Gutenberg School of Management and Economics

& Research Unit “Interdisciplinary Public Policy”

Discussion Paper Series

Do Children Cooperate Conditionally? Adapting the Strategy Method for First-Graders

Florian Hett, Mario Mechtel, Henning Müller,

Felix Schmidt, Daniel Schunk, Valentin Wagner

January 2018

Discussion paper number 1803

Johannes Gutenberg University Mainz
Gutenberg School of Management and Economics
Jakob-Welder-Weg 9
55128 Mainz
Germany
https://wiwi.uni-mainz.de/

Contact Details:

Florian Hett
Goethe University Frankfurt,
60323 Frankfurt am Main
Germany
hett@econ.uni-frankfurt.de

Mario Mechtel
Institute of Economics
Leuphana University Lüneburg,
Universitätsallee 1, 21335 Lüneburg
Germany
mario.mechtel@leuphana.de

Henning Müller
Department of Economics / FAIR
NHH Bergen,
Helleveien 30, 5045 Bergen
Norway
henning.mueller@nhh.no

Felix Schmidt
Department of Economics
Johannes Gutenberg University
Jakob-Welder-Weg 4
55128 Mainz
Germany
felix.schmidt@uni-mainz.de

Daniel Schunk
Department of Economics
Johannes Gutenberg University
Jakob-Welder-Weg 4
55128 Mainz
Germany
daniel.schunk@uni-mainz.de

Valentin Wagner
Düsseldorf Institute for Competition Economics,
Gebäude 24.31, Universitätsstraße 1, 40225 Düsseldorf,
Germany
wagner@dice.uni-duesseldorf.de

Do Children Cooperate Conditionally? Adapting the Strategy Method for First-Graders

Florian Hett[1], Mario Mechtel[2], Henning Müller[3], Felix Schmidt[4], Daniel Schunk[4], and Valentin Wagner[5]

January 31, 2018

Abstract

We develop a public goods game (PGG) to measure cooperation and conditional cooperation in young children. Our design addresses several obstacles in adapting simultaneous and sequential PGGs to children who are not yet able to read or write, do not possess advanced abilities to calculate payoffs, and only have a very limited attention span at their disposal. It features the combination of haptic offline explanation, fully standardized audiovisual instructions, computerized choices based on touchscreens, and a suitable incentive scheme. Applying our experimental protocol to a sample of German first-graders, we find that already 6-year-olds cooperate conditionally and that the relative frequency of different cooperation types matches the findings for adult subjects. We also find that neither survey items from teachers nor from parents predict unconditional or conditional cooperation behavior; this underlines the value of incentivized experimental protocols for measuring cooperation in children.

Acknowledgments: The authors are thankful to a group of elementary school teachers in Mainz who participated in a preparatory workshop, the administration, parents, children and teachers from the participating schools, as well as Ida Jakob, who provided valuable research assistance. The authors gratefully acknowledge financial support by the German Research Foundation (DFG), grant number SCHU 2828/4-1. Henning Müller gratefully acknowledges financial support by the Research Council of Norway through its Centres of Excellence Scheme, FAIR (projects No 262675, 262636 & 250170F10).

Keywords: Conditional cooperation, strategy method, public goods game, revealed preferences, measurement, children, ingroup bias, group identity

JEL-codes: H41, C71, C91

[1]Goethe University Frankfurt, hett@econ.uni-frankfurt.de
[2]Leuphana University Lüneburg, Institute of Economics, mario.mechtel@leuphana.de
[3]NHH Bergen, Department of Economics / FAIR, henning.mueller@nhh.no
[4]University of Mainz, Department of Economics, felix.schmidt@uni-mainz.de, daniel.schunk@uni-mainz.de
[5]Düsseldorf Institute for Competition Economics, wagner@dice.uni-duesseldorf.de

1 Introduction

A large literature across the social sciences shows that cooperation is a key driver of economic and social prosperity (Bowles and Gintis 2011, Fehr and Fischbacher 2003, Ostrom et al. 1999). This important role of cooperation in affecting the well-being of groups and societies implies the need to study how cooperative preferences develop and if and how they can be affected. What is the role of societal norms and culture as compared to genetic factors? Do formal or informal institutions play a substantial role? Can public policies directly or indirectly shape preferences for cooperation? Studying these questions empirically means to take a perspective covering the whole lifespan, explicitly including the very early years.

However, most of the existing evidence is limited to adults or adolescents, while we still know little about individual preferences for cooperation in (young) children. Studying such preferences requires the ability to measure them in a precise, accurate, and controlled way, in particular in settings and with subjects for which this is difficult. Such a measure will then allow to empirically analyze how cooperation changes in reaction to a broad variety of possible determinants and thus produce valuable evidence potentially informing the academic literature as well as policy-makers. In this paper, we develop an experimental protocol to measure cooperation that explicitly addresses first-graders who are not even able to read or write. Applying our method, we find that already 6-year-olds cooperate substantially and that a significant share of our subjects even behave as conditional cooperators.

While the general literature discusses several factors affecting individual and collective cooperation[1], methodological advancements in experimental economics carved out the pivotal role of *preferences for cooperation* (Bruhin et al. forthcoming, Fehr and Leibbrandt 2011, Kosfeld 2017, Rustagi et al. 2010). In particular, the literature shows that individual beliefs about what others do matter a great deal for actual behavior (Fischbacher and Gächter 2010, Fischbacher et al. 2001, 2012, Herrmann and Thöni 2009, Kocher et al. 2008, Kurzban and Houser 2005).[2] Behavior often takes the form of *conditional cooperation*: Many individuals are in principle willing to cooperate, but only if others do so as well. This conditionality of behavior implies that the mere fact that an individual does not cooperate is not sufficient to conclude that she has a low preference for cooperation because she might as well be a conditional cooperator who has pessimistic beliefs about the behavior of others. To really measure individual *preferences* for cooperation, one needs to control for such beliefs with an appropriate experimental design. For adult subjects, the so called strategy method (Selten 1967) has become the preferred tool of choice.

Implementing the strategy method to measure conditional cooperation in young children imposes a series of challenges. First, in comparison to adult subjects, the attention span of young children is severely limited and important cognitive abilities might not be readily developed, which among other things complicates the implementation of non-trivial payment schemes. Another obvious and crucial obstacle is the inability to read or write, which affects the communication of instructions and recording of choices while

[1] Previous studies have, for instance, documented the role of group size (Olson 1965) and (punishment) institutions (Ostrom 1990). In another vein, Axelrod and Hamilton (1981) illustrate how individual beliefs over the behavior of others crucially affect the establishment of cooperation.

[2] See also Chaudhuri (2011) for a survey of the literature on conditional cooperation.

at the same time maintaining a high level of experimental control. Finally, it is difficult to incentivize children at this age without introducing too much heterogeneity in the valuation of payoffs. We address these challenges with a novel public goods game design (Ledyard 1995) that (i) combines advantages of haptic offline games and computer-based testing, (ii) uses motivating and diversified auditory instructions and touchscreen input devices, accounting for children's ability to concentrate in this age-group, and (iii) implements a simple and straightforward incentive structure.

As an application we elicit unconditional and conditional cooperation in a sample of first-graders in Germany. Whether and to which extent 6-year-olds already display cooperative behavior and in particular conditional cooperation is interesting from several perspectives. For instance, from the view of developmental psychology it matters to what extent these children command the cognitive sophistication of thinking in contingencies as required for conditional cooperation. More generally, this also holds for children's behavior in strategic situations, i.e., decision environments featuring inter-dependencies of individual choices. Our results might also already serve as a piece of evidence in the above mentioned debate regarding the general determinants, malleability and origins of cooperative behavior, by providing data on cooperative behavior in a critical period of human development. As a second application, we also implement a treatment variation to test the sensitivity of cooperation to social context. Here, we study whether cooperation exhibits a form of ingroup bias, i.e., a differential treatment of others depending on their group membership, which is a robust finding for adult subjects across the social sciences (Chen and Li 2009, Hett et al. 2018, Tajfel and Turner 1986).

Our results are as follows. We see substantial levels of cooperation in unconditional (simultaneous) as well as conditional (sequential) versions of the public goods game (PGG). Importantly, when children can condition their behavior on the behavior of others, the average child clearly displays *conditional cooperation*, as contribution levels to the public good are the higher the more other children contribute. While this pattern captures the aggregate behavior of children, it masks substantial systematic heterogeneity at the individual level. This heterogeneity is captured by classifying behavioral patterns into different types of cooperation, including conditional cooperators, free-riders, as well as altruists. The distribution of these types by and large matches relative frequencies found with adult subjects (Fischbacher et al. 2001, Herrmann and Thöni 2009, Kocher et al. 2008, Kurzban and Houser 2005, Muller et al. 2008, Rustagi et al. 2010). We also compare behavior in our experiment with parent and teacher assessments of cooperativeness elicited via survey measures. We find that choices in our experiment, i.e., actual cooperative behavior, cannot be predicted by these corresponding survey items, emphasizing the need to implement experimental protocols such as ours in order to really capture cooperative behavior in children. Finally, our treatment variation reveals that cooperation behavior in our sample does not feature ingroup bias.

Our study relates to the broad literature on (conditional) cooperation[3], the literature aiming at measuring preferences and characteristics in children in general[4] and the literature on cooperation of children in public goods games in particular (Alencar et al. 2008, Cipriani et al. 2013, Harbaugh and Krause 2000, Peters et al. 2004, Vogelsang et al. 2014).[5] One of the first economic studies examining cooperative behavior in a public goods game is Harbaugh and Krause (2000) on 6–12 year old children (mean age of 8.8 years), illustrating the game with poker chips, a (private) cup and an (public) envelope. Further, they include non-integer payoffs, thereby potentially increasing the complexity of the experiment. Alencar et al. (2008) (mean age of children 8.7 years) and Cipriani et al. (2013) (average school grade 2.68) adopt the same design. In the latter study, the mere instructions already last for 40 minutes on average.[6] Moreover, experimenters instead of children play through the mock simulations of the game. Peters et al. (2004) use a laboratory setting to study whether children aged 8–16 contribute more to the public good when grouped with family members than with "strangers".[7] Vogelsang et al. (2014) use a mechanical device illustrating the allocation mechanism.[8] An advantage of this apparatus is the visualization of the doubling process in case children pull the public rope. While this kind of apparatus seems to be an age-appropriate way to explain the public goods game, its implementation and use is quite involved and only one participant can be tested at a time.[9]

None of the papers above (and to the best of our knowledge also no other study) implements the measurement of *conditional* cooperation of children with the strategy method.[10] Cipriani et al. (2013) and Vogelsang et al. (2014) argue that conditional cooperation can be measured by analyzing whether the per-round contribution depends on the contribution of other group members in previous rounds. However, such an approach is not able to pin down individual beliefs about the behavior of others but rather assumes a quite specific form of adaptive belief formation. This might also be a reason behind the mixed results found in these studies: While Vogelsang et al. (2014) find that children are, to some degree, conditional cooperators, Cipriani et al. (2013) find no support for children being conditional cooperators. In fact, our

[3]The literature on public goods games shows that individuals contribute to a public good although the contribution of a rational and selfish player would be zero (Ledyard 1995, Zelmer 2003). Further, those who contribute tend to be conditional contributors, this is, they contribute more the more they believe others will contribute (Chaudhuri 2011, Fischbacher et al. 2001, Kocher et al. 2008).

[4]See Angerer et al. (2015), Czermak et al. (2016), Kosse et al. (2017), Schildberg-Hörisch et al. (2015).

[5]There are several studies examining cooperation among children in two-player-games such as dictator games or prisoners dilemma games (Angerer et al. 2016, Cárdenas et al. 2015, Fan 2000, Gummerum et al. 2010, Sally and Hill 2006, Yu et al. 2016). However, we concentrate on the literature on public goods games as we are explicitly interested in cooperation in groups.

[6]This matters in particular as choices are obviously made only after these 40 minutes, when cognitive capacity might already be severely exhausted.

[7]Children receive instructions in written form, play practice rounds and the game is implemented with the VCM software package by the Economic Science Laboratory at the University of Arizona.

[8]A wooden table covered with a Plexiglas serves as a distributing apparatus. Children can pull two ropes, resulting in gumballs to roll off the apparatus into private or public boxes.

[9]The findings of the cited literature can be summarized as follows: the level of contribution of children is similar to that of adults (Harbaugh and Krause 2000); children contribute more to a public good when in groups with family members (Peters et al. 2004); parents' attitudes toward free-riding do not have a significant effect on their children's cooperation behavior (Cipriani et al. 2013); group size matters, this is, small-group children cooperate significantly more than children in large groups (Alencar et al. 2008).

[10]We are aware of one study in psychology using a virtual public goods game to analyze conditional cooperation of children 9–16 years of age (Keil et al. 2017). However, the experimental design of Keil et al. (2017) has several limitations: the game involved deception, the authors cannot compare the conditional behavior of their children to an unconditional situation and the "currency" were virtual pizza slices, hence children need to calculate exchange rates (the value of the gift children could choose at the end increased with the number of slices).

results precisely show that while children do not display this kind of supposed conditional cooperation, i.e., reacting to contribution levels in *previous* rounds, they do so once beliefs about others' contributions are fixed by properly implementing the strategy method.

To summarize, our study differs from the existing literature in the following extents: (i) we are able to measure conditional cooperation of children who are two years younger than children in previous studies[11], (ii) we develop a computerized measure of cooperation, thereby increasing the degree of control, standardization, and time-efficient implementability, (iii) our measure already functions with participants who are not yet able to read or write, have very limited calculation abilities, and a low attention span, (iv) our design further improves understanding by visualizing the "heart" of the public goods game—the multiplication of contributions to the public good, and (v) our measurement can be easily replicated and extended, e.g., by including a third-party-punishment stage. While previous studies do to some extent address one or the other of the aforementioned points, a key feature of our study design is to simultaneously integrate all issues within one measurement protocol.

The remainder of the paper is organized as follows. We first describe the experimental protocol we developed and explain how it addresses substantial challenges in measuring (conditional) cooperation in children. We then present the corresponding results, discuss their implications and highlight some potential avenues for future research. Finally, we conclude.

