



Human Capital and Economic Opportunity Global Working Group

Working Paper Series

Working Paper 2016-011

The Formation of Prosociality: Causal Evidence on the Role of
Social Environment

Fabian Kosse
Thomas Deckers
Hannah Schildberg-Hörisch
Armin Falk

April, 2016

Human Capital and Economic Opportunity Global Working Group
Economics Research Center
University of Chicago
1126 E. 59th Street
Chicago IL 60637
www.hceconomics.org

The Formation of Prosociality: Causal Evidence on the Role of Social Environment

Fabian Kosse¹, Thomas Deckers¹, Hannah Schildberg-Hörisch¹
and Armin Falk^{*,1,2}

* Corresponding author, armin.falk@uni-bonn.de

¹ University of Bonn, Institute for Applied Microeconomics

² Behavior and Inequality Research Institute (briq)

April 2016

Abstract

This study presents descriptive and causal evidence on the role of social environment for the formation of prosociality. In a first step, we show that socio-economic status (SES) as well as the intensity of mother-child interaction and mothers' prosocial attitudes are systematically related to elementary school children's prosociality. In a second step, we present evidence on a randomly assigned variation of the social environment, providing children with a mentor for the duration of one year. Our data include a two-year follow-up and reveal a significant and persistent increase in prosociality in the treatment relative to the control group. Moreover, enriching the social environment bears the potential to close the observed developmental gap in prosociality between low and high SES children. Our findings suggest that the program serves as a substitute for prosocial stimuli in the family environment.

Keywords: Formation of preferences, prosociality, social preferences, trust, social inequality.

JEL-Codes: D64, C90

*Armin Falk gratefully acknowledges financial support through the Leibniz Programme of the German Science Foundation (DFG) and the Benckiser Stiftung Zukunft.

1 Introduction

Prosociality is a particularly important aspect of human personality and affects a wide range of economic decisions and outcomes, e.g., the provision of public goods, contract enforcement, management of commons, governmental and judicial efficiency or economic growth (Fehr and Gächter, 2002; Henrich et al., 2004; Knack and Keefer, 1997; Ostrom et al., 2002; La Porta et al., 1997; Guiso et al., 2009; Cooper and Kagel, 2009; Burks et al., 2016). Prosocial behavior is not only a crucial factor for the functioning of societies but also an important skill that affects individual success and well-being (Carpenter and Seki, 2011; Becker et al., 2012; Algan et al., 2013). Despite its fundamental importance, and significant advances in understanding the *consequences* of prosocial motivation, little is known about how prosociality *forms*. This paper therefore provides evidence on the formation of prosocial motivation, in particular on the causal role of social environment.

Our research strategy builds on the conceptual framework suggested by Cunha and Heckman (2007).¹ They highlight *childhood* as critical and sensitive period for the formation of personality², and identify *parental background* and *investments* as the two primary drivers of personality formation. In this paper, we study both drivers in a sample of elementary school children: In terms of parental background we study the role of socio-economic status, patterns of social interaction between mother and child, as well as intergenerational transmission of prosociality. To study the causal impact of investments we randomly assigned children to an enriched social environment in the form of a mentoring program. This mentoring program is a well-established non-profit program called “Balu und Du” (German for “Baloo and You”). The program provides children with a mentor for the duration of one year. Conceptually, the idea of the program is to extend a child’s horizon and to foster the acquisition of new skills and experiences through social interactions between mentor and child. By interacting with the mentor the child experiences that an unrelated and new contact person takes responsibility and spends effort and time with him. Hence, the mentor enriches a child’s social environment, and serves as a potential prosocial role model.

Before and after the intervention, children and their parents were interviewed by trained interviewers. Children participated in incentivized choice experiments and answered a short questionnaire. Mothers completed an extensive questionnaire cov-

¹For biological and genetic determinants see, e.g., Fehr and Fischbacher (2003), Kosfeld et al. (2005) and Cesarini et al. (2008, 2009).

²For descriptive evidence on development patterns of prosociality see, e.g., Harbaugh and Krause (2000), Sutter and Kocher (2007), Almås et al. (2010) and Fehr et al. (2008, 2013).

ering socio-economic background information, interaction patterns and assessments of personality regarding their children and themselves. To yield a comprehensive measure of a child’s prosociality we collected data on three facets: altruism, trust and other-regarding behavior in everyday life. All measures are based on established instruments that have been shown to predict real-life societal and individual success (Falk et al., 2015; Becker et al., 2012). Our measure of altruism is children’s behavior in three incentivized experiments. Trust was measured using a well-established and experimentally validated three-item trust questionnaire. Moreover, we asked mothers to assess their child’s general other-regarding behavior in everyday life. Our measures thus provide a comprehensive characterization of prosocial disposition using different data collection methods and statements from children and mothers. For the main analysis, we collapse the three facets into one joint measure of children’s prosociality. In a similar vein, we generate a measure of prosociality for mothers and mentors. As for children, this measure consists of the three facets altruism, trust and other-regarding behavior.

