assigned different roles in the game.⁶

Table 1: Sampling

Grade (age)	Elementary school 1	Elementary school 2	High school		
			3 year program	4 year non-gymn. program	4 year gymn. program
3rd (9-10)	3 classes	4 classes			
5th (11-12)	3 classes	3 classes			
7th (13-14)	3 classes	4 classes			
1st (15-16)			3 classes	3 classes	3 classes
3rd (17-18)			3 classes	3 classes	3 classes

During and after the experiment, the players remained anonymous to each other, although they were aware that the other players were students of similar age but not from their class. Each student was given a code that served as identification during the experimental procedure and no names were used. During the experiment, only the experimenters and the children were allowed to be in the classroom. The environment during the experiment resembled the environment during a normal school test. The tables where children were sitting were separated and children were not allowed to talk to each other (Fig. B1 in Online appendix).

⁵Unlike elementary school education in Croatia, high school education offers multiple study programs, which students choose based on their preference and prior academic achievement.

⁶In two grade levels in elementary school 2, we added an additional class to have a balanced sample size across the two schools. In these two cases, children from the extra class were assigned the same role in the game as in one of the other three classes.

Given the relatively complex mechanism of the third-party punishment game, we devoted special focus to its implementation, and to ensuring that children of all ages are able to understand the game. In particular, we put special focus on the design of the decision sheets and the instructions. To develop them, we worked together with an experienced child psychologist. Moreover, we tested the instructions and the decision sheets with a small group of 3rd-graders prior to the study. As a final product of this process, instead of a list, we designed decision sheets which contain 11 ordered pages. On each page, there was one potential transfer from the dictator with depicted punishment options and vividly-displayed outcomes for each of those options (Fig. B2 in Online appendix). This ensured that subjects of all ages and different cognitive abilities were fully aware of the potential payoff outcomes conditional on their choices. Furthermore, the experimental instructions comprised oral instructions, which were supported with posters on the blackboard that were incorporated in the structured explanation of the game (for instructions, see Online appendix).⁷ To ensure standardization, alongside the main explanation of the game, the posters also contained several examples that were thoroughly explained for each session. Moreover, several aforementioned design choices were made explicitly to minimize issues with comprehension: We limited the punishment decisions to a maximum of 3 decisions as in Bernhard et al. (2006), and decided not to run the study with grades lower than grade 3. Finally, to be able to identify those that had difficulties with understanding the game, each participant had to answer three control questions, each comprising three sub-questions (items). If a child had problems solving any of the control questions, an experimenter explained the rules of the game again just to that child, and assisted in answering the first control question through explanation. The second and third questions were used as a control of understanding (for control questions, see Online appendix). Out of 217 punishers, 180 answered the control questions correctly and have completed their decision sheets, while out of 209 dictators, 193 managed the same. The rest were excluded from the analysis.⁸

As the number of participating children across classrooms did not perfectly match, 8 dictators were each matched with 2 punishers, and 2 recipients were each matched with 2 dictators. After the experimental game, subjects answered a short questionnaire

⁷Oral instructions were read by the same experimenter in all experimental sessions.

⁸We report robustness analyses where we include the entire sample, and also, where we use inverse-probability-weighted least-square estimates to correct for minor imbalances across age cohorts and across subjects that have successfully answered the control questions or have failed to do so, in the Online appendix. See section B.1 in the Online appendix for a detailed description, and the robustness analyses in Tables B2, B3, B4, B5, B6 and B7. Our findings remain robust in these additional analyses. Also note that we fully excluded 2 subjects from our data due to their mental disabilities, and have also fully excluded subjects that gave an incompletely filled decision sheet, i.e., did not indicate a punishment choice on each dictator's transfer. Only 1 such subject successfully answered the control questions, while the remaining 4 have failed. The two children with mental disabilities were unable to finish the task, and were paid the mean earning of their age cohort.

and took a cognitive abilities test comprising 16 matrices items. The experiment was approved by the school principals of all three schools. Parents of participating children gave written consent for participation after they were informed about the nature and possible consequences of the study. With the consent, parents also filled out a small questionnaire with demographic and socio-economic questions. The tokens that students earned in the game were converted to Croatian kunas for payment and were paid out within 2 weeks after the experiment. The value of a token was adjusted to the age cohort, whereby students from 3rd grade of elementary school earned 3 kunas for a token, students from 5th and 7th grade of elementary school earned 4 kunas for a token, and students from 1st and 3rd grade of high school earned 5 kunas for a token, where $1 \text{ HRK} \approx 0.17 \text{ USD}$ at the time of the experiment. In addition to the earnings in the game, each subject also received a participation fee equivalent to the value of 2 tokens.

3 Results

3.1 Behavior of the punishers

We present the average punishment behavior for each dictator’s transfer in Figure 2. A visual inspection reveals that in all age cohorts, subjects are willing to punish. Moreover, the amount of punishment does not appear to differ much across the age cohorts, and the punishment exhibits a decreasing pattern as it declines with the dictator’s transfer. Finally, in all age cohorts, we observe a certain amount of punishment on the egalitarian norm, which seems to be most pronounced for the youngest cohort.

Before analyzing the punishment patterns, we first look at the overall amount of punishment, irrespective of the transfer level. We take the average punishment across all 11 potential transfers, and compare it across the age cohorts. As it was suggested in the visual inspection, we observe that the average punishment stays relatively stable across age cohorts, and it is always close to $2/3$ of a token (0.773, 0.614, 0.625, 0.723 and 0.680 for age cohorts 9-10, 11-12, 13-14, 15-16 and 17-18, respectively).

Next we turn to the analysis of the punishment patterns, and in particular, of norm enforcing behavior. To do so, we make a simple comparison of the punishment on the egalitarian norm, selfish deviations from the norm, and generous deviations from the norm as in Fehr and Fischbacher (2004). We regress the punishment amount on variables D_{neg} and D_{pos} , where D_{neg} is $\max(0, 5 - transfer)$, and represents the relation between punishment and selfish (negative) deviations from the norm, while D_{pos} is $\max(0, transfer - 5)$, and represents the relation between punishment and generous (positive) deviations from the norm. We show the results of the regression model in Table 2.