2 Study Design

Our study tackles three main issues, i.e., (i) developing a (computer-based) measurement protocol of conditional and unconditional cooperation suitable for young children, (ii) testing the magnitude and heterogeneity of cooperation behavior in first-graders, and (iii) testing children's ingroup-sensitivity of cooperation. Behavioral economists' standard tool to measure (preferences for) cooperation in groups is the public goods game (PGG). Adopting the PGG for the application with 6-year-olds imposes numerous challenges. Therefore, after explaining the general structure of the PGG, we discuss how our measurement strategy deals with these issues. We then describe the treatment variation addressing the ingroup-sensitivity of cooperation. Further details on the procedures of our study can be found in Section B.3 in the online appendix.

2.1 The Public Goods Game and the Strategy Method

The public goods game captures the fundamental social dilemma of cooperation: if each member of a group does what is best for her, it does not lead to maximizing well-being of the group as a whole (Hardin 1968, Olson 1965). We follow Cipriani et al. (2013), Peters et al. (2004) and Vogelsang et al. (2014) and put children into groups of 4 represented by 4 different colors.[12] Each child receives an endowment of

[11] One exemption is the study of Vogelsang et al. (2014) with children of nearly the same age as in our study. However, in addition to Vogelsang et al. (2014), our study design allows to simultaneously implement the strategy method and to test a large group of children.

[12] We abstain from using just a two-person version of the PGG, which basically resembles the prisoner's dilemma game, as we are interested in measuring cooperation in groups.

five pairs (tuples) of coins. Investments into the public good are doubled and, subsequently, evenly split among the players. Player i's payoff y_i is given by $y_i = y - c_i + \alpha G/4$, where y denotes each player's individual endowment, c_i is player i's contribution to the group account G, $G = \sum_{i=1}^{4} c_i$, and $\alpha = 2$ holds. The PGG is a social dilemma game as contributing 0 is the dominant strategy (if players only focus on their own payoff), while the efficient solution results if all players contribute their full endowment.

As argued above, individuals holding the same *preferences* for cooperation can display entirely different choices in a PGG, just because their *beliefs* about the behavior of others differ. For instance, a player with preferences for *conditional cooperation* will contribute to the public good only if she believes that others do so as well. Hence, just because she has rather pessimistic beliefs, her behavior in the simultaneous PGG might make her choices appear like choices by an uncooperative type. As one therefore cannot rule out that differences in individual beliefs and not preferences are driving individual heterogeneity in cooperation behavior, this represents a serious problem when inferring preferences for cooperation from behavior in the simultaneous PGG. The strategy method eliminates any differences in beliefs simply by allowing players to report contribution levels conditional on each potential combination of actions of others, i.e., making the game sequential instead of simultaneous. By then implementing the actual actions of others as well as the corresponding conditional reaction of the player under consideration, one can cleanly identify differences in preferences for cooperation because beliefs as a potential confound are ruled out by design.

—Table 1 about here—

Children play eight rounds of the PGG (see Table 1). The first four rounds are played as unconditional version, i.e., players choose their contribution to the public account simultaneously and do not know what other children do. In rounds 1–3, players can choose any integer contribution out of the set of $[0, ..., 5]$ tuples. After each round, players learn about their resulting payoff.[13] In rounds 5–8, we implement the strategy method to achieve full control over children's beliefs regarding other players' actions. To reduce complexity, we limit the action space in these 4 rounds to "contribute nothing" and "contribute everything". Children make their contribution decisions conditional on whether all others contribute 0 (round 5), one child contributes 5 and all others 0 (round 6), two children contribute 5 and one child 0 (round 7), or all children contribute 5 (round 8). To test whether this change in the action space affects contributions at the individual level, round 4 also allows only to contribute 0 or 5, but in the unconditional version of the PGG. Subsequently, the outcome of one randomly determined round is paid to each participant.

2.2 Adapting the Public Goods Game and the Strategy Method to Children

Measuring (conditional) cooperation with first-graders poses several challenges. In particular, they include explaining the complexity of the payment consequences, the limited attention span of children, creating standardized instructions for children who cannot read or write, and finding appropriate (non-monetary)

[13] Repeating the game allows us to analyze the persistence of individual behavior and whether children's choices are affected by the previous round(s') outcome(s). Cipriani et al. (2013) use that approach to analyze whether individual contributions depend on the others players' contributions in the *previous* round. This strategy, however, does not solve the problem of potential differences in individual beliefs about what others do in the *current* round.

incentives. While many of these issues do exist in one way or another also with other non-standard or even standard subject pools, many of the problems are amplified and sometimes even interact with each other when running the PGG with children. In the following, we therefore discuss how our design addresses these issues.

Complexity of Payment Consequences

One of the most demanding aspects of the PGG—even for adults—is the calculation of payoffs resulting from the players' contribution choices. This is particularly difficult if the players can neither read (you cannot assist them with any payoff table) nor do simple calculations (you cannot assist them with a payoff function). To address this, we do the following: Each player receives an endowment of 5 tuples of 2 golden coins.[14] These tuples can either be put into the "private basket" (a colored box in front of each player) or the "group basket" (a larger box in the colors of all 4 players).[15] Using coin tuples allows an easy visualization of how the number of coins in the group basket is doubled and subsequently distributed to players. Children can easily see that every coin tuple a child keeps for herself will result in two coins for herself and that every coin tuple she invests in the public good will first be doubled and then results in one coin for each of the four children, including herself.

The chosen calibration (5 tuples of 2 golden coins each, $\alpha = 2$) allows for an easy visual presentation, at the same time ruling out non-integer payoffs. The initial endowment of 5 coin tuples which each child receives at the same time (i) gives room for variation in contributions (0 to 5 tuples), (ii) encompasses numbers easy to handle for children in this age-group (e.g., fingers on one hand), and (iii) makes payoffs very easy to "calculate".

Limited Attention Span

The scope for experiments with children is typically quite restricted due to their limited attention span. Explanations must therefore be concise and varied, in order to keep the children focused. Our setup combines advantages of offline games and computer-based testing[16] and uses motivating and diversified instructions that account for children's ability to concentrate in this age-group. The game is explained both offline and online using a very similar visual presentation. Measurement directly follows the onscreen explanation. In pretests, this two-stage explanation proved to be very effective in (i) keeping the children focused on the game and (ii) allowing them to understand the incentive structure.

First, children assemble around a table on which the PGG is explained by the experimenter (the same person in all sessions). To foster understanding and to increase attention, following the first explanation of the game, the experimenter uses standardized questions to interact with the children and to actively engage them during the explanations. In doing so, the experimenter lets the children actively participate in the process by letting each of them put coin tuples into the baskets and asking what will happen next. Second, players move to their individual computer screens. On their screen, they see an animated video sequence

[14] See online appendix Figure C.1 for a picture.
[15] See Figure C.2 in the online appendix.
[16] The software is programmed using oTree (Chen et al. 2016).

of the game they saw on the table (see Figure C.3 in the online appendix).[17] The game is explained again and the children play a trial round.[18] Following this second round of explanations, the children privately make their decisions in the eight rounds of the PGG.

By switching modes of explanation and having a rich visualization of the PGG, we make sure that the children can follow instructions and the different rounds of the game although they have limited concentration spans. Using the computer, we also ensure a very high motivation and an intuitive user interface for making choices. A complete session including the payoff lasts approximately 30 minutes in total, such that it is guaranteed that each of the different parts is short enough to keep concentration at a high level.

Inability to Read or Write

While instructions and payoff functions can be easily displayed in a PGG with adults, this is not the case with 6-year-olds who cannot read. Contrary to adults, they are at the same time not able to report their chosen contribution levels by writing down or typing in numbers. Our offline and onscreen settings allow solving this problem in an intuitive way. Accompanying the video animations, the offline instructions are repeated with other words. The children hear these auditory instructions via headphones. Importantly, the doubling of coins in the group basket and the final allocation is supported by visual animations, both for the doubling of coins as well as the equal distribution to all four players. The children also see an animation of one exemplary scenario of how the game could be played.[19]

Following the instructions, the children learn how to use the visual surface of the game. They use their computer touch screens to make choices based on the same animations. They can simply use their fingers to select their decision.[20] The chosen visualization and auditory instructions guarantee a high degree of standardization, allow for a comprehensible explanation of the rules of the game and its payoff consequences, and enable children to make their decisions privately although they are not able to write. Children played a trial round in which they could familiarize themselves with the interface. To check for their understanding of the game, children are thereafter asked whether they understood the game completely. If yes, they can click a green thumb pointing up, if not, there is a red thumb pointing down.[21] Only 15 out of 129 children (11.6%) asked for additional explanation from the experimenter after listening to the audio instructions. Thus, we are quite optimistic that our effort to explain the game in a way suitable for children was successful (see Section 4 for a more detailed discussion on the problem of proper understanding of the game).

[17] Each child is seated in front of a computer (notebook and large-scale touchscreen with 22"). Visual separators are placed between each table to prevent children from looking on the other screens.

[18] See the online appendix Section C.2 for the auditory instructions.

[19] In our online scenario one player contributes 0 coin tuples, one player contributes 5, and the remaining two players contribute 3. Furthermore, the offline explanation consists of two scenarios. One in which all players contribute to the public good and another one in which all but one player contribute.

[20] Coin tuples are initially placed in a neutral position between a player's individual basket and the group basket. There are two arrows to move coin tuples into the own basket and the group basket. Children can switch their decision as many times as they want and then confirm their final allocation with a green check at the right hand side of the screen. A bar which is getting smaller over 30 seconds to nudge children to make their decision is placed at the bottom of the screen. This is important to keep up the motivation of the group as children can only proceed to the next round once every player made her decision. However, there was no real consequence of overrunning the 30 seconds.

[21] If children click on the red thumb, an experimenter privately explains the game once more.

Incentives

The translation of experimental currency into money (or an equivalent) is much more demanding with primary school children than adults. Simply exchanging golden coins against money is likely to add much noise as 6-year-olds have a large degree of heterogeneity regarding their parents' dealings with money, their financial capability and children's degree of being used to deal with money in everyday life. All these factors would make it difficult to apply induced value theory (Smith 1976) in our setting when relying on real money. Furthermore, ethical reasons might speak against exchanging golden coins against money.

Some experimental studies therefore exchange an experimental currency against small gifts such as toys, sweets, books, etc. (cf. Angerer et al. 2016, Harbaugh and Krause 2000, Schildberg-Hörisch et al. 2015, Schunk et al. 2017). However, there is still potential for a significant degree of heterogeneity concerning the individual children's valuation for these goods. In the run-up of the study we consulted with elementary teachers in a one-day workshop to discuss the applicability of our study design. During this workshop, the teachers suggested to use the golden coins themselves as incentives as they would make it very easy for children to understand the direct consequences of their choices. This suggestion was in line with our experiences from previous studies (Schunk et al. 2017) that demonstrated that golden coins themselves carry a high intrinsic value for primary school children. Also, no exchange rate has to be applied to evaluate final incentives.

In consequence, we used the following, very simple incentive structure in our experiment: We randomly select one of the eight PGG rounds for each child and she earns the number of coins from that round, paid out in "real" golden coins. Given the above considerations, the advantages of this way of incentivizing are straightforward: (i) incentives are very direct—there is no need for children to apply any kind of calculations, prices, exchange rates, etc., (ii) incentives are perfectly discrete, (iii) incentives are extremely salient to children because they are tangible, and (iv) Levitt et al. (2016) show that *young* children react stronger to non-monetary rewards compared to financial incentives.

2.3 Group Identity

Across many studies in social psychology as well as economics, adult behavior displays *ingroup bias*, i.e., a differential treatment of others depending on their group membership (Chen and Li 2009, Hett et al. 2018, Tajfel and Turner 1986). A smaller, but growing literature indicates that already children might be prone to this kind of behavior (Angerer et al. 2016, Peters et al. 2004). While these studies rely on naturally existing groups like classes and families, our study setup allows to investigate whether ingroup bias is already triggered by a treatment that much closer resembles the classical minimal group paradigm (Tajfel 1970). Contrary to the mentioned studies, we therefore randomly assign children to groups *within classes* and interaction takes place anonymously.[22]

Children of the same class are randomly assigned either into a "turtle group" or an "elephant group". Corresponding cartoon animal badges are distributed to children prior to the experiment (see Figure C.4

[22] In Section B.2 in the online appendix we describe our randomization procedures in detail.

in the appendix). Children are picked up in their classroom and guided to the experimenter room in separated elephant and turtle groups. During the offline explanation of the game, the two groups have to stand around a table—each animal group on one side of the table. Additionally, we created a starting screen for the online experiment showing the two animals. Children are requested to click on their animal in order to start the game. Furthermore, we placed the animal icon of each player next to its basket throughout the experiment. Thus, children could see to which animal group the three other players belonged (see Figure C.3 in the appendix). We repeat in the auditory instructions that in the ingroup treatment all other players belong to the same animal group and in the outgroup treatment that two players belong to the other animal group.

This procedure shows that while being minimal, our group treatment is quite strong and salient. In the *ingroup treatment*, all four children who play the PGG together belong to the same group (either the turtle or the elephant group), while in the *outgroup treatment*, two children from the turtle group and two children from the elephant group play the PGG together. If ingroup bias in cooperation among children exists, average contributions to the group account will be higher in the ingroup treatment than in the outgroup treatment.