Our sample was recruited using official registry data and consists of families interested in participating in the mentoring program and the interviews. More specifically, we study three distinct groups, which differ in terms of socio-economic and treatment status: Based on socio-demographic background information (household income, parental education and single parent status) we classify families as either low or high socio-economic status (SES). Among low SES families, we randomly assigned a subset of families to participate in the intervention. This is the treatment group (Treatment Low SES). The remaining families with low SES background form our intervention control group (Control Low SES). The third group consists of families with high SES background (Control High SES). Comparing outcomes between these groups allows us to study the two potential drivers of personality formation, parental background and investment, respectively. The comparison between Control Low SES and Control High SES informs about the role of parental background, i.e., differences in socio-economic endowments, social interaction patterns and intergenerational transmission of prosociality. Comparing Control Low SES and Treatment Low SES provides causal evidence on the potential role of investments in the form of providing an enriched social environment. In total, we collected three waves of data. After a baseline wave of interviews (wave 1), the treatment was randomly assigned and implemented. The one-year treatment period was closely followed by a post-treatment wave of interviews (wave 2), and a two-year follow-up wave (wave 3). In the main analysis we focus on post-treatment wave 2 data, but also use wave 1 to study baseline balance and systematic attrition, as well as wave 3 data to investigate

whether the observed effects are enduring.

Our main findings can be summarized as follows. First, we find a pronounced developmental gap in elementary school children, depending on SES: Children from Control High SES households are significantly more prosocial than children from Control Low SES households. The effect amounts to 22.6% of a standard deviation in prosociality. Moreover, we provide evidence on the intergenerational transmission of prosociality, and on the importance of mother-child interactions for the formation of prosociality. This sets the stage for our main analysis: Our second result shows that children whose social environment was randomly enriched through participation in the mentoring intervention score 27.3% of a standard deviation higher on the prosociality measure than children from the control group. Moreover, the observed high-low SES developmental gap in prosociality is closed, i.e., children from Treatment Low SES and Control High SES score very similarly on the prosociality measure. A particular important question concerns persistence, i.e., whether the effect of the mentoring program is enduring. Our third result shows that this is the case. Using wave 3 data indicates a general increase in prosociality for all three groups, and in particular persistence of the high-low SES gap and the treatment effect. Two years after the end of the intervention prosociality in Treatment Low SES is significantly higher than in Control Low SES, and there is no significant difference between Treatment Low SES and Control High SES.

Our findings reveal that prosociality is generally malleable, and confirm that parental background and investments are key drivers in the development process. They also lend support for the effectiveness of childhood interventions (Kautz et al., 2014; Rodríguez-Planas, 2012). In other words, interventions such as mentoring programs have the potential to systematically affect character formation (Heckman et al., 2013) and to close developmental gaps arising from “the accident of birth” (Heckman, 2008, 2013). In this sense our findings inform the discussion on social mobility, societal inequality and the intergenerational persistence of life outcomes (Putnam, 2015; Aizer and Currie, 2014; Currie and Moretti, 2003; Case et al., 2002).

The remainder of the paper is organized as follows. In the next section we discuss recruitment, our measures, details of the mentoring program and implementation of the intervention. Section 3 shows our results concerning parental background and investments, and section 4 concludes.

2 Study Design

This section introduces the design of the study. We first report how children and their families were recruited, how we classified them in terms of SES and how children were randomly assigned to treatment, resulting in our three groups Control Low SES, Treatment Low SES and Control High SES, respectively. We then describe the intervention, i.e., scope and concept of the mentoring program. Finally, we discuss the setting and details of the interviews, including a description of our survey and experimental measures.

2.1 Recruitment of sample

Figure 1 presents a flow chart of the timing, sampling and procedural details of our study. Recruiting started in summer 2011. We used official registry data to obtain more than 95% of the addresses of families (with children of age seven to nine) living in the German cities of Bonn and Cologne. Invitations to take part in the mentoring program and the interviews were sent to all families with children born between 09/2003 and 08/2004 and to one third of the families with children born between 09/2002 and 08/2003. Birth dates were chosen such that children of the younger cohort were typically in second grade.³ In summer 2011, families were contacted via postal mail. We announced the possibility to take part in the mentoring program and the interviews. We informed parents that participation in the mentoring program was not guaranteed due to limited capacity (which was the case). We asked them to send back a short questionnaire concerning socio-economic characteristics of the household, and to sign a non-binding letter of intent to take part in the interviews and the mentoring program, if interested. We received 1,626 complete and valid responses.⁴ Using responses to the questionnaire we categorized households as either low or high socio-economic status (SES) households, respectively. SES reflects the level of resources available at the household level, i.e., material, educational and time resources. Accordingly, a household was classified as low SES if at least one of the three following criteria was met: (i) *Low income*: Equivalence income of the household is lower than 1,065 Euro. This corresponds to the 30% quantile of the

³Nearly all children in our sample (99%) attend public schools. The quality of public elementary schools in Germany is very homogeneous. Teachers are typically centrally allocated to schools, have obtained the same qualifications, and are paid the same salaries.

⁴An additional requirement was that the families speak some German. We did not exclude families with migration background, however. In fact, 34.4% of the participating children have at least one parent who was not born in Germany.