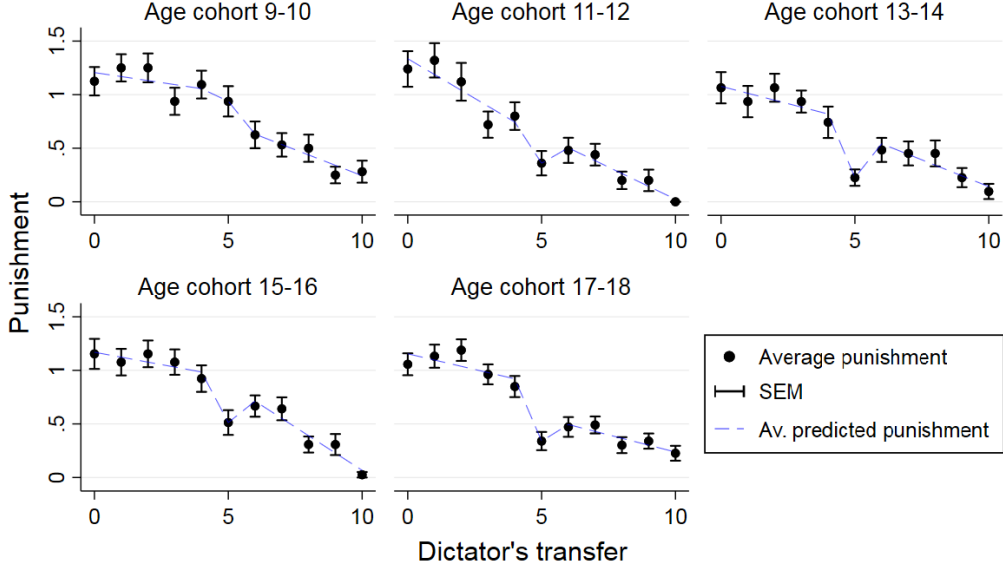


Figure 2: The punishment behavior across age cohorts. The black dots show the average observed punishment. Error bars show standard errors of the means. The blue dashed line shows the prediction of punishment based on our piecewise linear model (Table 3), which allows for a different increase in punishment on marginal deviations in comparison to further deviations from the egalitarian norm.

The table reveals that in all age cohorts, subjects significantly increase their punishment amount as the dictators selfishly deviate from the egalitarian norm (coefficient for D_{neg}), and significantly decrease their punishment amount as the dictators exceed the norm (coefficient for D_{pos}). Furthermore, the table also suggests that in all cohorts, there is a significant punishment on the egalitarian transfer (coefficient for Constant).⁹ Hence, while we observe that children increase their punishment when dictators give less than 5 tokens, which is in line with norm-enforcing behavior, we also observe non-negligible punishment on the egalitarian norm, and decreasing punishment on transfers that exceed it.¹⁰ However, the punishment on the egalitarian transfer, which is not consistent with egalitarian norm-enforcement, seems to decrease for older cohorts. Indeed, if we look at the actual average punishment on the egalitarian transfer in Figure 2, we observe that the 9-10 year-olds exhibit a punishment of 0.938 tokens, which then strongly and significantly decreases to 0.360 tokens with the 11-12 year-olds (two-sided t-test,

⁹See Tables B2 and B5 in the Online appendix for robustness checks. The reported results remain robust.

¹⁰In our paradigm, the low punishment on very high transfers is mechanically supported by the game design: As the dictator's transfer increases, the incentive to punish decreases since the dictator cannot end with less than 0 tokens. For example, if a dictator gives 8 tokens and leaves 2 for himself, it is sufficient to punish him with 1 token to ensure that he ends with 0 tokens. Hence, if there is punishment on the egalitarian transfer, this design aspect will mechanically support a negative coefficient for D_{pos} . Similarly to our findings, Fehr and Fischbacher (2004) also find a significant decrease in punishment when giving more than the egalitarian norm with adult subjects.

Table 2: OLS regressions estimates of the effect of egalitarian norm deviations on punishment

Dependent var:	Punishment				
Age cohort:	9-10	11-12	13-14	15-16	17-18
	(1)	(2)	(3)	(4)	(5)
D_{neg}	0.071** (0.028)	0.168*** (0.033)	0.126*** (0.041)	0.093*** (0.030)	0.136*** (0.022)
D_{pos}	-0.143*** (0.028)	-0.092*** (0.021)	-0.070** (0.027)	-0.132*** (0.021)	-0.070*** (0.019)
Constant	0.896*** (0.110)	0.521*** (0.085)	0.531*** (0.116)	0.767*** (0.083)	0.579*** (0.074)
Observations	352	275	341	429	583
R-squared	0.198	0.313	0.185	0.216	0.200
Adj. R-squared	0.193	0.307	0.180	0.212	0.198

Clustered standard errors at the individual level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

$p = 0.035$, $N(9-10) = 32$, $N(11-12) = 25$), and stays similar with older cohorts (0.225, 0.513, 0.340 tokens for cohorts 13-14, 15-16 and 17-18, respectively). A potential cause of this decrease might be a decline in the proportion of subjects punishing the egalitarian transfer, which could also allow for an increase in the proportion of subjects that are enforcing the norm. To investigate this conjecture, and get a better understanding of the punishment patterns we observe, we next turn to the analysis of punishment on the individual level.

For the analysis of punishment behavior on the individual level, we divide subjects into six categories: i) subjects that spend at least 1 token on punishing any of the selfish (negative) deviations, and at least 1 token on punishing any of the generous (positive) deviations from the norm and do not punish otherwise, ii) subjects that spend at least 1 token on punishing any of the selfish (negative) deviations from the norm and do not punish otherwise, iii) subjects that spend at least 1 token on punishing any of the generous (positive) deviations from the norm and do not punish otherwise, iv) subjects that do not punish on any of the possible transfers, and v) subjects that punish on the egalitarian norm. Notice that the first three categories focus on identifying *norm enforcers*, i.e., those that punish any deviation, selfish deviations, or generous deviations from the norm, respectively.¹¹ We present the results in Figure 3.

¹¹The categorization of norm enforcers can be interpreted in respect of the model used in Table 2. First category encaptures subjects for which the model predicts $D_{neg} > 0$ and $D_{pos} > 0$, second category encaptures subjects for which the model predicts $D_{neg} > 0$ and $D_{pos} = 0$, and the third category encaptures subjects for which the model predicts $D_{neg} = 0$ and $D_{pos} > 0$. We additionally impose that the punishment on the egalitarian norm is 0, as in these three categories we are interested