3 Results

We present our results in the following order: We start by analyzing behavior in the simultaneous version of the PGG (the first four rounds), where children cannot condition their behavior on the actions of others. Second, we consider choices made in the sequential version of the PGG (using the strategy method) and check whether average behavior is consistent with conditional cooperation. After that we move to the individual level to assess relative frequencies of different cooperation types including conditional cooperators. Next, we test whether survey items on children's cooperative behavior can predict actual cooperation in the experiment. Finally, we investigate our treatment variation to see whether cooperation in children is sensitive to ingroup bias.[23]

3.1 Public Goods Game: Unconditional Cooperation

Result 1 – Unconditional cooperation levels: *On average, children contribute substantially to the public good. Cooperation levels vary, with a substantial part of this variation being between children.*

—Figure 1 about here—

Figure 1 shows the average contributions of children to the public good in the unconditional rounds. We see that children contribute on average 1.87 units out of a maximum of 5. The contribution levels do not significantly vary over rounds, i.e., there are no aggregated patterns like a general increase or decrease in contribution levels. Further, limiting the action space of children to just "contribute everything" or "do not contribute" in the fourth round (R4) does not lead to differences in aggregate contributions, making cooperation levels comparable across the unconditional and conditional game variants.[24] The

[23] Section B.1 in the online appendix describes the composition and characteristics of our sample in detail.
[24] Table A.1 in the online appendix breaks down the distribution of contributions by round.

overall standard deviation of cooperation is 1.55 (see Table A.1 in the online appendix), when we focus only on rounds 1 to 3, where the action space was held constant. This variation can be decomposed into a between-children component of 1.26 and a within-child component of 0.90 (see also Figure A.1 in the online appendix). Hence, while also the variation within children is sizable, the fact that variation is more pronounced between children serves as a first indication for systematic interpersonal heterogeneity in cooperation.

Given that children receive feedback on the outcome of previous rounds before taking the next decision, there would in principle be scope to react to this information.[25] Table 2 investigates these potential patterns. We run OLS estimations with a child's unconditional contribution in rounds 2–4 as the dependent variable and include a set of standard child-specific control variables as well as session fixed effects capturing session related features like time of the day but also teacher characteristics as all children in one session are from the same class. Column (1) shows that own contributions in round 2 significantly correlate with own contributions in round 1. However, average previous contributions of the other three players do not affect own contributions at all. The same holds in rounds 3 and 4 (columns (2) to (5)). The results do not change when limiting the sample to those children who directly understood the game according to their answer to the respective control question (column (6)). This significant impact of previous own behavior on contribution levels reflects a certain robustness and persistence of cooperative behavior, indicating our measure to be informative and meaningful. In turn, the result that children do not systematically react to the behavior of others in previous rounds is in line with previous findings, for instance by Cipriani et al. (2013), who interpret this insensitivity to previous contributions as an absence of conditional cooperation. We will address this interpretation in the next section when we analyze choices in the strategy method.

—Table 2 about here—

3.2 Public Goods Game: Conditional Cooperation

Result 2 – Aggregated conditional cooperation levels: *On average, children are conditional cooperators. The more other children contribute, the more they are likely to contribute.*

—Figure 2 about here—

We now turn to behavior in the sequential PGG where children can decide whether to contribute or not *conditional* on the number of other children contributing. As the first two bars in Figure 2 show, average contributions hardly differ between the simultaneous PGG (rounds 1–4) and sequential PGG (rounds 5–8) when pooling choices across all rounds ($p = 0.43$).[26] However, the key question is whether cooperation varies systematically with the behavior of other children. The remaining bars in Figure 2 show that the average contribution in the unconditional PGG is in line with contributions when one or two other children

[25] Importantly, children could observe the doubling and distribution of the final amount of coins in the basket but there was no information displayed how much each individual player contributed.

[26] This also translates to the individual level: Table A.2 in the online appendix shows that behavior across the simultaneous and sequential PGG is strongly related within-child which indicates a certain consistency in behavior and adds further confidence to our measure.

also contribute (SM 1/2).[27] In turn, average contributions are significantly lower when all other children free-ride (1.01) and significantly higher when all other children contribute (2.83). These averages result from the fact that 20.16% of the children contribute even if none of the other players contribute (SM 0) and in contrast 56.59% of the children contribute if all other player contribute (SM 3).

Table 3 illustrates the same finding in a regression framework using OLS and random-effects (RE) models. The dependent variable is the contribution of a child in the sequential public goods game. Because each of the 129 children faces this decision four times, we observe 516 decisions. Column (1) shows that contributions significantly increase in the number of other children who contribute. Columns (3)–(5) display that conditional contributions are significantly lower if no other child contributes (SM0) and significantly higher if all other players contribute (SM3). In column (5), we drop all children who did not immediately and fully understand the auditory instructions. This does not change results.

—Table 3 about here—

Result 3 – Individual heterogeneity in conditional cooperation: *Heterogeneity in cooperation between children can be reasonably described by distinct types. Depending on classification, between one third and one half of the children are conditional cooperators.*

While aggregate behavior follows a clear pattern, the question arises to what extent this pattern holds once investigating heterogeneity across individuals. We therefore now look at how individual contribution levels react to contributions of others. Figure 3 breaks this down to show all cooperation patterns existing in our data. Out of the 129 children, 32 can be classified as free-riders (first row) and 8 as pure altruists (second row).[28] The third row shows patterns for which a child's own contribution is monotonously increasing in the number of other players who contribute. We classify these children as explicit conditional cooperators. The fourth row shows behavioral patterns that fit the conditional cooperation definition of Fischbacher et al. (2001), i.e., a positive rank correlation of own and others' contributions, but violate monotonicity. We classify these children as potential conditional cooperators.[29] Finally, the last two rows show the remaining behavioral patterns, including "hump shaped" types.

—Figure 3 about here—

Figure 4 depicts the corresponding shares of types in our data.[30] Overall, the distribution of types by and large matches relative frequencies found with adult subjects (Fischbacher et al. 2001, Herrmann and Thöni 2009, Kocher et al. 2008, Kurzban and Houser 2005, Muller et al. 2008, Rustagi et al. 2010).

—Figure 4 about here—

[27] In the online appendix, we show that average conditional contributions are not statistically different if one (SM1, average contribution: 2.17) or two (SM2, average contribution: 1.82) of the other children contribute (see Figure A.2). We therefore pool them in the main text to keep the presentation as simple as possible.

[28] Free-riders never contribute (independently of the other players' actions), while altruists always contribute.

[29] While our definition matches the one by Fischbacher et al. (2001), it is less appropriate in our case, as the number of scenarios participants could condition on is much lower. Hence, a positive correlation coefficient is not as informative in our setting as it is in theirs.

[30] Table A.3 in the online appendix shows that conditional cooperators are distributed across all seven classes.

Let us summarize the results on the conditionality of cooperation up to this point in relation to the literature. We replicate the finding by Cipriani et al. (2013) that children do not react systematically to outcomes of previous rounds of the unconditional PGG, which they interpret as absence of conditional cooperation. In turn, once properly implementing the strategy method and thus giving children the explicit possibility to condition their behavior, cooperation patterns clearly match the structure of conditional cooperation (in the aggregate and for a substantial fraction of children at the individual level as well). The combination of these two results underscores the subtle difficulty and hence also the relevance of designing a functioning measure for conditional cooperation suitable for young children. While already from a game-theoretic perspective (a preference for) conditional cooperation is not identical to reacting to previous rounds of the same game,[31] our results indicate that maintaining control over individual beliefs about others' behavior is also empirically crucial. It seems likely that this is particularly the case for children as the process of belief formation is cognitively quite demanding.

Result 4 – Predictive power of survey items: *A comprehensive battery of survey items filled out by teachers as well as parents is neither able to predict unconditional nor conditional cooperative behavior.*

Table 4 provides evidence for the necessity to measure cooperation on a revealed preference basis by means of an incentivized experiment. Our OLS estimations seek to explain children's contributions with survey information on children's cooperation behavior obtained from parents and teachers.[32] It turns out that survey information is unable to predict the actual cooperative behavior of the respective children. Parental as well as teacher evaluations[33] do neither correlate with average contributions over all eight rounds (columns (1) and (2)), nor with average unconditional (columns (3) and (4)) or average conditional contributions (columns (5) and (6)). Moreover, there is no correlation of the survey measures with a child being a conditional cooperator as displayed in columns (7) and (8). The R^2 is very low for all specifications which underlines that the survey items have very limited explanatory power. Overall, these results emphasize the need to implement incentivized experimental protocols such as ours in order to really capture actual cooperative behavior in children and not to rely on simple surveys.

—Table 4 about here—

3.3 Treatment Effects

Result 5 – No ingroup bias in cooperation: *Neither conditional nor unconditional cooperation behavior is affected by differences in group composition along the lines of group identities.*

Participants in the *ingroup treatment* played the PGG with three other children from their animal group (elephants or turtles), while participants in the *outgroup treatment* played together with one participant from the same animal group and two participants from the other group. Table 5 presents the results of

[31] As only one of the games is paid out, cooperating conditional on past behavior can induce entirely different final payoffs as compared to cooperation conditionally on contemporaneous behavior.

[32] See online appendix Section C.3.

[33] We use twelve survey items from the parent questionnaire (see Table B.4 in the online appendix) to calculate a "cooperation score" and asked teachers to evaluate children' s social behavior on a scale from 1 ("rather low social behavior") to 10 ("very distinct social behavior").

OLS estimations which use individual contributions as the dependent variable. The explanatory variable of main interest is a dummy that takes the value of 1 for those participants in the outgroup treatment. Columns (1) and (2) reveal that overall contributions across all eight rounds do not differ in a statistically significant magnitude between the treatments.[34] The same result turns out once we limit our analysis to the first round which could not be affected by any previous experience. Columns (5) to (8) support the absence of an ingroup bias for both a child's unconditional (rounds 1–4) and conditional (rounds 5–8) contributions.

The reported results indicate that cooperation behavior does not depend on differences in group composition as implemented by our treatment. In contrast, experimental results reported by Angerer et al. (2016) and Peters et al. (2004) suggest that children cooperate more with matching partners from an ingroup with which they are particularly familiar. Angerer et al. (2016) find a higher degree of cooperation for children aged 6–7 in one-shot prisoners' dilemma games if they play the game with children from their own school class compared to playing the game with children from a different class. Peters et al. (2004) find that children aged 9–16 contribute more to a public good when in groups with family members than in groups with strangers.

—Table 5 about here—

4 Discussion

Understanding of the Game

A major concern in our study is whether children really grasp the functionality of the PGG. We carefully choose features in our study design in order to teach children the according mechanism and the consequences of their choices for both their own payoff as well as their group's payoff (see Section 2.1 for details). However, for the validity of our measure it is important to know whether children indeed understand the game. While this question could also be asked in studies with adult samples (cf. Burton-Chellew et al. 2016), concerns about participants' understanding might be even more relevant in studies involving younger children.

A first comforting figure is the low share of children (11.6%) indicating to have problems with understanding the game after the instructions. Further, the fact that children indeed systematically respond to the strategy method in an arguably reasonable way also adds confidence regarding the validity of our measure. Yet, one might worry that this pattern simply reflects *imitation* rather than conditional cooperation. However, looking at the share of children contributing to the public good in each round of the strategy method, we find evidence that children's behavior is apparently not driven by imitation: Imitation would imply that children's behavior follows the majority, thereby contributing if and only if the share of children contributing is above 50 percent. Our data, however, quite clearly rejects such a discontinuous pattern. A final argument speaking against imitation is the fact that children do not seem to follow the particular

[34] In all even specifications we control for variables which were not fully balanced across treatments (see Table B.3 in the online appendix).

examples presented in the instructions.[35] Hence, we would argue that our results provide support for an overall good understanding of the game and only little support for imitation behavior.[36] Still, as in every experiment and in particular those including children, we cannot fully rule out the possibility of at least some participants not fully understanding the procedures of the game.

Group Size

A related question concerning our study design is why we choose groups of four instead of two. Having groups of two would substantially reduce the complexity of the game and potentially increase the likelihood of understanding. Nevertheless, we use groups of four because (i) we aim at measuring general cooperative behavior in groups. In real world settings (e.g., working in a team), cooperation is an activity mostly depending on choices of various other independent agents. In contrast, in a two-player version of the PGG the focus is arguably on the ability to coordinate on a socially optimal outcome in a bilateral manner. Moreover, (ii) when focusing on conditional cooperation, a setting with more than two players creates more interesting variation in choices as not only levels of cooperation can differ (intensive margin) but also the proportion of players who cooperate (extensive margin).

Ingroup Bias

We find that the degree of cooperation does not vary by group composition (ingroup vs. outgroup treatment). There are at least two potential explanations for this result: Either, children at this age do not reason along the lines of group differences at all, i.e., they are not prone to ingroup bias as opposed to adults. Alternatively, it might be that children did not perceive group differences, i.e., the treatment did not induce a sense of belonging to one of the groups. Given the data at hand, we cannot distinguish between these two potential explanations. However, contrary to our findings, the results reported by Angerer et al. (2016) and Peters et al. (2004) show that children contribute more when they play with children of their own class or with family members, i.e., a group definition along the line of "natural" or "real" groups. Their results suggest that children are, in principal, prone to ingroup bias. As described in Section 2.3, our group treatment was quite strong compared to the experimental designs of Angerer et al. (2016) and Peters et al. (2004).[37] Although we do not have data that allows to perform any sort of manipulation check in this regard, we would therefore rule out that our treatment was simply too weak. Rather, it seems that the distinction between the behavioral effects of naturally induced groups (like classes and families) are qualitatively different from groups broadly and spontaneously defined according to the minimal group paradigm (Tajfel 1970), as implemented in our study. This interpretation of our result also speaks to evidence from adult subjects that indeed shows that this distinction can matter a great deal for behavioral outcomes (Goette et al. 2012). It appears an interesting potential avenue for future research to more closely investigate such potential differences in the processes of group formation across age.

[35] In the final example—before children made their own decisions—two other players contribute 3 but the modal choice in the first round was contributing 2, while only 19.38% contribute 3 as the majority of children in the final example did.

[36] This view is further reassured by the fact that we replicate findings of previous studies, e.g., Cipriani et al. (2013) with respect to the irrelevance of average contributions in previous rounds, or Fischbacher et al. (2001) with respect to the distribution of types.

[37] Angerer et al. (2016), for example, simply announce that children will be playing against a child of a different class whereas our procedures of distributing animal badges, bringing children to the experimenter room in distinguished groups and displaying the animals throughout the online public goods game was quite extensive.