German income distribution.⁵ (ii) *Low education*: Neither mother nor father of the child have a school-leaving degree qualifying for university studies.⁶ (iii) *Single parent status*: A parent is classified as single parent if he or she is not living together with a partner.⁷ Households for which none of the three criteria applied were classified as high SES.

Low SES households form our target group and therefore we invited all low SES families to take part in the study. To be eligible for treatment, they had to participate in the first wave of interviews (fall 2011) and to provide written consent to allow the transmission of their addresses to the organization running the mentoring program. Out of 590 eligible families, 212 were randomly selected and constitute our intention-to-treat (ITT) group (Treatment Low SES). The remaining 378 families form the control group (Control Low SES). We used stratified random treatment assignment: Stratification considered 14 subgroups resulting from the combination of local (Bonn or Cologne) and SES criteria (low income and/or low education and/or single parent status). Stratification was used to ensure a proportional representation of all combinations of criteria in the ITT group, and that the number of selected children matched the local supply of mentors.⁸ After treatment assignment, we transmitted contact information of the ITT group to the mentoring organization who initiated the treatment.

As a second control group, we also invited 150 randomly chosen high SES families (among those who had answered the information letter) to take part in the wave 1 interviews (Control High SES). As for families with low SES background we asked them to provide their written consent to allow the transmission of their addresses to the organization running the mentoring program. 122 families took part in the first wave and gave written consent.

After the one-year treatment period, all families who had participated in wave 1 were invited to take part in the post-treatment wave (wave 2). 85.3% of them (607 out of 712) took part in this second wave of interviews. This is our core sample. Two years after the end of the treatment period, all families who had participated

⁵The distribution was calculated using the 2009 wave of the German Socio-Economic Panel (SOEP) (Wagner et al., 2007) and the cross-sectional weights provided in the SOEP data.

⁶Note that the German secondary school system consists of basically three tracks. Only graduation from the highest track (Gymnasium) qualifies for university studies.

⁷With respect to single parent status we expected that single-parent households dispose of less time resources to spend with their children than households with two parents. Using survey data (mother interviews) we can actually show this. Children living in single parent households spend 45.8% more time “alone at home” than children in households with two parents.

⁸Given the larger relative supply of mentors in Bonn, we also assigned a higher share of children in Bonn to the ITT group. Therefore, assignment into treatment was random conditional on city of residence. To account for this we report conditional treatment effects in the Appendix.

in the second wave were invited to take part in the third wave of interviews. 83.9% (509 out of 607) took part in this two-year follow-up.⁹