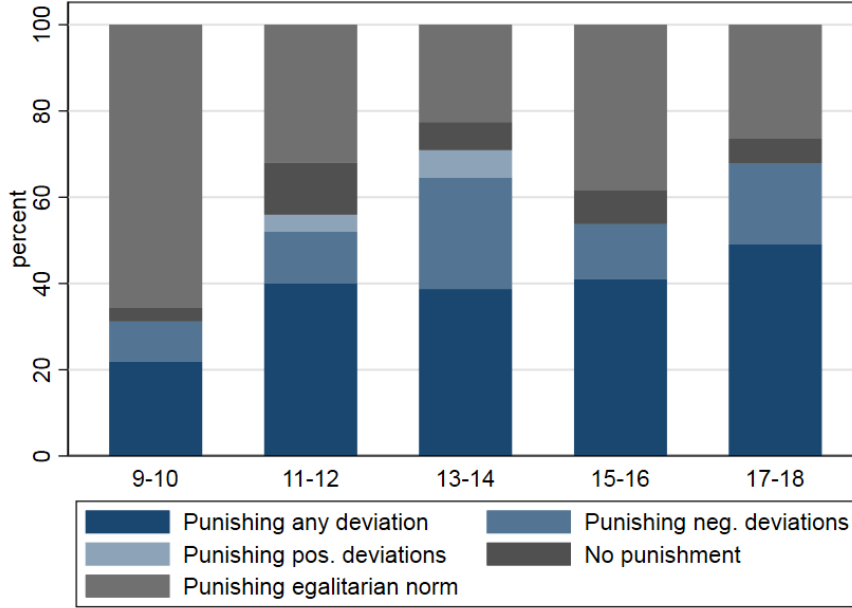


Figure 3: The distribution of punishment types across age cohorts. *Punishing any deviation*: Subjects that spend at least 1 token on punishing any of the selfish (negative) deviations, and at least 1 token on punishing any of the generous (positive) deviations from the norm and do not punish otherwise. *Punishing negative deviations*: Subjects that spend at least 1 token on punishing any of the selfish (negative) deviations from the norm and do not punish otherwise. *Punishing positive deviations*: Subjects that spend at least 1 token on punishing any of the generous (positive) deviations from the norm and do not punish otherwise. *No punishment*: Subjects that do not punish on any of the possible transfers. *Punishing egalitarian norm*: Subjects that punish on the egalitarian norm.

We observe that in the youngest cohort, only 31% of subjects enforce the egalitarian norm. The majority of those subjects are willing to punish both selfish and generous deviations from the norm (22%). Some are willing to punish only selfish deviations (9%), while no subject is willing to punish only generous deviations. We also observe that even 65% of 9-10 year-olds are willing to punish on the egalitarian norm. Moving to 11-12 year-old children, we observe a marked increase of norm enforcers, as more than 50% of subjects are willing to punish on some type of norm deviation. This change is primarily driven by the increase of subjects that are enforcing both selfish and generous deviations from the norm (40%), and less so by the increase of those that enforce only selfish deviations (12%). Furthermore, we also observe a very small proportion of subjects enforcing only generous deviations (4%). As more children start enforcing the egalitarian norm, less of them punish the normative behavior. In particular, 32% of 11-12 year-olds

in subjects that are punishing those that deviate from the norm, but are not punishing those that follow the norm.

punish the egalitarian norm, which is half less than with 9-10 year olds. Moving to older cohorts, we observe a similar pattern as with 11-12 year old subjects, with a slight increase of norm enforcers in the age cohorts 13-14 and 17-18, resulting in more than 60% of subjects willing to enforce the norm. Similarly to the 11-12 age cohort, the majority of norm enforcers in the older cohorts enforce both selfish and generous deviations from the norm. If we look at those willing to enforce only generous deviations, we observe them only in the 13-14 age cohort, and the proportion is again small (7%). Finally, if we focus on the subjects that do not punish on any of the 11 potential dictator's choices, we observe a small but existent proportion of such behavior across all five age cohorts (ranging from 3% to 12%).

A potential concern with the reported individual comparison is the definition of norm enforcers. In particular, one could argue that being willing to spend only 1 token to punish 1 out of 5 possible dictator's decisions (that deviate from the norm in selfish or generous direction) should not be considered as norm enforcement, as it might be insufficient to indicate willingness to enforce the norm. For this reason, we perform several robustness checks and repeat the analysis by defining norm enforcers in a stricter manner. Specifically, we redefine the first three categories as the following: i) subjects that spend a total of at least 2 tokens on punishing any of the selfish (negative) deviations, and a total of at least 2 tokens on punishing any of the generous (positive) deviations from the norm and do not punish otherwise, ii) subjects that spend a total of at least 2 tokens on punishing any of the selfish (negative) deviations from the norm and do not punish otherwise, and iii) subjects that spend a total of at least 2 tokens on punishing any of the generous (positive) deviations from the norm and do not punish otherwise. We show the distributions of different behavioral types in Figure A1 in the Appendix. Furthermore, we repeat the same analysis by redefining the three categories again in an even stricter manner, where a total of at least 3 tokens have to be spent on any generous or selfish deviations to indicate norm enforcement in that direction (Figure A2 in the Appendix), or that a total of at least 4 tokens have to be spent on any generous or selfish deviations to indicate norm enforcement in that direction (Figure A3 in the Appendix).

The robustness checks reveal two findings. First, the observed patterns stay robust. In all three robustness checks, we observe a similar age-related increase of norm enforcers as in Table 3. Moreover, the overall proportions of norm enforcers do not decrease by much, and for 11-12 age cohorts and older, the norm-enforcers are always the most common behavioral type, encapsulating more than 50% of subjects in majority of comparisons. Second, we observe that with the stricter definition of norm enforcers, the proportion of subjects that are willing to enforce both selfish and generous deviations declines but does not disappear, and the proportion of subjects that are willing to punish only selfish deviations increases. This leads to the conclusion that while the proportion of those

that enforce both generous and selfish deviations strongly depends on the strictness of our definition of norm enforcers, we robustly observe the existence of subjects that are willing to do so.¹²

Next, in Table A1, we apply a regression analysis to confirm the main observed pattern from the analysis of individual behavior: The age-related increase of norm enforcers when going from 9-10 year-olds to older age cohorts. We report 8 probit regression models. First two models investigate the age-related change of norm enforcers with our original definition of norm enforcers, while the latter 6 models include the definitions used for the robustness checks. For three out of four definitions of norm enforcers, we observe that all older age cohorts contain significantly more norm enforcers than the 9-10 age cohort (Models 1, 2, 3, 4, 7, 8). For one definition, we observe a significant increase in the proportion of norm enforcers between the 9-10 age cohort and 13-14 and older age cohorts, and a non-significant increase between the 9-10 and 11-12 age cohort (Models 5, 6). Hence, we confirm the age-related increase of the proportion of norm enforcers observed in the graphs, and also show that this increase is robust when controlling for individual differences in, e.g., gender and parental income.¹³

Altogether, we observe that a small but existent proportion of 9-10 year-olds are willing to enforce deviations from the norm. The proportion of norm enforcers increases in the following years, becoming the most common behavioral type with 11-12 year-olds, and it stays so for older cohorts. Moreover, in most of the comparisons across our robustness checks for cohorts 11-12 and older, the norm enforcers represent the majority of the subjects. Finally, we observe that (almost) all norm enforcers enforce selfish deviations from the norm, but some of them are also willing to enforce generous deviations.