Determinants of Cooperative Behavior

There is evidence that preferences for prosocial behavior—a prerequisite for cooperation behavior—are not only the product of innate psychology but reflect norms of a society (Henrich et al. 2010) that are socially transmitted from one generation to the next. Yet, it remains an open question how and at what age norms, such as a norm to (conditionally) cooperate, develop in children. Our study, which was conducted in the third month of the first grade, sheds light on this question by showing that at least for a substantial fraction of the children under study, the development of norms—even norms that are as complex as a norm for *conditional* cooperation—starts already before children get in touch with the schooling system or in the first months of their exposure to the schooling system. In line with recent papers such as Raby et al. (2015), this is further evidence underlining the importance of the early years for the development of social behavior. The fact that the extent to which prosocial behavior is transmitted genetically is rather low as compared to shared and unshared environmental effects (see, e.g., Ebstein et al. (2010) who reviews a number of studies with mono- and dizygotic twins on this issue) serves as a further piece of evidence that underlines the importance of early family environments as well as institutions such as child care centers, preschool, and the early school years for the development of cooperation behavior.

5 Summary and Conclusion

The present study takes an important step towards establishing a new instrument to measure cooperative behavior in children, specifically measuring conditional cooperation using the strategy method. We carefully design our instrument to overcome core problems with existing measurement methods, namely issues with the complexity of payoff calculations, children's limited attention span, and their inability to read and write. Moreover, we use a very simple and straightforward incentive structure to enable the elicitation of children's true underlying preferences for cooperation in a public goods game.

Our results demonstrate substantial levels of cooperation as well as a clear pattern of conditional cooperation amongst children in the first grade in Germany. We report alignment of our results to previous studies with respect to levels of unconditional cooperation (and the irrelevance of average levels of contribution in previous rounds) and the distribution of types in cooperative behavior (i.e., free-riders, conditional cooperators, etc.) which has so far only been investigated in adolescent and adult populations. Also, we argue that limited understanding or imitation behavior cannot (fully) explain our findings. Finally, a treatment eliciting ingroup bias demonstrates that—in our setup—cooperation behavior is not affected by differences in group composition.

We see potential for further research across at least two dimensions. First, it seems relevant to apply our measure in other settings and with other subjects in order to hopefully confirm the findings documented here.[38] This could and should also explicitly include validation studies that correlate behavioral measures of cooperation with field data in the spirit of Rustagi et al. (2010). Second, our study indicates that conditional cooperation behavior already develops at a young age. Thus, our measure of conditional

[38]The oTree code from the software used in the present study is provided online so that researchers can replicate and further extend our findings.

cooperation should be used for studying the following questions: (i) How does conditional cooperation (and, more generally, strategic interaction with others) develop over age? (ii) What are the main determinants of cooperative behavior? If norms are important, how do these norms evolve? And finally, (iii) to which extent is cooperative behavior shaped by the design of our (educational) institutions and by peer group composition?

References

Alencar, A. I., J. de Oliveira Siqueira, and M. E. Yamamoto (2008). Does group size matter? cheating and cooperation in brazilian school children. *Evolution and Human Behavior 29*(1), 42–48.

Angerer, S., D. Glätzle-Rützler, P. Lergetporer, and M. Sutter (2015). 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 (2016). Cooperation and discrimination within and across language borders: Evidence from children in a bilingual city. *European Economic Review 90*, 254 – 264.

Axelrod, R. and W. D. Hamilton (1981). The evolution of cooperation. *Science 211*(4489), 1390–1396.

Bowles, S. and H. Gintis (2011). *A cooperative species: Human reciprocity and its evolution*. Princeton University Press.

Bruhin, A., E. Fehr, and D. Schunk (forthcoming). The Many Faces of Human Sociality – Uncovering the Distribution and Stability of Social Preferences. *Journal of the European Economic Association*.

Burton-Chellew, M. N., C. E. Mouden, and S. A. West (2016). Conditional cooperation and confusion in public-goods experiments. *Proceedings of the National Academy of Sciences 113*(5), 1291–1296.

Cárdenas, J. C., A. Dreber, E. von Essen, and E. Ranehill (2015). Cooperativeness and competitiveness in children. *Journal of Behavioral and Experimental Economics 59*, 32–41.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics 14*(1), 47–83.

Chen, D. L., M. Schonger, and C. Wickens (2016). otreean open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance 9*, 88–97.

Chen, Y. and S. X. Li (2009). Group identity and social preferences. *American Economic Review 99*(1), 431–57.

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.

Currie, C., M. Molcho, W. Boyce, B. Holstein, T. Torsheim, and M. Richter (2008). Researching health inequalities in adolescents: The development of the Health Behaviour in School-Aged Children (HBSC) Family Affluence Scale. *Social Science & Medicine 66*(6), 1429–1436.

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.

Ebstein, R. P., S. Israel, S. H. Chew, S. Zhong, and A. Knafo (2010). Genetics of Human Social Behavior. *Neuron 65*(6), 831 – 844.

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

Fehr, E. and U. Fischbacher (2003). The nature of human altruism. *Nature 425*(6960), 785–791.

Fehr, E. and A. Leibbrandt (2011). A field study on cooperativeness and impatience in the tragedy of the commons. *Journal of Public Economics 95*(9), 1144 – 1155.

Fischbacher, U. and S. Gächter (2010). Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments. *The American Economic Review 100*(1), 541–556.

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

Fischbacher, U., S. Gächter, and S. Quercia (2012). The behavioral validity of the strategy method in public good experiments. *Journal of Economic Psychology 33*(4), 897 – 913.

Goette, L., D. Huffman, and S. Meier (2012). The impact of social ties on group interactions: Evidence from minimal groups and randomly assigned real groups. *American Economic Journal: Microeconomics 4*(1), 101–15.

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.

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

Hardin, G. (1968). The Tragedy of the Commons. *Science 162*(3859), 1243–1248.

Henrich, J., J. Ensminger, R. McElreath, A. Barr, C. Barrett, A. Bolyanatz, J. C. Cardenas, M. Gurven, E. Gwako, N. Henrich, C. Lesorogol, F. Marlowe, D. Tracer, and J. Ziker (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science 327*(5972), 1480–1484.

Herrmann, B. and C. Thöni (2009). Measuring conditional cooperation: a replication study in russia. *Experimental Economics 12*(1), 87–92.

Hett, F., M. Kröll, and M. Mechtel (2018). Choosing Who You Are: The Structure and Behavioral Effects of Revealed Identification Preferences. *Working Paper*.

Keil, J., A. Michel, F. Sticca, K. Leipold, A. M. Klein, S. Sierau, K. von Klitzing, and L. O. White (2017). The pizzagame: a virtual public goods game to assess cooperative behavior in children and adolescents. *Behavior research methods 49*(4), 1432–1443.

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

Kosfeld, M. (2017). The role of leaders in inducing and maintaining cooperation: The cc strategy. *Unpublished Manuscript*.

Kosse, F., T. Deckers, H. Schildberg-Hörisch, and A. Falk (2017). The formation of prosociality: causal evidence on the role of social environment. DICE Discussion Papers 266, Heinrich-Heine-Universität Düsseldorf, Düsseldorf Institute for Competition Economics (DICE).

Kurzban, R. and D. Houser (2005). Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences of the United States of America 102*(5), 1803–1807.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. Kagel and A. Roth (Eds.), *The Handbook of Experimental Economics*. Princeton: Princeton University Press.

Levitt, S., J. List, S. Neckermann, and S. Sadoff (2016). The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance. *American Economic Journal: Economic Policy 8*(4), 183–219.

Muller, L., M. Sefton, R. Steinberg, and L. Vesterlund (2008). Strategic behavior and learning in repeated voluntary contribution experiments. *Journal of Economic Behavior & Organization 67*(3), 782 – 793.

Olson, M. (1965). *The Logic of Collective Action*. Harvard University Press.

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge ; New York: Cambridge University Press.

Ostrom, E., J. Burger, C. B. Field, R. B. Norgaard, and D. Policansky (1999). Revisiting the commons: Local lessons, global challenges. *Science 284*(5412), 278–282.

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.

Raby, K. L., G. I. Roisman, R. C. Fraley, and J. A. Simpson (2015). The enduring predictive significance of early maternal sensitivity: Social and academic competence through age 32 years. *Child Development 86*(3), 695–708.

Rustagi, D., S. Engel, and M. Kosfeld (2010). Conditional Cooperation and Costly Monitoring Explain Success in Forest Commons Management. *Science 330*(6006), 961–965.

Sally, D. and E. Hill (2006). The development of interpersonal strategy: Autism, theory-of-mind, cooperation and fairness. *Journal of Economic Psychology 27*(1), 73 – 97.

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

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. *Working paper*.

Selten, R. (1967). Die strategiemethode zur erforschung des eingeschränkt rationalen verhaltens im rahmen eines oligopolexperimentes. In H. Sauermann (Ed.), *Beiträge zur experimentellen Wirtschaftsforschung. J.C.B. Mohr (Paul Siebeck)*, pp. 136–168. Tübingen.

Smith, V. L. (1976). Experimental Economics: Induced Value Theory. *The American Economic Review 66*(2), 274–279.

Tajfel, H. (1970). Experiments in Intergroup Discrimination. *Scientific American 223*(5), 96–103.

Tajfel, H. and J. C. Turner (1986). The social identity theory of intergroup behavior. In S. Worchel and W. Austin (Eds.), *The Social Psychology of Intergroup Relations*, pp. 7–24. Chicago: Nelson-Hall.

Vogelsang, M., K. Jensen, S. Kirschner, C. Tennie, and M. Tomasello (2014). Preschoolers are sensitive to free riding in a public goods game. *Frontiers in Psychology 5*.

Yu, J., L. Zhu, and A. M. Leslie (2016). Children's sharing behavior in mini-dictator games: The role of in-group favoritism and theory of mind. *Child Development 87*(6), 1747–1757.

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics 6*(3), 299–310.

Tables and Figures

Tables

Table 1: Overview of the PGG rounds

Round	**1** (R1)	**2** (R2)	**3** (R3)	**4** (R4)	**5** (SM0)	**6** (SM1)	**7** (SM2)	**8** (SM3)
A. PGG variant								
simultaneous, unconditional	x	x	x	o				
sequential, conditional (strategy method)					o	o	o	o
B. "Others" contribute (strategy method)								
# of Contributors:					0	1	2	3

Notes: **A** shows in which round children played the unconditional PGG (x) and the strategy method (o). *Action space*: x = $c_i \in [0, 1, 2, 3, 4, 5]$; **o** = $c_i \in [0, 5]$. **B** shows for each round of the strategy method how many of the other players contributed to the public good.

Table 2: Unconditional Contributions Across Rounds

	Contr. R2	Contr. R3		Contr. R4		
	(1)	(2)	(3)	(4)	(5)	(6)
Own contribution round R1	0.573***	0.657***	0.644***	0.706***	0.692***	0.751***
	(0.12)	(0.10)	(0.10)	(0.15)	(0.14)	(0.18)
Avg. contr. group members R1	0.015		-0.132		0.116	-0.051
	(0.13)		(0.13)		(0.26)	(0.28)
Avg. contr. group members R2		-0.252	-0.189		-0.537	-0.389
		(0.20)	(0.20)		(0.32)	(0.34)
Avg. contr. group members R3				-0.112	-0.000	0.066
				(0.28)	(0.30)	(0.29)
Control variables	yes	yes	yes	yes	yes	yes
Session fixed-effects	yes	yes	yes	yes	yes	yes
Understanding only	no	no	no	no	no	yes
Observations	118	118	118	118	118	106
R^2	0.404	0.452	0.454	0.328	0.346	0.370

Notes: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. The dependent variable is child *i*'s unconditional contribution in the public goods game (PGG) in round 2 (column 1), round 3 (columns 2 and 3), and round 4 (columns 4–6). *Own contribution round 1* measures *i*'s contribution to the PGG in round 1. *Avg. contr. group members Rj* represents the average contribution of the 3 other group members in round *j*. The vector of *control variables* includes a dummy variable that takes the value of 1 when child *i* is *female* and is 0 otherwise; a variable that captures *i*'s *number of siblings*; the dummy variable *Sports* that takes the value of 1 if *i* does some sports additional to the regular school sports classes and is 0 otherwise; the (*extended*) *FAS score* which is measured by five survey items: *Does your family own a car?* (No = 0, Yes, One = 1, Yes, two or more = 2); *Does your child has a bedroom for her/himself?* (No = 0, Yes = 1); *During the past 12 months, how many times did you travel abroad with your family?* (Not at all = 0, Once = 1, Twice = 2, More than twice = 3); *How many computers does your family own?* (None = 0, One = 1, Two = 2, Three = 3. More than three = 4); *How many books do you have at home?* (None/a few = 0, A complete single board of a book shelf = 1, A complete book shelf = 2, Two complete book shelves = 3, More than two complete book shelves = 4); a dummy variable which takes the value of 1 if at least one of *i*'s *parents was born in a country other than Germany* and is 0 otherwise; and a dummy variable which takes the value of 1 if at least one of *i*'s *parents holds a university degree* and is 0 otherwise. In specification (6), all children who did not understand the audio instructions at the first attempt are excluded from the analysis. Standard errors clustered at the group level are in parentheses.

Table 3: Conditional Contributions

	Conditional Contribution				
	(1) POLS	(2) RE	(3) POLS	(4) RE	(5) RE
Other players' contributions	0.102***	0.102***			
	(0.01)	(0.02)			
SM0 (dummy variable)			-0.988***	-0.988***	-1.075***
			(0.20)	(0.22)	(0.24)
SM3 (dummy variable)			0.833***	0.833***	0.724***
			(0.25)	(0.22)	(0.24)
Understanding only	no	no	no	no	yes
Observations	516	516	516	516	456
R^2	0.055		0.070		
R^2 overall model		0.055		0.070	0.068

Notes: Coefficient estimates from pooled ordinary least squares estimations (columns 1 and 3) and random-effects panel estimations (columns 2, 4, and 5). * significant at 10%; ** sign. at 5%; *** sign. at 1%. The dependent variable is child *i*'s conditional contribution in each of the four rounds (rounds 5–8 overall) of the public goods game (PGG). As 129 children participated in the experiment, we end up with 516 observations. *Other players' contributions* captures the sum of other players' contributions to the public good (0 in round 5, 5 in round 6, 10 in round 7, 15 in round 8). *SM(0)* and *SM(3)* are dummy variables that take the value of 1 if 0 (3) other children contributed their complete endowment. In the regressions based on these dummy variables (columns 3–5), the reference situation is 1 or 2 other children contributing. No further control variables are included in the estimations as all variation is within-child. In specification (5), all children who did not understand the audio instructions at the first attempt are excluded from the analysis. Heteroskedasticity-robust Huber-White standard errors are in parentheses. In columns (1) and (3), standard errors are clustered at the session level.