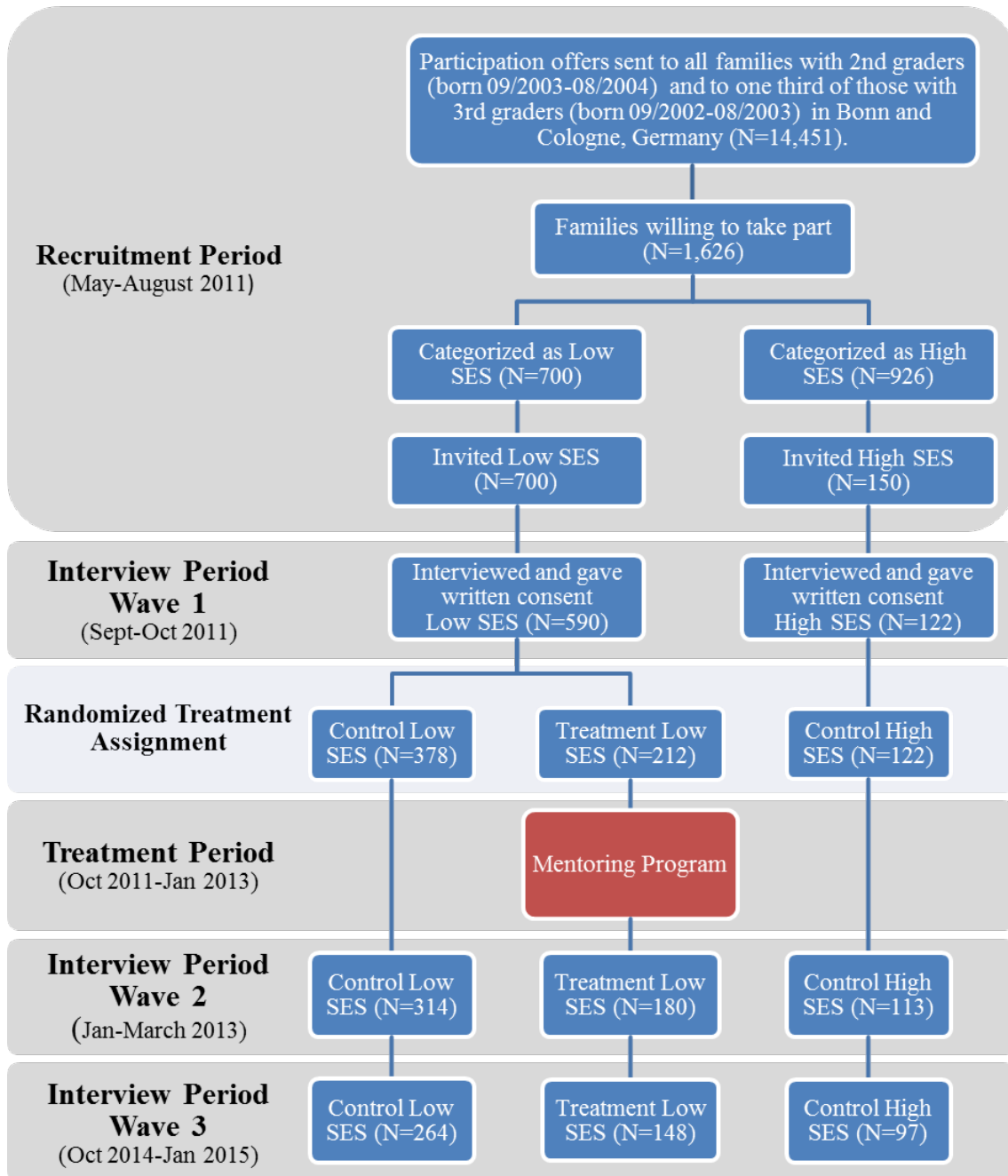


Figure 1: Flow chart of sampling and procedural details.

2.2 The mentoring program

The intervention which we implemented is a well-established non-profit mentoring program called “Balu und Du” (German for “Baloo and you”). In this program,

⁹See section 3.4 for a discussion on baseline balance and attrition.

elementary school children are provided with a mentor for the duration of one year. The mentors – called Baloos – are mainly university students (aged from 18 to 30) who voluntarily care for their mentees – called Mowglis. The conceptual idea of “Balu und Du” is to focus on “informal learning”, a concept that integrates and reinforces learning processes in children’s everyday life. According to informal learning, the mentors act as role models and “benevolent friends” who encourage the acquisition of new ideas and skills by enriching the social environment of the children.¹⁰

On a practical level, a mentored child typically spends one afternoon per week in one-to-one interaction with his or her mentor. During this time, they engage in joint activities, which are adapted to the individual needs, strengths, weaknesses, and interests of the child (and mentor). Examples are visiting the zoo, museum, or playground, cooking, doing handicraft, ice skating, or just having a conversation. Hence, a child spends time with an additional attachment figure and role model, gains new experiences, and learns that he or she is valued by the fact that the mentor regularly spends time with him or her.

The “Balu und Du” mentoring program is embedded in a professional structure. On a weekly basis, mentors complete an online diary in which they report the activities they engage in as well as potential problems of the mentor-child relationship. Program coordinators read and comment these diaries, and provide support. These coordinators are trained and paid professionals in education science or psychology and provide supervision and advice to mentors. They also organize biweekly monitoring meetings where mentors receive suggestions for activities with the mentored child, and discuss potential problems. To date, the mentoring program “Balu und Du” has arranged and supervised more than 7,500 mentor-child relationships in more than 50 different locations in Germany.¹¹

The mentoring program is designed to last up to 12 months. In our sample

¹⁰For further details see Müller-Kohlenberg and Drexler (2013). The informal, low-threshold and relation-based character of the intervention under study is complementary to more specifically targeted and/or school-based interventions. For an overview and discussion see Kautz et al. (2014). Prominent examples for high intensive curriculum-based programs are the Perry Preschool Program and the Abecedarian Program (ABC) (Kautz et al., 2014). Alan and Ertac (2014) represents a recent example of a training program targeted to change preferences, for an early example see Bandura and Mischel (1965). Effects of school-related interventions on social behavior are studied, e.g., in Bettinger and Slonim (2006) and Rao (2013). Chetty et al. (2016) explore the Moving to Opportunity Program (MTO) and represent an example for studying the effects on life outcomes in response to fundamental changes in the social environment by moving to a better neighborhood.

¹¹The program has been honored with several public awards, e.g., by the Robert Bosch Foundation in 2011 and by the federal government of North Rhine-Westphalia (Germany) in 2006. More details about the mentoring program can be found on www.balu-und-du.de and in an overview article by Müller-Kohlenberg and Drexler (2013).

the actual average duration of mentor-mentee relations was 9.3 months (Figure A1). Variation in duration is mainly due to unforeseeable events such as moving decisions of parents or mentors due to job change. On average, treated children met their mentor 22.8 times (std. dev. 11.9), typically for a whole afternoon.

We transmitted household addresses of all randomly selected families to “Balu und Du”. The actual matching process of mentor and mentee is part of the program and was conducted by “Balu und Du”. Each child in the intention-to-treat (ITT) group could potentially be matched but not all selected children were effectively matched with a mentor. A mentor-mentee match was successfully implemented for 74% of the ITT group children. For 8% of the children matches were actually initiated by “Balu und Du” but could not be realized due to refusals, availability, or coordination problems between mentors and families (e.g., pregnancy of mentor, moving of mentor or family, etc.). These children never met a mentor. Another 18% of children could not be matched to a mentor by “Balu und Du” due to an unexpected (local) shortage of mentors. These households were never contacted by “Balu und Du”. In the main analysis we focus on ITT effects and discuss local average treatment effects (LATE) in section 3.4.