As a final step of our analysis of punishment behavior, we employ an alternative analysis for investigating potential changes at the aggregate level across age cohorts. As reported in Table 2, we find that on the aggregate level, subjects on average increase their punishment to selfish deviations and decrease their punishment to generous deviations. Moreover, there is no clear development pattern of the degree of this increase or decrease across age, respectively. We complement this analysis by building an alternative piecewise linear model that takes into account the changes we observe in the punishment patterns. We observe that the main changes in the punishment behavior primarily happen in response to transfers of 4, 5 and 6 (see Figure 2), hence, we design a model

¹²As the strictness of the definition of norm enforcers increases, the subjects that enforce both generous and selfish deviations remain visible only in older age cohorts.

¹³See Tables B3 and B6 in the Online appendix for robustness checks. The age-related increase of norm-enforcers remains robust. Using inverse probability weighting (Table B3) yields highly similar results as in Table A1, while taking both subjects that failed and succeeded in answering the control questions in the analysis (Table B6) shows a slightly more gradual increase of norm-enforcers, as significant differences occur between the 9-10 age cohort and 13-14 and older age cohorts.

that contrasts the punishment decisions on deviations to 4 and 6 tokens with punishment decisions on larger deviations. To do so, we take the same regression model as in Table 2 as a base model, and we adjust it by adding two dummy variables: i) $Marginal_{neg}$, which estimates the change in the punishment amount when going from a transfer of 5 to 4, whereby it is 1 for $transfer < 5$ and 0 for $transfer \geq 5$, and ii) $Marginal_{pos}$, which estimates the change in the punishment amount when going from a transfer of 5 to 6, whereby it is 0 for $transfer < 6$ and 1 for $transfer \geq 6$. Note that while more motives could be driving these observed changes, we use these changes *only* to identify potential development in the aggregate punishment pattern across age cohorts.¹⁴

We report the results of the model in Table 3, and to visualize our findings, we plot the predictions of the model in Figure 2. Due to higher number of estimated coefficients, we report False-discovery rate adjusted p-values (Benjamini and Hochberg, 1995; Anderson, 2008). We observe that for the youngest cohort, both marginal selfish deviation and further selfish deviations from the norm are answered with a small (and insignificant) punishment increase (Model 1, coefficient for $Marginal_{neg}$ and D_{neg}). Looking at the 11-12 age cohort, the punishment increase to marginal deviation, together with the increase to further deviations becomes larger, and both become significant (Model 2, coefficient for $Marginal_{neg}$ and D_{neg}). Turning to the 12-13 age cohort, we observe again a smaller and insignificant effect of larger selfish deviations, but a much stronger punishment increase when reacting to the marginal selfish deviation from the norm, in comparison to the 11-12 cohort (Model 3, coefficient for $Marginal_{neg}$ and D_{neg}). Similarly, this pattern of strong increase of punishment on marginal selfish deviations, and weak increase on larger selfish deviations stays stable with the older cohorts (Model 4, 5; coefficient for $Marginal_{neg}$ and D_{neg}). If we look at the generous (positive) deviations, we observe that the punishment on larger generous deviations exhibits a decreasing pattern across all cohorts (coefficient for D_{pos}).¹⁵ The punishment to marginal deviation qualitatively changes, as in the 9-10 age cohort, children (insignificantly) decrease their punishment

¹⁴A reasonable explanation to the observed changes is the limitation of the action space in our design. Indeed, similar non-smoothness of the punishment curve can also be seen in all of the treatments of Bernhard et al. (2006), who similarly to our design, limit the action space of punishers to 0, 1 and 2 monetary units. If in such a limited punishment space the subject sufficiently increases the punishment amount with further deviations from the norm (e.g., for each token less, punishment increases by one), then by design, there can occur a “kink” in the pattern of that subject as his punishment cannot increase any further after he punishes with the maximal amount. In this way, a stronger increase of punishment on the marginal deviation might potentially be suggestive of a harsher punisher. In contrast to our design and the design of Bernhard et al. (2006), studies that allow third-parties to spend more than 40% of their endowment on punishment, and allow them to gradually increase the punishment amount by having a tighter grid of actions, report smooth punishment curves (e.g., Fehr and Fischbacher, 2003, 2004). Alternative to this explanation, it is possible that as subjects get older and come closer to adulthood, they start seeing any norm deviation as a violation worthy of strong punishment, while the degree of violation matters less. This, however, goes against common findings in studies with adults (e.g., Fehr and Fischbacher, 2003, 2004).

¹⁵Similarly as in Table 2, this decrease is mechanically supported by the design of the game.

Table 3: OLS regressions estimates of the effect of marginal and larger egalitarian norm deviations on punishment

Dependent var:	Punishment				
Age cohort:	9-10	11-12	13-14	15-16	17-18
	(1)	(2)	(3)	(4)	(5)
D_{neg}	0.037 (0.039) [0.434]	0.148*** (0.033) [0.001]	0.065 (0.054) [0.239]	0.046 (0.035) [0.195]	0.058* (0.028) [0.053]
$Marginal_{neg}$	0.081 (0.178) [0.651]	0.236** (0.105) [0.042]	0.529*** (0.151) [0.003]	0.426*** (0.140) [0.007]	0.523*** (0.114) [0.001]
D_{pos}	-0.097** (0.037) [0.035]	-0.120*** (0.025) [0.001]	-0.100** (0.036) [0.013]	-0.162*** (0.029) [0.001]	-0.064** (0.025) [0.025]
$Marginal_{pos}$	-0.209 (0.144) [0.260]	0.264* (0.151) [0.094]	0.416*** (0.136) [0.009]	0.362* (0.172) [0.053]	0.219* (0.111) [0.053]
Constant	0.938*** (0.142) [0.001]	0.360*** (0.115) [0.001]	0.226*** (0.077) [0.001]	0.513*** (0.116) [0.001]	0.340*** (0.085) [0.001]
Observations	352	275	341	429	583
R-squared	0.204	0.319	0.211	0.232	0.222
Adj. R-squared	0.195	0.309	0.202	0.224	0.217

Clustered standard errors at the individual level in parentheses. False-discovery rate adjusted p-values in brackets (Benjamini and Hochberg, 1995; Anderson, 2008), adjusted across all estimated coefficients within each age cohort. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

to marginal generous deviation, while with the older cohorts, they significantly increase it (coefficient for $Marginal_{pos}$). As can be seen from Figure 2, this change is primarily driven by the stark decrease of punishment on the egalitarian norm when going from 9-10 to 11-12 years of age.¹⁶

Altogether, the alternative model indicates that at the aggregate level, the punishment pattern develops until 13-14 years of age, and stays stable until adulthood.