Table 4: Predictive Power of Survey Items

	Avg. contr.		Avg. uncond. contr.		Avg. cond. contr.		Cond. coop. (d)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cooperation score	0.015		0.006		0.024		-0.008	
	(0.02)		(0.02)		(0.02)		(0.01)	
Teacher rating		0.029		0.056		0.002		-0.009
		(0.08)		(0.08)		(0.09)		(0.02)
Teacher fixed-effects	no	yes	no	yes	no	yes	no	yes
Observations	106	122	106	122	106	122	106	122
R^2	0.006	0.016	0.001	0.008	0.011	0.026	0.015	0.036

Notes: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. The dependent variable is child *i*'s average contribution over all 8 rounds of the public goods game (PGG) in columns (1) and (2), *i*'s average unconditional contribution in rounds R1–R4 of the PGG (columns 3 and 4), *i*'s average conditional contribution in rounds SM0–SM3 of the PGG (columns 5 and 6), and the dummy variable *Cond. coop. (d)* which takes the value of 1 if child *i* is classified as being a conditional cooperator (columns 7–8). The cardinal variable *Cooperation score* is the standardized average of the twelve survey items from the parent questionnaire (see Table B.4 in the appendix). The variable *Teacher rating* is based on teachers' evaluation of child *i*'s social behavior on a scale from 1 ("rather low social behavior") to 10 ("very distinct social behavior"). Teacher ratings are available also for those children for whom we do not have information from the parent questionnaire. We include teacher fixed-effects in all estimations that rely on teacher ratings to control for unobserved heterogeneity between teachers. Heteroskedasticity-robust Huber-White standard errors are in parentheses.

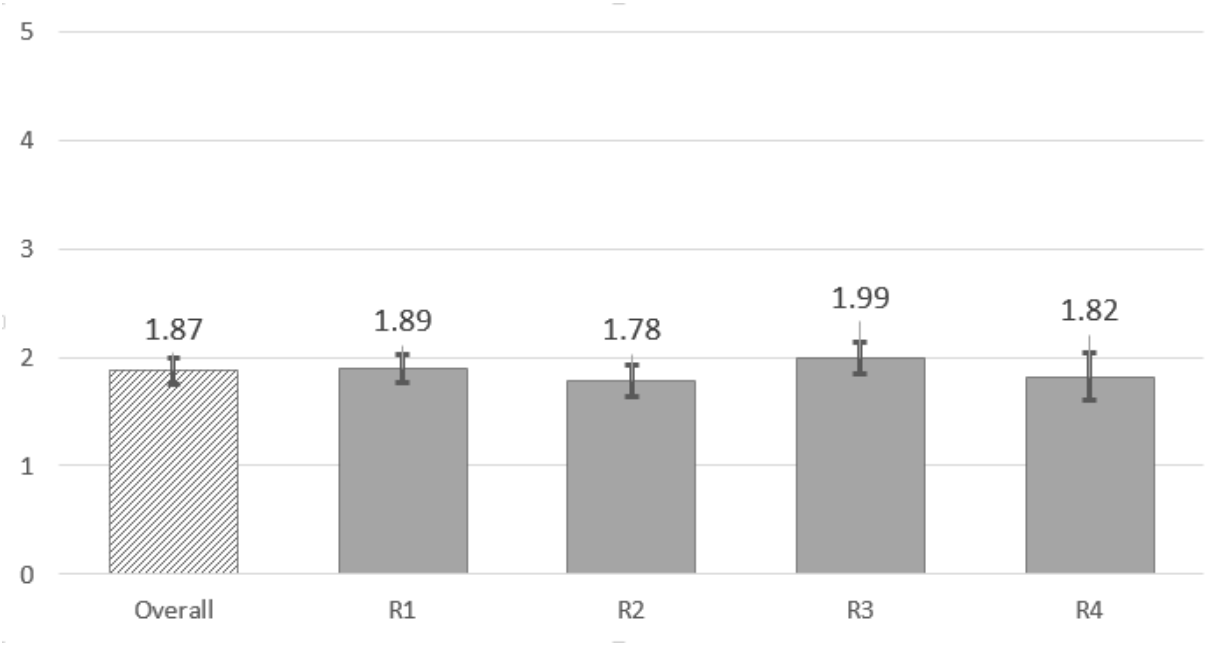
Table 5: Group Membership Effects on Contributions

	Avg. contr.		R1 contr.		Avg. uncond. contr.		Avg. cond. contr.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outgroup treatment (d)	-0.009	-0.132	0.122	0.147	0.029	-0.020	-0.047	-0.244
	(0.18)	(0.26)	(0.25)	(0.27)	(0.21)	(0.26)	(0.20)	(0.29)
Teacher rating		-0.002		0.073		0.058		-0.062
		(0.09)		(0.09)		(0.09)		(0.10)
Female (dummy)		0.047		0.331		0.034		0.061
		(0.23)		(0.29)		(0.26)		(0.25)
# of siblings		0.051		0.173		0.089		0.012
		(0.11)		(0.13)		(0.13)		(0.11)
Parent born abroad (d)		0.314		0.521*		0.354		0.274
		(0.35)		(0.27)		(0.37)		(0.39)
Further control variables	no	yes	no	yes	no	yes	no	yes
Observations	129	118	129	118	129	118	129	118
R^2	0.000	0.028	0.002	0.078	0.000	0.042	0.000	0.025

Notes: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. The dependent variable is child *i*'s average contribution over all 8 rounds of the public goods game (PGG) in columns (1) and (2), *i*'s contribution in the very first round (R1) of the PGG (columns 3 and 4), *i*'s average unconditional contribution in rounds R1–R4 of the PGG (columns 5 and 6), and *i*'s average conditional contribution in rounds SM0–SM3 of the PGG (columns 7 and 8). The explanatory variable of main interest is *Outgroup treatment (d)* which takes the value of 1 if child *i* participated in the *outgroup treatment* and is 0 otherwise (*ingroup treatment*). The table additionally reports the coefficients of all explanatory variables that are not perfectly balanced between treatment and control. This set of variables includes: the variable *Teacher rating* which is based on teachers' evaluation of child *i*'s social behavior on a scale from 1 ("rather low social behavior") to 10 ("very distinct social behavior"); the dummy variable *Female (dummy)* that takes the value of 1 when child *i* is female and is 0 otherwise; the variable *# of siblings* that captures *i*'s number of siblings; and the dummy variable *Parent born abroad (d)* which takes the value of 1 if at least one of *i*'s parents was born in a country other than Germany and is 0 otherwise. The vector of *Further control variables* includes the dummy variable *Sports* that takes the value of 1 if *i* does some sports additional to the regular school sports classes and is 0 otherwise; the (*extended*) *FAS score* which is measured by five survey items: *Does your family own a car?* (No = 0, Yes, One = 1, Yes, two or more = 2); *Does your child has a bedroom for her/himself?* (No = 0, Yes = 1); *During the past 12 months, how many times did you travel abroad with your family?* (Not at all = 0, Once = 1, Twice = 2, More than twice = 3); *How many computers does your family own?* (None = 0, One = 1, Two = 2, Three = 3. More than three = 4); *How many books do you have at home?* (None/a few = 0, A complete single board of a book shelf = 1, A complete book shelf = 2, Two complete book shelves = 3, More than two complete book shelves = 4); and a dummy variable which takes the value of 1 if at least one of *i*'s *parents holds a university degree* and is 0 otherwise. Heteroskedasticity-robust Huber-White standard errors are in parentheses.

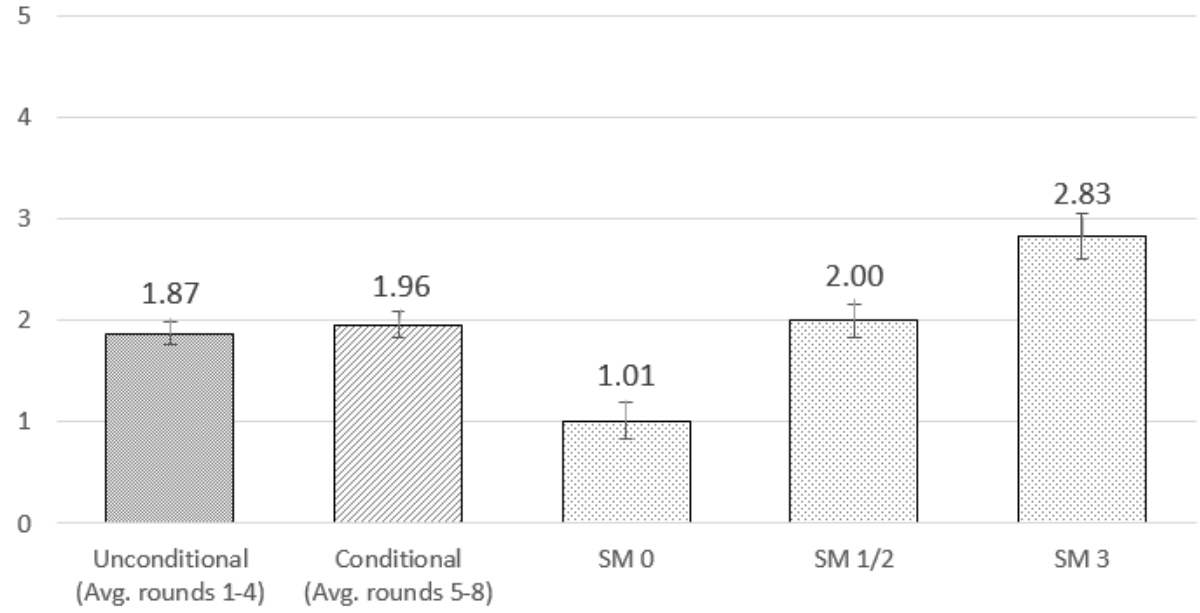
Figures

Figure 1: Unconditional Average Contributions



Notes: Average contributions to the public good in the simultaneous part of the game. Contributions are measured in coin tupels (from 0 to 5) placed in the group basket. The left-hand bar shows mean contributions over rounds 1–4, the other bars show mean contributions for each respective round. Error bars present standard errors for the respective statistic.

Figure 2: Average Contributions

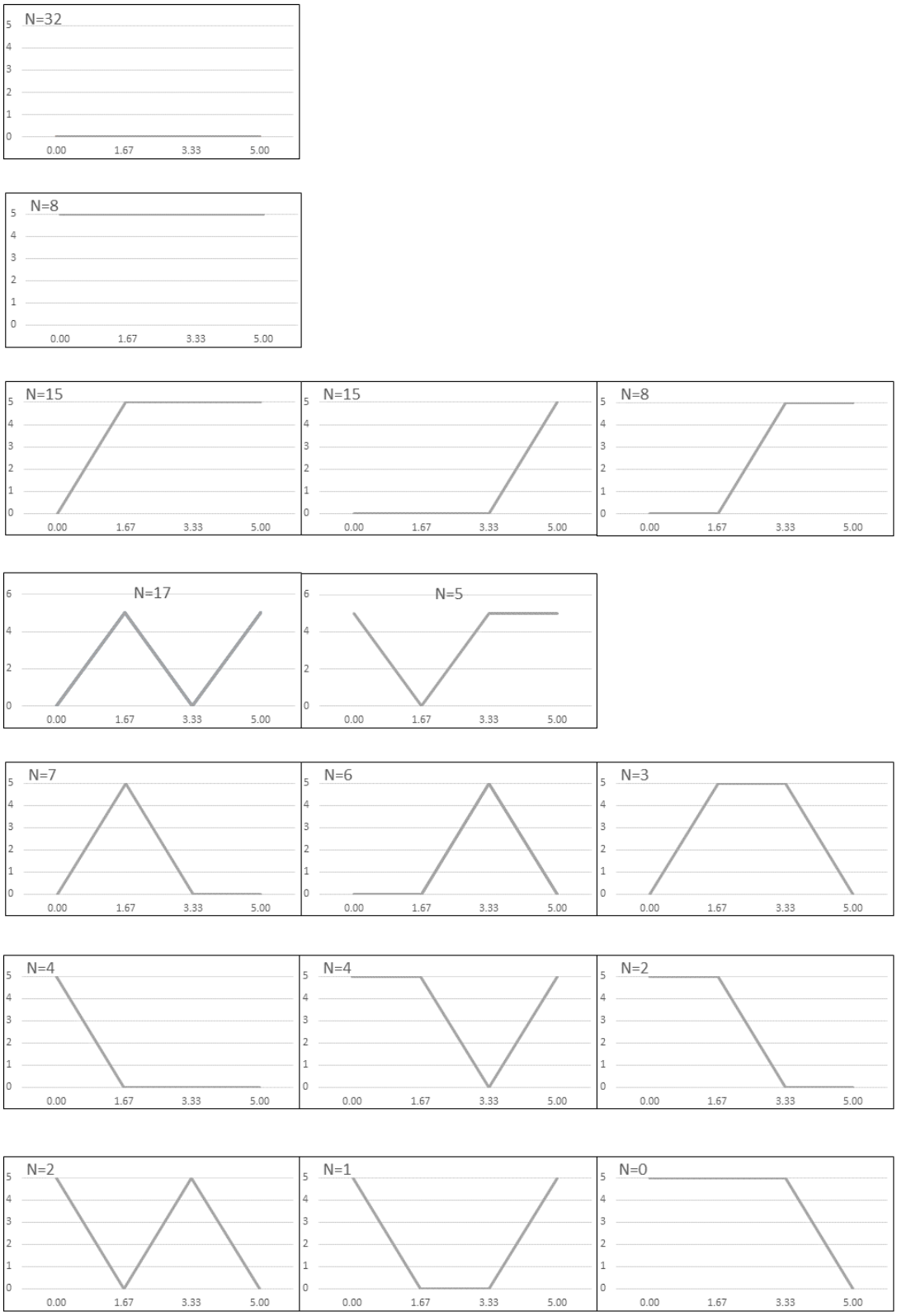


Paired t-test:

Unconditional	=	Conditional	p=0.43
SM 0	=	SM 1/2	p<0.01
SM 1/2	=	SM 3	p<0.01
SM 0	=	SM 3	p<0.01

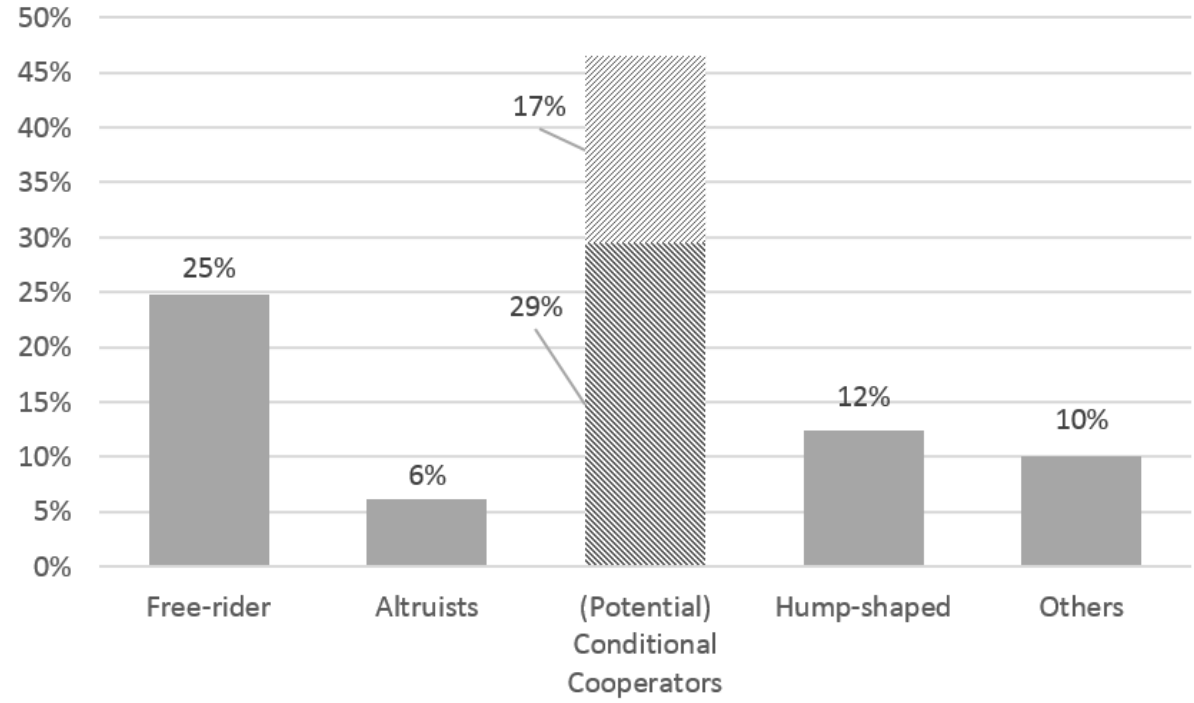
Notes: Average contributions to the public good over the eight rounds of the game. Contributions are measured in coin tupels (from 0 to 5) placed in the group basket. The left-hand bar shows mean contributions over rounds 1–4 (i.e., in the simultaneous part of the game). The next bar shows average contributions in rounds 5–8 (i.e., in the sequential part of the game). The remaining bars show mean contributions for rounds 5, 6 and 7 (pooled), and 8, respectively (refer to the online appendix, Figure A.2 for details on rounds 6 and 7). Error bars present standard errors for the respective statistic.

Figure 3: Own Contribution Level for Each Average Contribution Level of Other Group Members



Notes: Subjects are classified as follows: First row: **Free-Riders**; second row: **Altruists**; third row: **Conditional Cooperators**; fourth row: **Potential Conditional Cooperators** (ρ >0); fifth row: **Hump-Shaped**; sixth and seventh row: **Other patterns**. See Section 3.2 for detailed definitions of types.

Figure 4: Type Distribution



Notes: Explicit conditional cooperators represent 29% of the participants and potential conditional cooperators 17%. See Section 3.2 for detailed definitions of types.

Supplementary Online Appendix

A Additional Tables and Figures

A.1 Detailed Results on Rounds 1–8

In Table A.1, we provide summary statistics for each single round played in the public goods game. While average contributions in rounds 1–4 remain fairly stable (around 1.9 out of 5), contributions in the four rounds based on the strategy method (SM0–SM3) clearly vary depending on the other players' contributions. Note that already in R4 the action space is limited to choosing only 0 or 5 tuples for contribution—which does not affect the average contribution level as compared to rounds R1–R3.

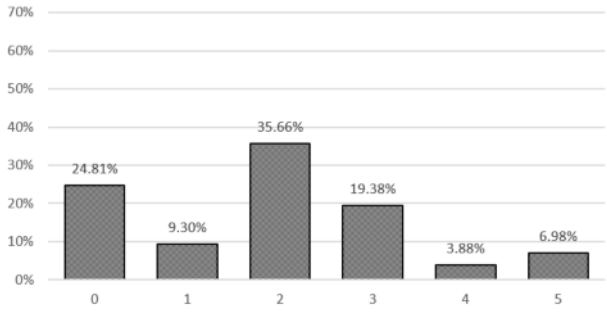
Table A.1: Summary Statistics: Contributions

	Mean	**Std. Dev.**	**Min.**	**Max.**	**N**
Avg. contribution	1.915	2.146	0	5	1032
Avg. unconditional contribution (R1–R4)	1.872	1.803	0	5	516
Avg. unconditional contribution (R1–R3)	1.889	1.549	0	5	387
Avg. conditional contribution (SM0–SM3)	1.957	2.443	0	5	516
Avg. contribution R1	1.891	1.437	0	5	129
Avg. contribution R2	1.783	1.566	0	5	129
Avg. contribution R3	1.992	1.642	0	5	129
Avg. contribution R4	1.822	2.416	0	5	129
Avg. contribution SM0	1.008	2.014	0	5	129
Avg. contribution SM1	2.171	2.488	0	5	129
Avg. contribution SM2	1.822	2.416	0	5	129
Avg. contribution SM3	2.829	2.488	0	5	129

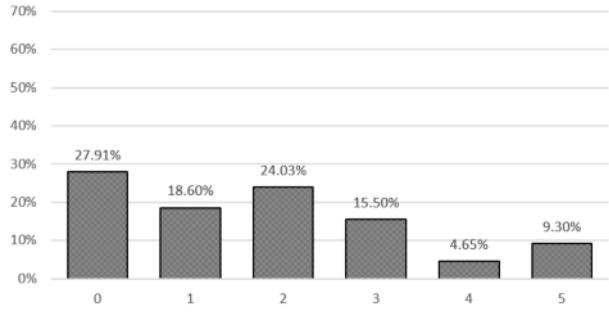
Notes: For a definition of the public goods game's rounds (R1–R4, SM0–SM3), see Table 1.

To add to the information from Table A.1, we provide Figure A.1. The distribution of contributions remains fairly stable across rounds 1–3. However, while the mode in the first round is to contribute 2, in rounds 2 and 3 the mode is contributing 0. Still, in all rounds the majority of children contributes a positive amount to the public good. In round 4, the choice set was limited to contributing 0 or 5. As reported above, the average contribution level remains unaffected by this variation in the action space.

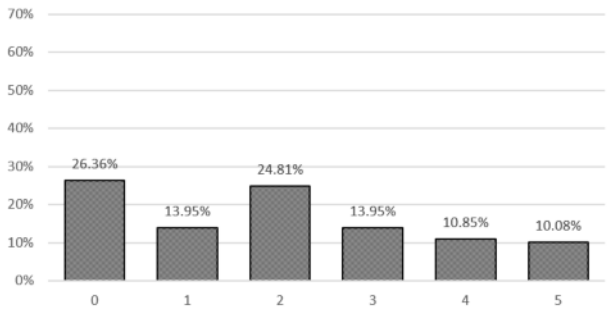
Figure A.1: Unconditional Contributions (Rounds 1–4)



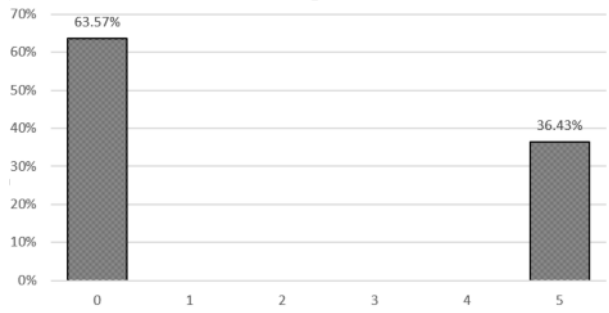
(a) Distribution of Contributions R1



(b) Distribution of Contributions R2



(c) Distribution of Contributions R3



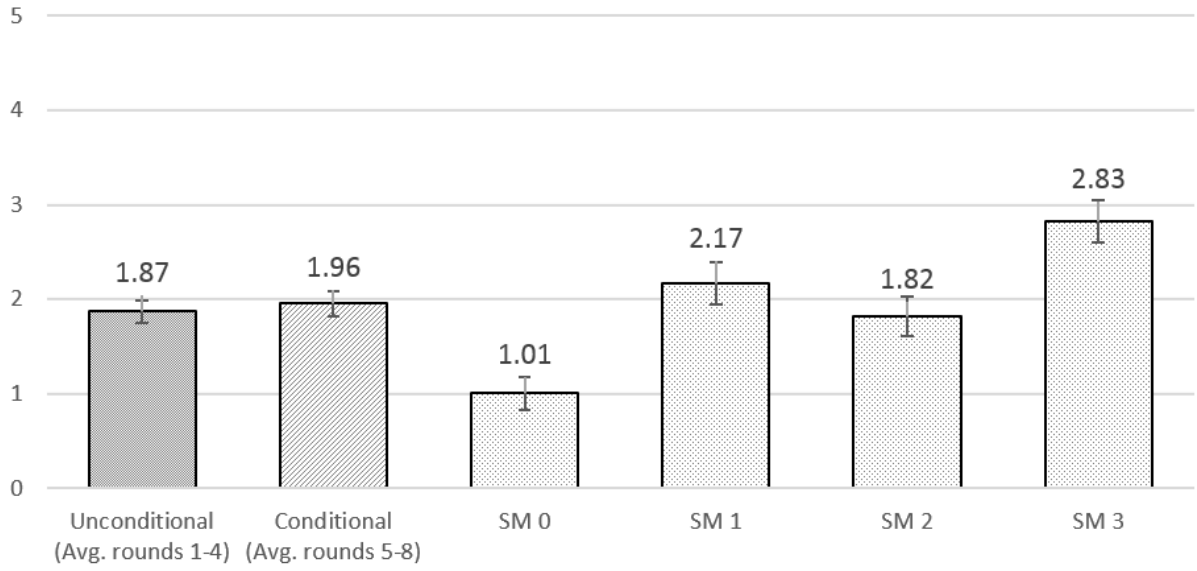
(d) Distribution of Contributions R4

Notes: Detailed distributions of contributions to the public good in rounds 1–4. Percentages refer to the share of children that contributed 0, 1, 2, 3, 4, or 5 coin tupels to the group basket, respectively. In round 4, the choice set was restricted to 0 or 5 coin tupels.

A.2 Comparing Conditional and Unconditional Cooperation

Figure A.2 adds to the analysis provided in Figure 2 by separating the average contribution levels in rounds SM1 and SM2 (i.e., one or two other players contribute to the public good). Average contribution in SM1 amounts to 2.17, average contribution in SM2 is equal to 1.82. A paired t-test reveals no significant difference between SM1 and SM2 ($p = 0.21$), while t-tests comparing SM0 with SM1 and SM2 with SM3 both reveal significant differences in average contribution levels (p-values <0.01). We therefore pool SM1 and SM2 in the main analysis.

Figure A.2: Average Contributions



Paired t-test:

SM 0	=	SM 1	p<0.01
SM 1	=	SM 2	p=0.21
SM 2	=	SM 3	p<0.01
SM 0	=	SM 2	p<0.01
SM 0	=	SM 3	p<0.01
SM 1	=	SM 3	p<0.01

Notes: Average contributions to the public good over the eight rounds of the game. Contributions are measured in coin tupels (from 0 to 5) placed in the group basket. The left-hand bar shows mean contributions over rounds 1–4 (i.e., in the simultaneous part of the game). The next bar shows average contributions in rounds 5–8 (i.e., in the sequential part of the game). The remaining bars show mean contributions for rounds 5, 6, 7, and 8, respectively. Error bars present standard errors for the respective statistic.

To support the notion that children's choice patterns display a certain stability or consistency, we present Table A.2. On an individual level, average contributions when using the strategy method (i.e., rounds 5–8) are strongly correlated with average contribution levels in the unconditional rounds (i.e., rounds 1–4), regardless of whether we include control variables or not. The same applies when looking at the correlation of a participant's contribution in round 4 (in which, like in the subsequent rounds based on strategy method, the choice set was limited to 0 or 5) with average contributions in the strategy method rounds.