2.3 Setting of interviews and experiments

In all interviews the child was accompanied by one parent. In 95% of the cases the interviewed parent was the biological mother. For convenience we therefore use the term “mother” for the adult who was interviewed. In waves 1 and 2, interviews took place at central locations in either Bonn or Cologne. For that purpose we rented two large flats, one in each city. The interviews and experiments were conducted by trained interviewers with a background in psychology or education science. In wave 3, interviews took place at participants’ homes. In this wave, interviews and experiments were conducted by experienced and trained interviewers of the same professional surveying company that administers the SOEP (Wagner et al., 2007). At no point in time, interviewers were informed about the purpose of the study or the group assignment of the participating families (neither treatment/control nor high/low SES). The interviews and experiments were conducted according to a detailed protocol (see Appendix), which was identical across all three waves. In total, interviews lasted about one hour. For participation in the interview, mothers received 35 Euros in wave 1 and 45 Euros in waves 2 and 3, respectively. The experiments run with children were incentivized using an experimental currency called “stars”. At the end of the interview, children exchanged their stars for toys

(in waves 1 and 2) or money (in wave 3).

Toys were arranged in four categories which visibly increased in objective value and subjective attractiveness to children (see Figure A2 in the Appendix). During the experiments, children knew that more stars would result in the option to choose toys from a higher category.¹² In wave 3, children had reached an age where toys are no longer appropriate as incentive. We therefore changed to money in this wave, with one “star” corresponding to 0.3 Euro.

We took care to create a pleasant, non-stressful interview situation by seating a mother and her child in the same room. To avoid interaction between the two, however, a standardized seating plan ensured that mother and child could not directly see each other. In addition they were not allowed to communicate. One experimenter ran the experiments with only one child at a time. During the experiments, mothers completed a comprehensive survey covering topics such as basic information about the child, mother assessments of personality and attitudes of the child, socio-economic background of the family, details on how parents spend time with the child including joint activities, and personality of the mother (Big 5, trust and altruism).

2.4 Measures of prosociality and social interaction

Altruism: Incentivized dictator game experiments

Children’s altruism was measured using three incentivized versions of dictator games with children in the role of the decision maker (“dictator”). The participating child had to allocate amounts of the experimental currency (stars) between him- or herself and another anonymously matched child of the same age. In particular, we conducted one binary dictator game and two continuous versions of dictator games with varying receivers (see Appendix for experimental protocols).

In the binary dictator game children had to decide between two possible allocations of two stars, regarding themselves and another unknown child from the same city (see Fehr et al., 2008). One option was that both, decision maker and receiver, receive one star (1,1). The other option implied that the decision maker received two stars, while the receiver received no star (2,0). In both continuous versions, the decision makers were endowed with six stars and could choose how to distribute the six stars between themselves and the other child, respectively. The two continuous versions differ by the characteristics of the receiver. In one version, the receiver was an unknown child who lives in a city nearby. In the other version, the receiver was

¹²We made sure that each additional star that would not result in a higher category was nevertheless valuable: These stars were exchanged into “Lego” bricks.

an unknown child who lives in an “African country” and who cannot live with his or her parents because they are either “poor or ill” or “have passed away”.

In all three versions, the allocations were physically shown and the children answered control questions. Our measure of altruism is the average share of stars given in each of the three dictator games. In our wave 2 sample, the mean share is 0.390 with standard deviation 0.156 ($N = 606$; all children answered the control questions correctly¹³, one missing observation).

Importantly, all decisions had real consequences for the participating children and the anonymous receivers. We cooperated with three charities (in Cologne, Bonn and Togo, respectively) to implement the allocation decisions as described. All children in the role of receivers (Cologne, Bonn and Togo) benefited from receiving stars in form of toys. To benefit receivers in Cologne and Bonn we collaborated with two established local charity organizations. To implement the dictator game outcomes with children who “live in Africa” we collaborated with an SOS Children’s Village in Togo.

Trust: Questionnaire answers of the child

Children answered three questions concerning trust. These survey items are taken from the SOEP (Wagner et al. 2007) and have been experimentally validated (Fehr et al. 2002, Falk et al. 2016). We slightly adapted these items to make them appropriate for children in the age range under study. In particular, the statements read as follows: “One can trust other people”, “Other people have good intentions towards me”, and “One can rely on other people, even if one does not know them well”. The statements were read out aloud by the interviewer and children indicated how much they agree to the statements using a five-point Likert scale ranging from “totally disagree” to “totally agree”. As shown in Figure A3, the scale was printed on an extra sheet of paper and additionally visualized. To further facilitate understanding, the interviewer explained the procedure using a simple neutral example item (“I like Spaghetti”). The average response to the three items is our measure of a child’s trust. In our wave 2 sample, the mean is 3.193 with standard deviation 0.765 ($N = 607$; all 607 children answered all three trust questions).

Other-regarding behavior: Mother survey

As part of the mother survey, every mother assessed her child’s other-regarding behavior in everyday life using the Prosocial Scale of the Strength and Difficulties Questionnaire (SDQ) (Goodman, 1997). The SDQ is a well-established behavioral screening survey consisting of five subscales. The Prosocial Scale includes five items

¹³Only one child needed a repetition of the rules.

which read as follows¹⁴: My child is... “considerate of other people’s feelings”, “shares readily with other children”, “helpful if someone is hurt, upset or feeling ill”, “kind to younger children”, “often volunteers to help others”. Responses were given on a seven-point Likert scale ranging from “does not apply at all” (1) to “applies completely” (7). The average answer to these five items is our measure of a child’s other-regarding behavior in everyday life. In our wave 2 sample, the mean is 5.837 with standard deviation 0.972 ($N = 605$; according to the manual we include an observation if at least three out of five items are completed; two missing values remain).