3.2 Behavior of the dictators

In the last step of our analysis, we inspect the sharing behavior of the dictators. As we observe that the development of norm-enforcing behavior primarily happens in the period from 9-10 to 13-14 years of age, we investigate whether this is reflected in the dictators' behavior. Fig. 4 shows the average transfer by age cohort. Children from the youngest cohort transfer the smallest average amount (1.864 tokens). Regressing the amount of transfer on the age cohort yields a positive coefficient ($p = 0.039$, OLS regression model, $N = 193$).¹⁷ However, this increase in transfers seems to be non-linear. It primarily occurs from 11-12 to 13-14 years of age, where the average transfer increases from 1.909 to 3.233 ($p = 0.016$, two-sided t-test, $N(11-12) = 33$, $N(13-14) = 30$). For the older cohorts, the transfer decreases slightly in comparison to the 13-14 age cohort, although the decrease is not statistically significant ($p > 0.179$ for both 15-16 and 17-18 age cohorts, two-sided t-test, $N(13-14) = 30$, $N(15-16) = 61$, $N(17-18) = 47$). This pattern indicates that the development of the dictators' behavior roughly mirrors that of the punishers' behavior.

4 Discussion

In this section, we discuss potential mechanisms behind our findings. First, we focus on the punishment behavior of our youngest cohort: The 9-10 year-old children use the punishment tool but not to enforce the egalitarian norm, as they gradually decrease their punishment with the dictator's transfer and still punish the equal share. These findings relate to the discussion on the development of other-regarding behavior. On the one hand, our findings are at odds with studies reporting prevalent egalitarian behavior at the age of our youngest cohort. Most popularly, Fehr et al. (2008) use three simple games to classify 3-8 year-old children in other-regarding types. They observe that the majority of children (60%) become egalitarian by the age of 7-8. They find that other preferences,

¹⁶See Tables B4 and B7 in the Online appendix for robustness checks. The reported results remain robust.

¹⁷The result remains robust when including both subjects that answered the control questions and those that failed to do so ($p = 0.041$, OLS regression model, $N = 209$).

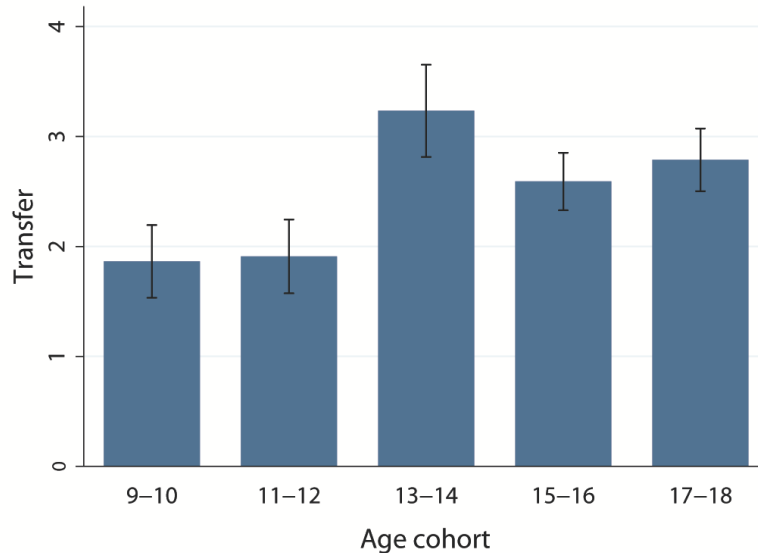


Figure 4: Average dictator’s transfer per age cohort. Error bars show standard errors of the means.

such as spiteful preferences where the subject’s utility decreases with the other’s payoff, diminish by that age. On the other hand, several studies report that spiteful preferences persist longer, offering a potential explanation for our findings. Fehr et al. (2013) also classify other-regarding types using a more similar experimental procedure to ours¹⁸, finding that the largest proportion of subjects are still spiteful (42%) at the age of 8-9 (alternatively, these subjects could also care about advantageous inequality, i.e. they prefer to have more than others). Moreover, McAuliffe et al. (2014) also find high levels of spiteful behavior for similarly-aged children as our youngest cohort, while both McAuliffe et al. (2014) and Fehr et al. (2013) report that this type of behavior decreases with age. Hence, it seems unclear when spitefulness exactly diminishes, and in line with the latter studies, our youngest punishers could have spiteful preferences and punish the dictator simply to reduce his payoff, or they could also care about advantageous inequality and punish the dictator to have more than him.

Furthermore, the comparison with Jordan et al. (2014) and McAuliffe et al. (2015) also supports this explanation. In particular, McAuliffe et al. (2015)¹⁹ run an experiment with 5-6 year-old children in which a third party can punish unfair allocations between two players. In comparison to our design, the punisher always obtains a much larger

¹⁸While both Fehr et al. (2008) and Fehr et al. (2013) use the same binary games, in Fehr et al. (2013) i) the experiment was conducted in large group sessions in a classroom and not one-to-one with the experimenter, and ii) the children were incentivized with money rather than candy. The same holds true for our experiment.

¹⁹Jordan et al. (2014) run an experiment with the same game and similarly-aged children as in McAuliffe et al. (2015), but they investigate the influence of in- and out-group bias and its emergence.

endowment in the one-shot game, hence, he is not comparing himself with others in the same egalitarian way as in our design. In contrast to the findings with our youngest cohort, they observe that 6 year-old children already do not punish much on the fair choice. This difference could be explained by spiteful preferences, or preferences towards advantageous inequality.²⁰ In support of this explanation, the findings of Gummerum and Chu (2014) are more fitting with the behavior of our youngest cohort. While our design has several differences compared with that of Gummerum and Chu, the punisher’s endowment is not one of them.²¹ In particular, the authors investigate the development of concerns for outcomes and intentions with second- and third-party punishers. In order to manipulate intentions, they run several mini-third-party punishment games, where the dictator can only choose between two division options. If we look at the specific game where dictators choose between dividing equally or keeping most for themselves, the authors report a substantial average amount spent on punishing the equal share for 8 year-olds, which decreases for older subjects. This is in line with our findings.

Conclusively, while we cannot deduce a single motive, spiteful preferences or preferences towards advantageous inequality are a plausible explanation for the behavior of our youngest cohort.²² Moreover, the comparison between our study and that of McAuliffe et al. (2015) might offer a valuable insight. McAuliffe et al. indicate that children already care about the fair allocation between two subjects at a very young age. However, in an egalitarian setup where all subjects possess the same endowment, our study would suggest that a lot of children are still driven by other factors at the age of 9-10, i.e., they might still be driven by self-comparison motives such as spiteful preferences or preferences towards advantageous inequality. Hence, if children’s understanding of a normative violation exists at a younger age, it might be dominated by other factors until 11-12 years of age.

Next, we discuss our finding regarding above-egalitarian punishment. We observe that some subjects, in addition to enforcing the egalitarian norm when it comes to selfish deviations, also punish above-egalitarian deviations. This behavior is inconsistent with outcome-based models of other-regarding behavior (e.g. Fehr and Schmidt, 1999),

²⁰In the case of using spitefulness as an explanation, a subject’s utility component representing the influence of other’s payoff should be concave in the payoff differences (at least in the advantageous domain), i.e., reducing the other player’s payoff by 1 yields a larger utility increase if payoffs are close than if the other player has a much smaller payoff.