Table A.2: Average Conditional Contributions

	Average conditional contribution			
	(1)	(2)	(3)	(4)
Avg. uncond. contr.	0.739***	0.724***		
	(0.08)	(0.09)		
Contribution R4			0.338***	0.331***
			(0.05)	(0.06)
Controls	no	yes	no	yes
Observations	129	118	129	118
R^2	0.420	0.418	0.285	0.270

Notes: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. The dependent variable is child *i*'s average conditional contribution in the public goods game (PGG) in rounds 5–8 (SM0–SM3). *Avg. uncond. contr.* represents the average unconditional contribution of child *i* in rounds 1–4 (R1–R4). *Contribution R4* is child *i*'s contribution in round 4. The vector of *control variables* includes a dummy variable that takes the value of 1 when child *i* is *female* and is 0 otherwise; a variable that captures *i*'s *number of siblings*; the dummy variable *Sports* that takes the value of 1 if *i* does some sports additional to the regular school sports classes and is 0 otherwise; the (*extended*) *FAS score* which is measured by five survey items: *Does your family own a car?* (No = 0, Yes, One = 1, Yes, two or more = 2); *Does your child has a bedroom for her/himself?* (No = 0, Yes = 1); *During the past 12 months, how many times did you travel abroad with your family?* (Not at all = 0, Once = 1, Twice = 2, More than twice = 3); *How many computers does your family own?* (None = 0, One = 1, Two = 2, Three = 3. More than three = 4); *How many books do you have at home?* (None/a few = 0, A complete single board of a book shelf = 1, A complete book shelf = 2, Two complete book shelves = 3, More than two complete book shelves = 4); a dummy variable which takes the value of 1 if at least one of *i*'s *parents was born in a country other than Germany* and is 0 otherwise; and a dummy variable which takes the value of 1 if at least one of *i*'s *parents holds a university degree* and is 0 otherwise. In specification (6), all children who did not understand the audio instructions at the first attempt are excluded from the analysis. Standard errors clustered at the group level are in brackets.

A.3 Individual Cooperation Patterns in the Sequential PGG

In Table A.3 we present the distribution of conditional cooperators across classes in our sample. While there is clearly heterogeneity across classes, conditional cooperators are fairly distributed across classes and the behavioral pattern we observe is not driven by a single or very few classes.

Table A.3: Distribution of Conditional Cooperators over Classes

Class-ID	Conditional Cooperators		N
	No	Yes	
1	8	6	14
2	9	9	18
3	13	7	20
4	16	4	20
5	18	4	22
6	14	5	19
7	13	3	16
N	91	38	129

Notes: Conditional Cooperators defined by own contribution increases monotonously in the number of other contributors.

B Procedural Details

B.1 Sample Descriptives

Table B.1: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
School A (dummy)	0.442	0.499	0	1	129
Female (dummy)	0.508	0.502	0	1	120
# of siblings	1.358	1.143	0	9	120
Sports (dummy)	0.783	0.414	0	1	120
FAS Score	9.898	2.158	3	14	118
Parent born abroad (dummy)	0.176	0.383	0	1	119
Parent university (dummy)	0.412	0.494	0	1	119
Cooperation score	0.137	7.018	-26.037	11.351	106
Teacher rating	6.607	1.775	2	10	122
Understanding (dummy)	0.884	0.322	0	1	129

Notes: *School A (dummy)* is a dummy variable that takes the value of 1 if child *i* is from the first school in our data and is 0 if *i* is from the second school. The dummy variable *Female (dummy)* takes the value of 1 when child *i* is female and is 0 otherwise. The variable *# of siblings* captures *i*'s number of siblings. *Sports (dummy)* is 1 if *i* does some sports additional to the regular school sports classes and 0 otherwise. The (*extended*) *FAS score* (FAS, Currie et al. 2008) to capture the socio-economic position of the childrens' households is measured by five survey items: *Does your family own a car?* (No = 0, Yes, One = 1, Yes, two or more = 2); *Does your child has a bedroom for her/himself?* (No = 0, Yes = 1); *During the past 12 months, how many times did you travel abroad with your family?* (Not at all = 0, Once = 1, Twice = 2, More than twice = 3); *How many computers does your family own?* (None = 0, One = 1, Two = 2, Three = 3. More than three = 4); *How many books do you have at home?* (None/a few = 0, A complete single board of a book shelf = 1, A complete book shelf = 2, Two complete book shelves = 3, More than two complete book shelves = 4). *Parent born abroad (dummy)* takes the value of 1 if at least one of *i*'s parents was born in a country other than Germany and is 0 otherwise. *Parent university (dummy)* is 1 if at least one of *i*'s parents holds a university degree and 0 otherwise. The cardinal variable *Cooperation score* is the standardized average of the twelve survey items from the parent questionnaire (see Table B.4 in the appendix). The variable *Teacher rating* is based on teachers' evaluation of child *i*'s social behavior on a scale from 1 ("rather low social behavior") to 10 ("very distinct social behavior"). The cardinal *Cooperation Score* is the standardized average of the 12 survey items given in table B.4 in the appendix. The variable *Teacher rating* is based on teachers' evaluation of child *i*'s social behavior on a scale from 1 ("rather low social behavior") to 10 ("very distinct social behavior"). *Understanding (dummy)* is 1 if child *i* understood the audio instructions at the first attempt and 0 otherwise. The sample is unbalanced for the following reasons: 129 children took part in the experiment; 122 parents handed in the parent questionnaire, but item non-response reduces the number of observations per variable to 106–120; 122 teachers provided us with the *Teacher rating* information on those children for whom we received back the parent questionnaire.

The children participating in the study were recruited from two public elementary schools in the city of Bonn in the federal state of North Rhine-Westphalia (NRW), Germany. The study was conducted in seven classes. Parents' consent is a necessary legal prerequisite in NRW to conduct scientific studies with under-aged children and 144 of 174 parents (82.76%) gave their consent. Out of the 144 potential participants, 7 children were sick on the testing day, 3 left the classroom during the session, and 5 children of the last class could not be tested due to a lack of time, leaving us with 129 children. All recruited classes were classes of grade 1 with a class size of on average 24.86 children aged about 5–7 years. Table B.1 summarizes important background characteristics of the children and their families, based on a parent questionnaire.

B.2 Randomization

Randomization was performed within class. First, children were randomized into groups of four. These groups were than randomized either into the ingroup or the outgroup treatment. Thereafter, children were assigned to the animal groups. All four children of a group in the ingroup treatment were randomized into the turtle or the elephant group. Children within each outgroup treatment were randomized into the turtle or the elephant group such that the outgroup treatment always consisted of two turtles and two elephants. Children in the ingroup vs. outgroup treatment are equally distributed over being turtles vs. elephants, compare Table B.2. Table B.3 presents randomization checks for observable characteristics. Differences of means between the control and the treatment group are mainly insignificant. The variables *# of siblings*, *Parent born abroad (dummy)*, and *Teacher rating* are not perfectly balanced between the treatments, which is not surprising given the sample size. Controlling for these variables in our regressions analyzing the treatment effect does not alter our results.

Table B.2: Number of Participants per Treatment and Animal Group Assignment

	Elephant	Turtle	Total
Ingroup	29	36	65
Outgroup	34	30	64
Total	63	66	129

Table B.3: Group Identity Treatment: Differences between Ingroup and Outgroup

	(1) Ingroup mean	(2) Outgroup mean	(3) Difference	(3) p-value
Female (dummy)	0.52	0.50	0.02	(0.86)
# of siblings	1.17	1.55	-0.38*	(0.07)
Sports (dummy)	0.75	0.82	-0.07	(0.38)
FAS Score	9.59	10.20	-0.61	(0.13)
Parent born abroad (dummy)	0.25	0.10	0.15**	(0.03)
Parent university (dummy)	0.36	0.47	-0.11	(0.22)
Cooperation score	0.65	-0.42	1.07	(0.43)
Teacher rating	6.97	6.26	0.71**	(0.03)
Understanding (dummy)	0.86	0.91	-0.04	(0.43)
Observations	65	64	129	

Notes: The dummy variable *Female (dummy)* takes the value of 1 when child *i* is female and is 0 otherwise. The variable *# of siblings* captures *i*'s number of siblings. *Sports (dummy)* is 1 if *i* does some sports additional to the regular school sports classes and 0 otherwise. The (*extended*) *FAS score* is measured by five survey items: *Does your family own a car?* (No = 0, Yes, One = 1, Yes, two or more = 2); *Does your child has a bedroom for her/himself?* (No = 0, Yes = 1); *During the past 12 months, how many times did you travel abroad with your family?* (Not at all = 0, Once = 1, Twice = 2, More than twice = 3); *How many computers does your family own?* (None = 0, One = 1, Two = 2, Three = 3. More than three = 4); *How many books do you have at home?* (None/a few = 0, A complete single board of a book shelf = 1, A complete book shelf = 2, Two complete book shelves = 3, More than two complete book shelves = 4). *Parent born abroad (dummy)* takes the value of 1 if at least one of *i*'s parents was born in a country other than Germany and is 0 otherwise. *Parent university (dummy)* is 1 if at least one of *i*'s parents holds a university degree and 0 otherwise. The cardinal variable *Cooperation score* is the standardized average of the twelve survey items from the parent questionnaire (see Table B.4 in the appendix). The variable *Teacher rating* is based on teachers' evaluation of child *i*'s social behavior on a scale from 1 ("rather low social behavior") to 10 ("very distinct social behavior"). The cardinal *Cooperation Score* is the standardized average of the 12 survey items given in table B.4 in the appendix. The variable *Teacher rating* is based on teachers' evaluation of child *i*'s social behavior on a scale from 1 ("rather low social behavior") to 10 ("very distinct social behavior"). *Understanding (dummy)* is 1 if child *i* understood the audio instructions at the first attempt and 0 otherwise. p-values provided for t-tests of differences between variable means across treatments.

B.3 Procedures on a Testing Day

The study was conducted on four days in November 2017. In each session, up to 12 children took part simultaneously. Classes were therefore divided into two groups, with the first half participating in the first session and the second half in the second session. First, the game is explained on the basis of the offline version—in a board game manner. Thereafter, children play the online version of the game which is relevant for children's payoffs. This section focuses on procedural details, see Section 2.1 for a description of the basic experimental design.

During the offline explanation, children are told that they will all play a game in groups of four with three other children from their class but that interaction will be anonymous. The experimenter explains that the children will learn the game first offline at the table and then continue to play the game on computers. Each player receives her endowment of five tuples of golden coins (see C.1) and it is demonstrated that these tuples consist of two single coins (kept together with Velcro).

It is emphasized that (i) for each coin tuple in the group basket the experimenter will put an additional coin tuple in there[39], (ii) that the resulting coins in the group basket will be split equally between all four players, no matter how many coins each player contributed to the group basket[40], (iii) that the final amount of coins for them will consist of the coins they put into their own basket plus the coins from the group basket, and (iv) that there is no right or wrong decision in this game. After the instructions, two trial rounds follow. In the first round, all children are supposed to put all their five coin tuples into the group basket—the experimenter then doubles the amount and demonstrates the payoffs by distributing the coins back to the players. In the second round, all but one player contribute to the group basket. This way, the experimenter demonstrates the fact that all players, independent of their contribution, will receive the same amount of coins from the group basket and that a child who puts her coins into her own basket can be better off in the end of the game. The experimenter then summarizes the incentives.[41] Finally, children can ask questions about the game. When there are no more questions, the experimenter tells the children that they can later on simply raise their hand to ask questions privately.

Subsequently, the children go to the computers where they receive the onscreen instructions and additional explanations via headphones (see Section 2.2). Handling is very intuitive, the children can simply use their fingers to click on the decision they want to make. The graphical interface was designed to closely mirror the offline version of the game (see Figures C.2 and C.3). Explanations start with referring to the offline game played before and state that this game will work exactly the same way. To ensure a good and complete understanding, we also repeat the central features already known from the offline explanation. Namely (i) four players, all from the child's class but anonymous who exactly, (ii) explaining the coin tuples and the two different baskets, (iii) that coins in the group basket will be doubled and subsequently distributed equally to all four children, (iv) that children will finally receive the coins from their own basket plus the coins from the group basket, and (v) that there are no right or wrong decisions. In addition, we stress that no other child will be able to observe their decision, that we will play eight rounds

[39] We used this way of picturing it in order to avoid misunderstanding due to the children not verbally grasping the concept of "doubling" the amount of coins in the basket.

[40] We specifically mentioned that "even if a player puts no coin tuple into the group basket", she will receive her share.

[41] The experimenter uses a predefined text to ensure highest possible standardization across groups. He mentions that "the more coins a child puts into her own basket, the more coins she receives. The more coins there are in the group basket, the more coins every child receives. The most coins for everybody together are achieved when all children put their coins into the group basket. But for each individual child it is better to put more coins into her own basket."

of the game, and that in the end of the game one of the rounds will be randomly selected for payoff (see Section C.2 for the exact wording of the instructions).

In the following, children see another example of how the game could take place. We use a mixture of possible strategies: the blue player (which always was the child's position) puts three coin tuples into the group basket, the yellow player likewise, the green player puts all five coin tuples into the group basket, while the red player puts all coin tuples into her own basket. The doubling of the coins in the group basket and the equal distribution to all the players are then animated. After each child confirms that she understood the game, the actual experiment starts. After all decisions are made, the children privately receive their respective payoffs in golden coins. On average, the offline instructions lasted for about 10 minutes, the computer-based game (including onscreen instructions) for 15 minutes, and the payoff took about five minutes.

B.4 Data Collection

In addition to collecting data on children's choices in the game, we collected background information on the children and their families with a parental questionnaire. At the end of the school day, we asked teachers to distribute the parental questionnaire and collect it within the next week. We received back 122 out of the 129 questionnaires. Table B.4 lists the exact items used in the parental questionnaire to assess children's cooperative behavior. The *Cooperation score* is the standardized average of these twelve survey items (some of them reversed). Moreover, we asked teachers to privately give us their assessment of children's cooperative behavior on a scale from 1 ("rather low social behavior") to 10 ("very distinct social behavior").

Table B.4: Survey Items to Construct Variable "Cooperation score".

	Mean	Std. Dev.	Min.	Max.	N
My child...					
(1) is able to, if necessary, set own wishes aside	4.867	1.566	1	7	120
(2) is more of an individualist, does not prefer to cooperate with others (reversed)	5.203	1.517	1	7	118
(3) is considerate	5.291	1.408	2	7	117
(4) likes to share with other children (sweets, toys, pencils etc.)	5.425	1.37	1	7	120
(5) is a loner, plays on her/his own most of the time (reversed)	5.908	1.309	1	7	120
(6) is helpful if others are hurt, ill or grieved	6.069	1.235	1	7	116
(7) has at least one good friend	6.445	1.212	1	7	119
(8) acts cooperatively in a group with other children	5.597	1.167	2	7	119
(9) in general, is popular with other children	5.692	1.256	2	7	120
(10) is kind to younger children	6.158	1.25	1	7	120
(11) helps others voluntarily (parents, educators, other children)	5.458	1.396	1	7	120
(12) rather profits from others in group works (reversed)	4.722	1.43	1	7	115

Notes: Adults evaluated their child on a Likert scale ranging from 1 ("does not fit at all") to 7 ("fits perfectly"). 122 parents handed in the parent questionnaire, but only 106 replied to all 12 questions. For aggregation, the scale is reversed for items labeled accordingly. The cardinal variable *Cooperation score* is the standardized average of the twelve parental survey items.