Joint measure: Prosociality

In sum, we obtained three facets of prosociality – altruism, trust and other-regarding behavior – combining experimental choice and survey data, and exploiting responses of the child and the mother. Each single measure is based on multiple responses (three dictator games, three trust questions and five other-regarding behavior items), which reduces measurement error. For the main analysis, we collapse the three facets of the underlying prosocial disposition into one joint measure of prosociality. This measure is the equally weighted score of the standardized measures of the three facets. To ease comparability, for all analyses in this paper (waves 1, 2 and 3) we use wave 2 means and standard deviations for standardization. To limit missing observations, if one of the three facets is missing the joint measure is calculated as equally weighted score of the two observed facets.

Prosociality of mothers and mentors

In order to obtain measures of prosociality for mothers and mentors, respectively, we proceed as similarly as possible as for the children: We construct an equally weighted score using standardized measures of altruism, trust and other-regarding behavior. All measures are collected using established and validated survey items. Altruism was measured using the question “How would you assess your willingness to share with others without expecting anything in return, for example your willingness to give to charity?” (Falk et al., 2016). Trust was measured using the two items “In general, one can trust people” (Fehr et al., 2002) and “As long as I am not convinced otherwise I always assume that people have only the best intentions” (Falk et al., 2016). Responses were given on an 11-point Likert scale. Other-regarding behavior was measured using the Big Five dimension Agreeableness in form of a three-item version (seven-point Likert scale) (Lang et al., 2011). The items read as follows: I see myself as someone who “is sometimes somewhat rude to others” (reversed),

¹⁴We used the wording of the official German SDQ version, see www.sdqinfo.org.

“has a forgiving nature” and “is considerate and kind to others”. Note that the Agreeableness items use a similar wording as the Prosocial Scale of the SDQ which we use to measure children’s other-regarding behavior. Moreover, Agreeableness is theoretically and empirically related to concepts of other-regarding behavior (Becker et al., 2012). As for children, we construct an equally weighted score of the three measures as our measure of prosociality for mothers and mentors ($N(\text{Mothers})= 597$, $N(\text{Mentors})= 98$, missing observations due to incomplete or missing questionnaires).

Social interaction patterns

We hypothesized that potential differences in children’s prosociality could arise from differences in parent-child interactions, and the intense social interaction between children and their mentors. Therefore, we asked mothers how they spend time together with their child. We focus on joint activities similar to those mentors and mentees engaged in: Having a conversation; having a snack together (e.g., a cake); playing board or card games; playing music together or going to music lessons. For each item mothers were asked: “How many times during the last 14 days have you or the main caregiver done the following activities together with your child?”. Answers were given on a 4-point Likert scale. To yield one measure that captures the intensity of social interaction of mother and child, the standardized items are aggregated using confirmatory factor analysis (direct maximum likelihood estimator).

3 Results

In presenting our results we first discuss the impact of parental background. We compare the prosociality of untreated children, i.e., Control Low SES and Control High SES, and report the effects of SES, mothers’ prosociality as well as interaction patterns between mothers and their children. In section 3.2 we then show our main finding, the causal effect of the intervention on prosocial dispositions in children. In section 3.3, we explore heterogeneous treatment effects, connecting our findings on parental background and the intervention. For our main analyses we use the wave 2 sample ($N = 607$; 47.0% girls; age at the start of the program: Mean = 7.76 years, std. dev. = 0.48). However, to study long-run effects of the intervention and persistence of parental background over a two-year developmental period, we also refer to wave 3 data. Finally, in section 3.4, we make use of wave 1 data to provide several robustness checks. This includes tests for baseline balance, the absence of selective attrition, and a confirmation of our main results using inverse probability weighting and difference-in-differences estimations.

3.1 Descriptive results on the effect of parental background: SES, social interaction, and mothers' prosociality

We hypothesized that the relatively high level of material, educational and time resources available in high SES households has a positive effect on prosociality, relative to low SES households (Bauer et al., 2014; Benenson et al., 2007). Differences in resources could in particular manifest themselves in different patterns of daily interaction and thus affect the development of prosociality. Social reinforcement in form of intensive interactions and feedback provision is central within the process of learning and adopting social behavior (Skinner, 1953; Bandura, 1986). Building on recent empirical and theoretical work (e.g., Bisin and Verdier, 2001; Dohmen et al., 2012; Kosse and Pfeiffer, 2012; Alan et al., 2013; Zumbuehl et al., 2013) we further expected a positive impact of mothers' prosocial attitudes, i.e., an intergenerational transmission of prosociality.