²¹Gummerum and Chu (2014) implement several mini-third-party games, where the dictator always has two possibilities how to divide 10 points. He can choose to give 2 points or an alternative allocation. Depending on the game, the alternative allocation can be 0, 2, 5 or 8 points. The punisher is endowed with 5 points, and can punish the dictator with up to 5 points. Each token reduces the dictator’s payoff by 2.

²²Another explanation for the punishment pattern that we observe is that young children only approve of overly-generous transfers. However, this explanation lacks support in the literature and might also seem unintuitive.

but could be explained with social norm based models (Krupka and Weber, 2013; Kimbrough and Vostroknutov, 2016). While both types of models emphasize the normative value of the equal share, the latter posit that each action has a social-normative value that indicates its appropriateness. Furthermore, Krupka and Weber (2013) and Kimbrough and Vostroknutov (2016) show through norm elicitation methods that in the dictator game people on average approve more of the egalitarian transfer than above-egalitarian transfers. More specifically, there is a consensus about the appropriateness of the egalitarian transfer, yet less so with above-egalitarian transfers. While some subjects approve of transferring high amounts, others do not. This is in line with our observation that only some subjects engage in punishment of above-egalitarian behavior. Alternatively, one can explain these results through descriptive norms, i.e., each deviation from the usual behavior encounters punishment (Irwin and Horne, 2013), or one can think of the above-egalitarian punishment as “derogation of the do-gooders” (Monin, 2007).

Finally, we discuss the finding that dictators’ transfers primarily increase in the period when the behavior of punishers develops, suggesting a potential connection between the two. This provides supporting evidence that punishment is one of the core mechanisms contributing to the commonly-observed increase of prosocial behavior with age (Harbaugh et al., 2003; Benenson et al., 2007; Fehr et al., 2008, 2013; Bauer et al., 2014; Kosse et al., forthcoming). Indeed, the mere threat of punishment is crucial for the prosocial and cooperative behavior of adults (Fehr and Gächter, 2002; Herrmann et al., 2008). Alternatively, it is possible that the observed increase of dictators’ transfers is not due to the emergence of punishment. An underlying variable, e.g., the internalization of norms could be driving the development both norm-enforcing punishment and prosocial behavior as we observe it. While investigating the internalization of norms is a very interesting question, it remains rather intangible to actually confirm when a behavior starts being a norm. Nevertheless, several studies indicate an understanding of norms at a very young age (e.g. Fehr et al., 2008; Schmidt and Sommerville, 2011), which questions the argument that the internalization of norms could be solely responsible for the observed simultaneous changes in dictators’ and punishers’ behavior.

5 Conclusion

In this paper, we have investigated the emergence and development of egalitarian norm enforcement in childhood and adolescence. We took the most commonly-used third-party punishment game for adults, adapted it to ensure children’s understanding and ran an experiment 9-18 year-olds. We show that a small proportion of subjects are already enforcing the egalitarian norm at the age of 9-10, and that this behavior strongly develops in the following years, as the norm enforcers become the most prevalent type starting

with 11-12 years of age, and the punishment pattern fully develops until 13-14 years of age. Our results suggest punishment as one of the core mechanisms that underpins children's development of prosocial behavior with increasing age. Furthermore, our findings suggest that the negative valuation of exceeding the egalitarian norm in contrast to following the norm, which was reported in studies with adults (Krupka and Weber, 2013; Kimbrough and Vostroknutov, 2016), has roots in a relatively young age.

References

- Almås, I., Cappelen, A., Sørensen, E., Tungodden, B., 2010. Fairness and the development of inequality acceptance. *Science* 328, 1176–1178.
- Anderson, M.L., 2008. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association* 103, 1481–1495.
- Andreoni, J., Bernheim, B.D., 2009. Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica* 77, 1607–1636.
- Bauer, M., Chytilová, J., Pertold-Gebicka, B., 2014. Parental background and other-regarding preferences in children. *Experimental Economics* 17, 24–46.
- Benenson, J.F., Pascoe, J., Radmore, N., 2007. Children’s altruistic behavior in the dictator game. *Evolution and Human Behavior* 28, 168–175.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)* , 289–300.
- Bernhard, H., Fischbacher, U., Fehr, E., 2006. Parochial altruism in humans. *Nature* 442, 912–915.
- Brocas, I., Carrillo, J.D., Kodaverdian, N., 2017. Altruism and strategic giving in children and adolescents. CEPR Discussion Paper No. DP12288 .
- Camerer, C., 2003. Behavioral game theory: Experiments in strategic interaction. Princeton: Princeton University Press.
- Deckers, T., Falk, A., Kosse, F., Pinger, P., Schildberg-Hörisch, H., 2017. Socio-economic status and inequalities in children’s iq and economic preferences. IZA Discussion Paper No. 11158.
- Engel, C., 2011. Dictator games: A meta study. *Experimental Economics* 14, 583–610.
- Falk, A., Fischbacher, U., 2006. A theory of reciprocity. *Games and Economic Behavior* 54, 293–315.
- Falk, A., Kosse, F., 2016. Early childhood environment, breastfeeding and the formation of preferences. HCEO Working Paper 2016-036.
- Fehr, E., Bernhard, H., Rockenbach, B., 2008. Egalitarianism in young children. *Nature* 454, 1079–1083.
- Fehr, E., Fischbacher, U., 2003. The nature of human altruism. *Nature* 425, 785–791.
- Fehr, E., Fischbacher, U., 2004. Third-party punishment and social norms. *Evolution and Human Behavior* 25, 63–87.
- Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. *Nature* 415, 137–140.