C Further Material

C.1 Public Goods Game and Treatment Details

Figure C.1: Golden Coins Used as Incentive and Currency in the PGG.

Figure C.2: The Public Goods Game for Children—Offline Version

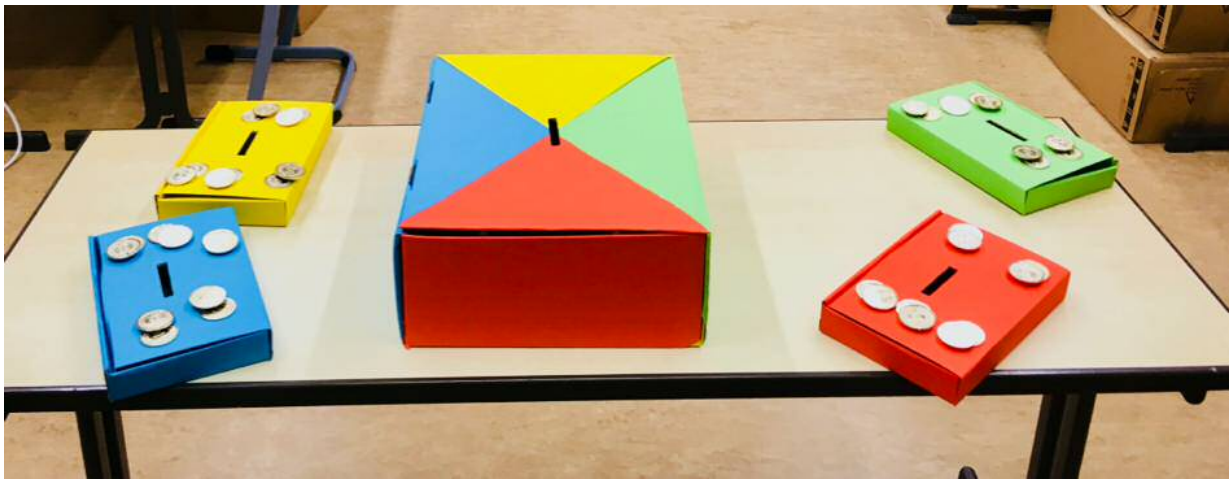


Figure C.3: The Public Goods Game for Children—Online Version with Ingroup Conditions (left and middle picture) and Outgroup Condition (right picture)

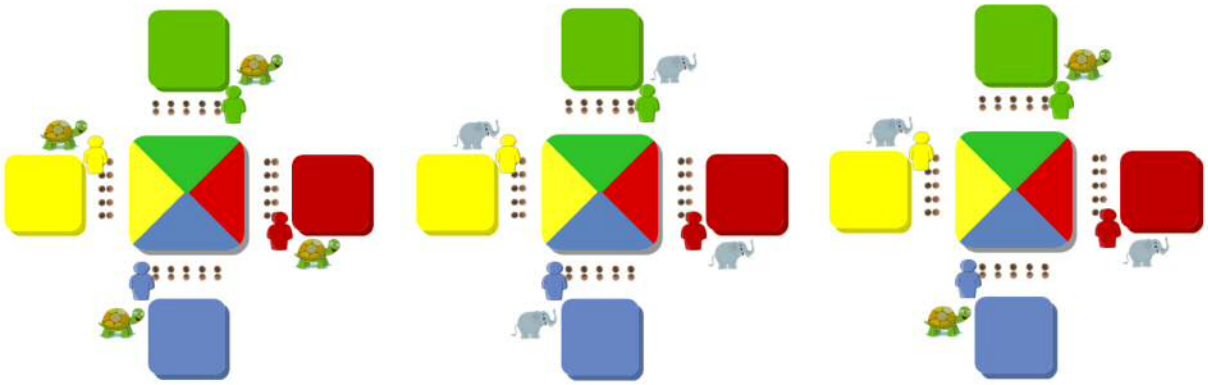


Figure C.4: Bagdes for Children to Assign Group Identity

C.2 Exact Instructions from the Public Goods Game

In the following, we provide the original auditory instructions used in the public goods game (translated from the original German text). Importantly, in the German original version we try to use a very simple and child-appropriate language with short sentences, few subclauses, and easy wording.

This game is played in a group of four and you will play with three of your classmates. Before the game, we divided your class into the elephant group and the turtle group. Every child got a badge which shows his or her groups animal. You are part of the "elephant"/"turtle" group.

Outgroup: One of the three other children who will play the game with you also belongs to the group of the "turtles"/"elephants". But the two other children are part of the other group, the group of the "elephants"/"turtles". So, there are two turtles and two elephants in your game.

Ingroup: The three other children who will play the game with you all belong to the same group as you do, the group of the "elephants"/"turtles".[42]

After the game, none of the other children will find out how you decided during this game—that is a promise!

The game works in the same way as the game we just played before. Look, here you see your five pairs of coins [*coins flash*]. This is your blue basket [*individual basket flashes*]. And here you can see the colored group basket [*group basket flashes*]. Remember: All coins which you put into the group basket will be doubled! After that, all coins will be split equally to all four children. Thus, at the end you get all coins from of your own basket [*individual basket flashes*] and further coins from the group basket [*group basket flashes*].

In this game, there is no right or wrong decision. You simply distribute your pairs of coins as you like it. The other children will not find out how many pairs of coins you put in your own basket—that is a promise.

Pay attention, now we will show you how the game might go along: The blue player puts two of his pairs of coins into his own basket and three pairs of coins into the group basket. The yellow player does the same. The green player puts all of his pairs of coins into the group basket and the red player puts all of his pairs of coins into his own basket. Now, we double all coins which were put into the group basket [*animation of doubling*] and the coins will be split equally to *all* children [*animation of distributing the coins equally*]. Probably you already realized what happens: The more a single child puts into her own basket, the more coins she will get. The more coins are put into the group basket, the more everybody gets. So, the most coins for *everybody together* will be realised if as many children as possible put their coins into the group basket. But for *each individual* child it is better to put more pairs of coins into her own basket.

[42] Accordingly, the there were four versions of this part of the intsructions: Ingroup for an elephant, ingroup for a turtle, outgroup for an elephant, and outgroup for a turtle.

Now it's your turn! First of all, we practise the game. Here you can start. If you tap on the colored arrow [*colored arrow flashes*], a pair of coins will be put into the group basket and if you tap on the blue arrow [*blue arrow flashes*], a pair of coins will be added to your own basket. Attention, you only have 30 seconds to distribute your pairs of coins. Below, you see a line [*line at bottom of screen flashes*] how much time remains for distributing your pairs of coins. As soon as the round starts, the line begins to get smaller. Alright, now you can test the arrows on your own and you can distribute your pairs of coins to the baskets.

[*When all pairs of coins are distributed:*] If you still want to change something, tap either on the colored [*colored arrow flashes*] or the blue arrow [*blue arrow flashes*]. If you like your distribution you can fix your final decision by clicking on the green check mark [*green check mark flashes*].

Do you understand how the game works? Then tip on the green button. Or is there something you do not understand yet? Then tip on the red button. Then somebody will come to you and will shortly explain the game to you again.

[*If red button was pushed:*] Please raise your arm. Soon somebody will come to you and will explain the game again.

Great! Now, you will get your pairs of coins and we can start! All in all, we will play eight rounds. After the first three sessions, there will be a minor change in the rules of the game which we will explain to you then.

Now the first session starts.

[*Whenever the 30 seconds for the decision are over, the child hears the following:*] You have to take a decision. Put your pairs of coins either into the blue basket or into the colored group basket.

Great, all children have distributed their pairs of coins and the coins from the group basket will be distributed now [*animation of how coins are equally distributed to the four players.*] In this session you won that many coins!

Ok, now the next session starts. [*Repeated for rounds 2 and 3.*]

[*Before round 4 starts:*] In the next five sessions we will change the game a bit. Now you have to decide whether you want to put ALL of your pairs of coins into your own basket [*animation of coins going to own basket*] or ALL of your pairs of coins into the group basket [*animation of coins going to group basket*]. Now it's your turn!

[*Before round 5:*] We will now slightly change the game once more. In the next four sessions, we first show you situations how the other children potentially could have played and only *then* you can distribute

your pairs of coins. Attention, the first session starts soon. [*Simultaneous animation of how other players distribute, no player contributes to the group basket.*] Look, that is how the other children distributed their pairs of coins. All children put all of their pairs of coins into their own baskets and no child put his pairs of coins into the group basket. Now it is your turn, would you like to put your pairs of coins into the group basket or into your own basket?

[*Before round 6:*] Now, the second session starts. [*Simultaneous animation of how other players distribute, one other player contributes to the group basket.*] Look, that is how the other children divided their pairs of coins. One child put all of her pairs of coins into the group basket and the two other children put all of their pairs of coins into their own baskets. Now it is your turn again, what would you like to do?

[*Before round 7:*] Now, the third session starts. [*Simultaneous animation of how other players distribute, two players contribute to the group basket.*] Two children put all of their pairs of coins into the group basket and one child put all of his or her pairs of coins into his own basket. It's your turn!

[*Before round 8:*] Now, the fourth session starts. [*Simultaneous animation of how other players distribute, three players contribute to the group basket.*] All of the other children put their pairs of coins into the group basket and no child put her pairs of coins into her own basket. What are you going to do?

Finished! Great! Now we draw by lot for which round you will receive your coins. Look, that is the amount of coins you collected! Now, raise your arm, so that somebody comes to you to give you your coins.

C.3 Parent Questionnaire

Parent Questionnaire KIDS-COOP

Thank you for allowing your child to participate in our research project Measuring cooperation in elementary schools. For analyzing the data of the project, we would need some further information about you and your child. We would kindly ask you to fill in this questionnaire in block letters and as complete as possible. The questionnaire will be collected by the teachers and is anonymised.

Child-ID: XXX

Questions about your child

1. Your child is... (a) a girl (b) a boy.
2. How many older and younger siblings does your child have?
   - The child has... (number) older siblings
   - The child has... (number) twin siblings
   - The child has... (number) younger siblings
3. How many siblings live in the same household as your child?

   ... (number) siblings live in the same household

4. How would you assess your child: Is he or she in general a venturesome child or does he or she tries to avoid risk?

   [Answers on a scale from 1 (not at all venturesome) to 7 (very venturesome)]

5. Is your child rather impatient or patient?

   [Answers on a scale from 1 (very impatient) to 7 (very patient)]

6. Does your child reflect for a rather long time before acting, so that she or he is a not at all an impulsive person? Or does she or he act without reflecting before acting, so that she or he is very impulsive?

   [Answers on a scale from 1 (not at all impulsive) to 7 (very impulsive)]

7. In how far do the following statements match the personality of your child? Please consider your child's behavior over the last six months. Please answer all questions as far as possible, even if you do not feel certain or if a question seems unusual. [Answers could be given on scale ranging from 1 (does not fit at all) to 7 (fits perfectly)]

   My child...

- is able to – if necessary – set own wishes aside
- is more of an individualist, does not prefer to cooperate with others
- is considerate
- likes to share with other children (sweets, toys, pencils etc.)
- is a loner, plays on her/his own most of the time
- is helpful if others are hurt, ill or grieved
- has at least one good friend
- acts cooperatively in a group with other children
- is, in general, popular
- is kind to younger children
- is teased and bullied by others
- helps others voluntarily (parents, educators, other children)
- rather benefits from others in group works
- better gets along with adults than with other children

8. Did you child attend kindergarden (or similar) before attending school?

   (a) No (b) Yes

9. How old was your child when it started to attend kindergarden (or similar)?

   ... years and ... months

10. How many classmates did your child already know before attending school?

    ... classmates

11. Does your child do sports outside the sports lessons in school?

    (a) No (b) Yes, ... (type of sport), approx....hours a week.

Questions about you and your family

12. How would you assess yourself: Are you in general a cooperative person?

    [Answers on a scale from 1 (less cooperative) to 7 (very cooperative)]

13. Are you in general a venturesome person or do you try to avoid risk?

    [Answers on a scale from 1 (not at all venturesome) to 7 (very venturesome)]

14. Are you rather impatient or patient?

[Answers on a scale from 1 (very impatient) to 7 (very patient)]

15. Do you reflect for a rather long time before acting, so that you are a not at all an impulsive person? Or do your rather act without reflecting, so that you are very impulsive?

    [Answers on a scale from 1 (not at all impulsive) to 7 (very impulsive)]

16. Does your family own a car?

    (a) no (b) yes, one (c) yes, two or more

17. Does your child has a bedroom for her/himself?

    (a) yes (b) no

18. During the past 12 months, how many times did you travel abroad with your family?

    (a) not at all (b) once (c) twice (d) more than twice

19. How many computers does your family own? (including tablet computers, without smartphones)?

    (a) none (b) one (c) two (d) three (e) more than three

20. How many books do you have at home?

    (a) none/a few (b) a complete single board of a book shelf (c) a complete book shelf (d) two complete book shelves (e) more than two complete book shelves

21. Finally, we would like to ask you for some further personal questions

    Which relation do you have to your child?

    I am...

    - the biological mother
    - the biological father
    - the adoptive or foster mother
    - the adoptive or foster father
    - the stepmother
    - the stepfather
    - other .....

    Are you born in Germany or in another country?

    (a) Germany other .....

What is your highest degree?

- no graduation
- Board school/lower secondary education
- Secondary school certificate
- A levels/higher education entrance qualification
- other .....

Did you finish an apprenticeship or a course of studies?

- Yes, apprenticeship
- Yes, graduation at university/college
- No, no professional education

All of your information will, of course, be anonymised and will only be analyzed for scientific purposes. All of your information will be treated confidentially and personal data will not be handed over to third parties.

Thank you very much for your collaboration!