Table 1 reports our findings concerning parental background for the two untreated groups of children, Control Low SES and Control High SES. Column (1) shows that children from high SES households are significantly more prosocial than children from low SES households ($p < 0.05$). In terms of effect size, the difference amounts to 22.6% of a standard deviation. Column (2) indicates that spending more time on social interaction and feedback providing tasks promotes children's prosociality. Children who experience a one standard deviation higher intensity of social interaction with their mothers are 16.2% of a standard deviation more prosocial ($p < 0.01$). Column (3) shows that mothers' prosocial attitudes strongly affect children's prosociality. A one standard deviation increase in mother's prosociality is related to a 27.0% of a standard deviation increase in her child's prosociality ($p < 0.01$).

Our data further indicate that SES is correlated with the intensity of social interaction as well as mothers' prosociality. High SES families spend on average 35.9% of a standard deviation more time on socially interactive tasks than low SES families ($p < 0.01$, $N = 418$, two-sided t -test). High SES mothers are 20.4% of a standard deviation more prosocial than low SES mothers ($p < 0.1$, $N = 418$, two-sided t -test). In column (4) of Table 1 we therefore regress children's prosociality jointly on SES, social interaction and mothers' prosociality. While the coefficients of social interaction and prosociality of the mother remain significant, the coefficient of SES drops by about 40% and is no longer significant. This suggests that part of the high-low SES gap in children's prosociality can be explained by differences in social interaction and mothers' prosociality.

	Prosociality			
	(1)	(2)	(3)	(4)
High SES	0.226** [0.106]			0.136 [0.105]
Social Interaction		0.162*** [0.045]		0.111** [0.047]
Prosociality of Mother			0.270*** [0.047]	0.247*** [0.047]
Constant	-0.127** [0.059]	-0.066 [0.049]	-0.087* [0.048]	-0.121** [0.056]
Observations	418	418	418	418
Adjusted R^2	0.008	0.023	0.070	0.083

Table 1: Prosociality and parental background. Prosociality measures of mother and child are constructed as equally weighted scores of standardized measures of altruism, trust and other-regarding behavior, respectively. Prosociality and Social Interaction measures are standardized, High SES is a dummy. For reasons of comparability the sample is restricted to individuals with non-missing values in all used variables. Data come from wave 2. Coefficients are OLS estimates with robust standard errors in brackets. ***, **, * indicate significance at 1-, 5-, and 10-percent level, respectively.

Our definition of SES aims at reflecting the level of resources available at the household level, i.e., material, educational and time resources. Any definition of SES is debatable, however. It is therefore reassuring that our finding of lower levels of prosociality in low SES children in comparison to high SES children also holds if we consider alternative definitions of SES. Given our three criteria (income, education, and single parent status) there exist seven possible combinations for which we can estimate a potential SES effect. This is shown in Figure A4. For each definition of SES, we find that children from low SES families score lower on prosociality than children from high SES families. The figure also illustrates that parental education plays a particularly important role.¹⁵

¹⁵Another potentially important difference in family background concerns migration status. 34.4% of the children in our sample have a migration background, defined as mother and/or father not born in Germany. 23.4% of the children with migration background have Turkish roots and represent the largest group among the immigrants. The data indicates that neither general migration background nor specifically Turkish background is related to children's prosociality ($p > 0.7$, $N = 418$, two-sided t -test).

3.2 The causal effect of social environment on prosociality

Unlike the correlational evidence reported in the previous section, the randomized controlled implementation of the intervention allows studying the causal effect of social environment. Children participating in the mentoring program experience an enrichment of their social environment, including more intense social interactions and exposure to a prosocial role model. As we have shown above, both of these experiences are relatively scarce in low SES families in comparison to high SES families, contributing to the difference in prosociality between high and low SES. Moreover, evidence on social learning suggests an adaptation to the prosocial environment provided by the program (Skinner, 1953; Bandura, 1986). Hence, we hypothesized a positive effect of the mentoring program on children’s prosociality. We expected a particularly pronounced effect for children living in households where mothers score low on our prosociality score, and who experience relatively little social interactions.

The left panel of Figure 2 shows our main result. In wave 2 (post-treatment) children whose social environment was randomly enriched through participation in the mentoring intervention score 27.3% of a standard deviation higher on the prosociality measure than children from the control group ($p < 0.01$, $N = 494$, two-sided t -test). Figure 2 also indicates that the high-low SES developmental gap in prosociality is closed for children in the treatment group: Children from Treatment Low SES and Control High SES score very similarly on prosociality ($p = 0.651$, $N = 293$, two-sided t -test). The positive effect of the intervention is not only observed for the joint measure of prosociality but holds independently for the three facets of prosociality, altruism ($p < 0.05$), trust ($p = 0.05$) and other-regarding behavior ($p = 0.266$), see Table A1¹⁶.

¹⁶A joint F -test rejects the null-hypothesis of a zero treatment effect for the three facets of prosociality (altruism, trust, other-regarding behavior; $p = 0.010$, $F = 3.78$, $N = 491$).