- Fehr, E., Glätzle-Rützler, D., Sutter, M., 2013. The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence. *European Economic Review* 64, 369–383.
- Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics* 114, 817–868.
- Gummerum, M., Chu, M.T., 2014. Outcomes and intentions in children’s, adolescents’, and adults’ second-and third-party punishment behavior. *Cognition* 133, 97–103.
- Harbaugh, W.T., Krause, K., Liday, S.G.J., 2003. Bargaining by children. University of Oregon Economics Working Paper No. 2002-4 .
- Henrich, J., McElreath, R., Barr, A., Esminger, J., Barrett, C., Bolyanatz, A., Cardenas, J., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., Ziker, J., 2006. Costly punishment across human societies. *Science* 312, 1767–1770.
- Herrmann, B., Thöni, C., Gächter, S., 2008. Antisocial punishment across societies. *Science* 319, 1362–1367.
- Irwin, K., Horne, C., 2013. A normative explanation of antisocial punishment. *Social Science Research* 42, 562–570.
- Jordan, J., McAuliffe, K., Rand, D., 2015. The effects of endowment size and strategy method on third party punishment. *Experimental Economics* , 1–23.
- Jordan, J.J., McAuliffe, K., Warneken, F., 2014. Development of in-group favoritism in children’s third-party punishment of selfishness. *Proceedings of the National Academy of Sciences* 111, 12710–12715.
- Kimbrough, E.O., Vostroknutov, A., 2016. Norms make preferences social. *Journal of the European Economic Association* 14, 608–638.
- Kosse, F., Deckers, T., Pinger, P., Schildberg-Hörisch, H., Falk, A., forthcoming. The formation of prosociality: Causal evidence on the role of social environment. *Journal of Political Economy* .
- Kriss, P.H., Weber, R.A., Xiao, E., 2016. Turning a blind eye, but not the other cheek: On the robustness of costly punishment. *Journal of Economic Behavior & Organization* 128, 159–177.
- Krupka, E.L., Weber, R.A., 2013. Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association* 11, 495–524.
- Lergetporer, P., Angerer, S., Glätzle-Rützler, D., Sutter, M., 2014. Third-party punishment increases cooperation in children through (misaligned) expectations and conditional cooperation. *Proceedings of the National Academy of Sciences* 111, 6916–6921.
- Lewisch, P., Ottone, S., Ponzano, F., 2011. Free-riding on altruistic punishment? an experimental comparison of third-party punishment in a stand-alone and in an in-group environment. *Review of Law and Economics* 7.

- McAuliffe, K., Blake, P.R., Warneken, F., 2014. Children reject inequity out of spite. *Biology Letters* 10, 20140743.
- McAuliffe, K., Jordan, J.J., Warneken, F., 2015. Costly third-party punishment in young children. *Cognition* 134, 1–10.
- Monin, B., 2007. Holier than me? Threatening social comparison in the moral domain. *Revue internationale de psychologie sociale* 20, 53–68.
- Parks, C.D., Stone, A.B., 2010. The desire to expel unselfish members from the group. *Journal of Personality and Social Psychology* 99, 303–310.
- Schmidt, M.F., Sommerville, J.A., 2011. Fairness expectations and altruistic sharing in 15-month-old human infants. *PloS one* 6, e23223.
- Sutter, M., Feri, F., Glätzle-Rützler, D., Kocher, M.G., Martinsson, P., Nordblom, K., 2018. Social preferences in childhood and adolescence. a large-scale experiment to estimate primary and secondary motivations. *Journal of Economic Behavior & Organization* 146, 16–30.

Appendix A

A.1 Robustness checks

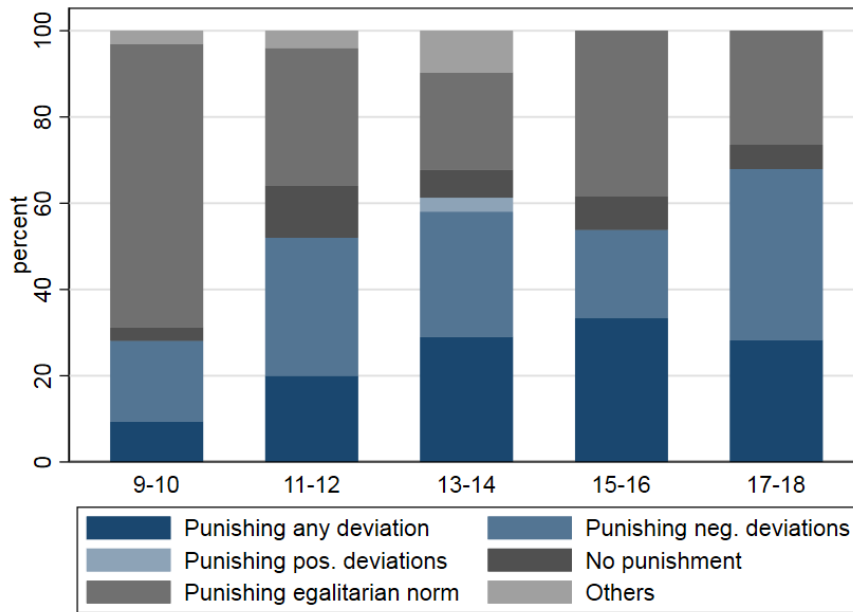


Figure A1: The distribution of punishment types across age cohorts - Robustness check 1. *Punishing any deviation*: Subjects that spend a total of at least 2 tokens on punishing any of the selfish (negative) deviations and a total of at least 2 tokens on punishing any of the generous (positive) deviations from the egalitarian norm and do not punish otherwise. *Punishing negative deviations*: Subjects that spend a total of at least 2 tokens on punishing any of the selfish (negative) deviations from the egalitarian norm and do not punish otherwise. *Punishing positive deviations*: Subjects that spend a total at least 2 tokens on punishing any of the generous (positive) deviations from the egalitarian norm and do not punish otherwise. *No punishment*: Subjects that do not punish on any of the possible transfers. *Punishing egalitarian norm*: Subjects that punish on the egalitarian norm. *Others*: Subjects that do not fit in any of the first five categories.

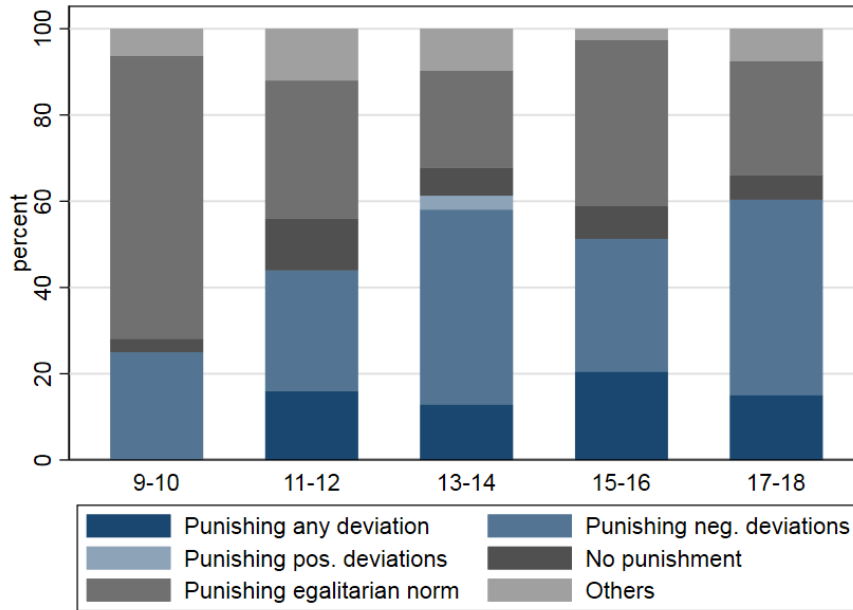


Figure A2: The distribution of punishment types across age cohorts - Robustness check 2. *Punishing any deviation*: Subjects that spend a total of at least 3 tokens on punishing any of the selfish (negative) deviations and a total of at least 3 tokens on punishing any of the generous (positive) deviations from the egalitarian norm and do not punish otherwise. *Punishing negative deviations*: Subjects that spend a total of at least 3 tokens on punishing any of the selfish (negative) deviations from the egalitarian norm and do not punish otherwise. *Punishing positive deviations*: Subjects that spend a total of at least 3 tokens on punishing any of the generous (positive) deviations from the egalitarian norm and do not punish otherwise. *No punishment*: Subjects that do not punish on any of the possible transfers. *Punishing egalitarian norm*: Subjects that punish on the egalitarian norm. *Others*: Subjects that do not fit in any of the first five categories.

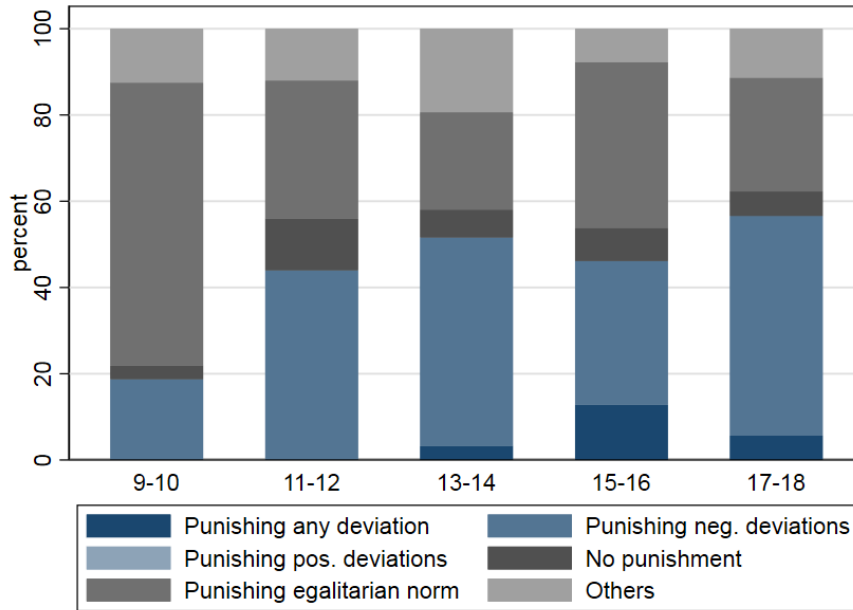


Figure A3: The distribution of punishment types across age cohorts - Robustness check 3. *Punishing any deviation*: Subjects that spend a total of at least 4 tokens on punishing any of the selfish (negative) deviations and a total of at least 4 tokens on punishing any of the generous (positive) deviations from the egalitarian norm and do not punish otherwise. *Punishing negative deviations*: Subjects that spend a total of at least 4 tokens on punishing any of the selfish (negative) deviations from the egalitarian norm and do not punish otherwise. *Punishing positive deviations*: Subjects that spend a total of at least 4 tokens on punishing any of the generous (positive) deviations from the egalitarian norm and do not punish otherwise. *No punishment*: Subjects that do not punish on any of the possible transfers. *Punishing egalitarian norm*: Subjects that punish on the egalitarian norm. *Others*: Subjects that do not fit in any of the first five categories.

Table A1: Probit regressions estimates of the effect of age cohorts on being a norm enforcer

Dependent var.:	Norm enf. [†] (≥ 1 tokens) (1)	Norm enf. (≥ 1 tokens) (2)	Norm enf. (≥ 2 tokens) (3)	Norm enf. (≥ 2 tokens) (4)	Norm enf. (≥ 3 tokens) (5)	Norm enf. (≥ 3 tokens) (2)	Norm enf. (≥ 4 tokens) (7)	Norm enf. (≥ 4 tokens) (8)
Basis: 9-10 age cohort								
11-12 age cohort	0.247* (0.129)	0.264** (0.130)	0.239* (0.128)	0.229* (0.126)	0.190 (0.125)	0.175 (0.121)	0.252** (0.121)	0.231** (0.113)
13-14 age cohort	0.397*** (0.116)	0.417*** (0.127)	0.332*** (0.118)	0.329*** (0.127)	0.363*** (0.116)	0.302** (0.125)	0.329*** (0.113)	0.255** (0.112)
15-16 age cohort	0.226** (0.114)	0.343** (0.150)	0.257** (0.113)	0.366** (0.145)	0.263** (0.111)	0.337** (0.142)	0.274*** (0.106)	0.364*** (0.129)
17-18 age cohort	0.367*** (0.104)	0.474*** (0.134)	0.398*** (0.102)	0.493*** (0.129)	0.354*** (0.102)	0.448*** (0.128)	0.379*** (0.097)	0.510*** (0.112)
Gender (1 if male)		-0.014 (0.081)		-0.045 (0.081)		-0.074 (0.081)		-0.091 (0.080)
Cognitive abilities [‡]		0.049 (0.040)		0.039 (0.040)		0.026 (0.041)		0.026 (0.040)
Pocket money (1 if yes) [§]		-0.018 (0.102)		-0.073 (0.103)		-0.040 (0.104)		-0.051 (0.103)
Family income		0.019 (0.027)		0.025 (0.027)		0.031 (0.027)		0.029 (0.027)
Siblings (1 if subject has a sibling)		-0.215* (0.121)		-0.251** (0.121)		-0.298** (0.119)		-0.245** (0.112)
Observations	180	163	180	163	180	163	180	163

The reported coefficients represent average marginal effects. Models (2), (4), (6) and (8) include controls for high school program fixed effects. *** p<0.01, ** p<0.05, * p<0.1. † Norm enforcer, ≥ 1 token (2, 3, 4 tokens) is 1 if a subject does not punish on the transfer of 5 but uses a total of at least 1 token (2, 3, 4 tokens) for punishing transfers smaller than 5 or uses at least 1 token (2, 3, 4 tokens) for punishing transfers larger than 5; and it is 0 otherwise. ‡ Cognitive abilities were measured with a 16-item matrices test. The value was normed for each of the age cohorts using the dictators', recipients' and punishers' test results to avoid confounding age and cognitive abilities effects. § As it is not common in the Croatian culture to give fixed weekly allowance to children, we only measure if a child receives money or not. In addition, we also include family income, measured in 7 categories. It takes a value of 1 if family income < 2000 HRK, 2 if family income ≥ 2000 and < 4000 HRK, 3 if family income ≥ 4000 and < 6000 HRK, ..., 7 if family income ≥ 12000 HRK.

Appendix B

Appendix B can be found in the Online appendix. It contains the analysis of potential imbalances across age cohorts and across subjects who failed or succeeded in answering the control questions, experimental instructions, control questions, and additional tables and figures.