Prosociality

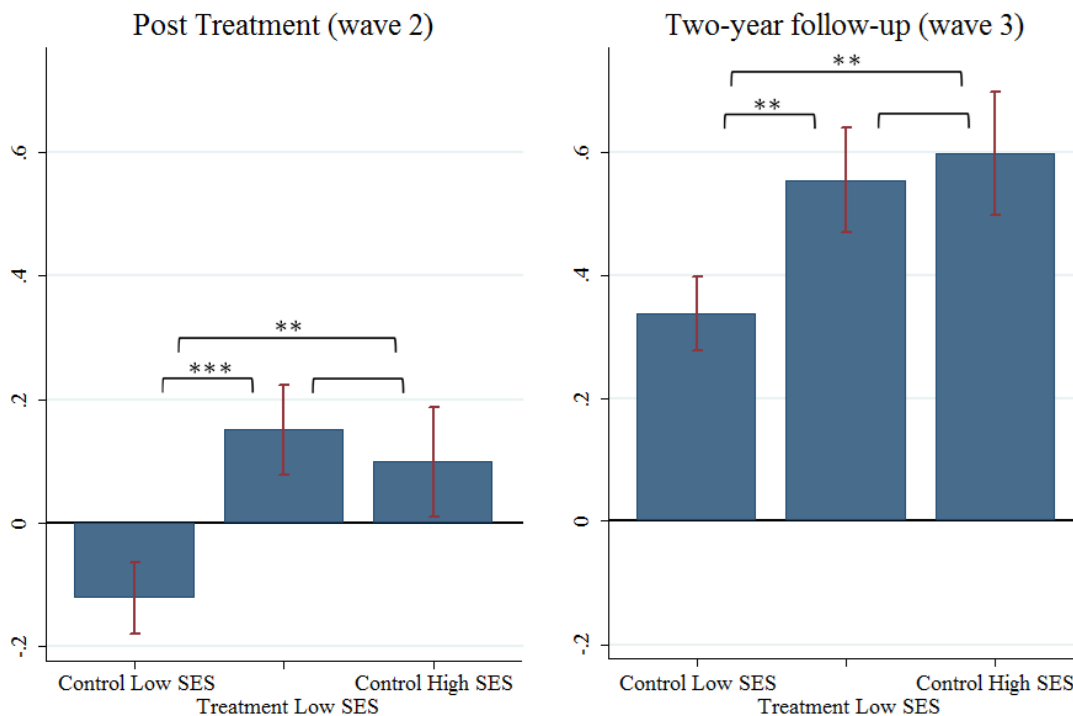


Figure 2: Significantly higher levels of prosociality for treated children compared to untreated children (Treatment Low SES vs. Control Low SES). There is no significant difference between Treatment Low SES and Control High SES. The scale on the y -axis indicates z -scores (i.e., standardized measures) of children’s prosociality. Standardization is conducted using the distribution of wave 2. Error bars show standard errors of the means. *** and ** indicate significant differences at the 1% and 5% level, respectively (two-sided t -tests). Left panel: $N(\text{Treatment Low SES vs. Control Low SES}) = 494$, $N(\text{Control High SES vs. Control Low SES}) = 427$, $N(\text{Treatment Low SES vs. Control High SES}) = 293$, right panel: $N(\text{Treatment Low SES vs. Control Low SES}) = 411$, $N(\text{Control High SES vs. Control Low SES}) = 358$, $N(\text{Treatment Low SES vs. Control High SES}) = 243$.

We now turn to the important question whether the observed effects of parental background and the intervention are short-lived or enduring. Even short and transitory changes in personality may be crucial, e.g., if they occur during critical transition periods. However, finding lasting effects with respect to temporary changes in the social environment would strengthen the relevance and credibility of our findings, and would underscore the potential of childhood interventions in general. To investigate this issue we now refer to our two-year follow-up data (wave 3), see Figure 2 (right panel). The figure illustrates three important findings. First, over the time span of two years, we observe a general increase in prosociality in our sample of elementary school children. The average increase amounts to 43.8% of a wave 2 standard deviation and holds for all groups under study, i.e., irrespective of SES and treatment. To the best of our knowledge this is the first documentation of

an increase in prosociality within children, complementing and supporting previous studies using cross-sectional data.¹⁷ Second, the high-low SES developmental gap in prosociality that we have seen in wave 2 persists to wave 3. The difference in prosociality between Control Low SES and Control High SES in wave 3 is similar as in wave 2 and adds up to a 25.9% of a wave 2 standard deviation ($p < 0.05$, $N = 358$, two-sided t -test). Third, and most importantly, the treatment effect reported in the left panel of Figure 2 turns out to be remarkably robust over time. Two years after the end of the intervention treated children display significantly higher levels of prosocial behavior than children from the control group (21.7% of a wave 2 standard deviation, $p < 0.05$, $N = 411$, two-sided t -test). As a consequence, prosociality in wave 3 does not significantly differ between treated low SES and untreated high SES children ($p = 0.749$, $N = 243$, two-sided t -test). In other words, the developmental gap that was closed in response to treatment remains closed more than two years after the intervention.

3.3 Heterogeneous treatment effects and mechanisms

How did the presence of mentors positively affect formation of prosociality? We hypothesized that mentors provide resources that are critical and scarce in a given family household: They constitute a positive prosocial role model and engage in intensive social interactions with children. To test this intuition we first explore heterogeneous treatment effects and then show that mentors do in fact provide missing resources that appear to be critical.

Table 2 shows OLS regression results. Prosociality of children (Treatment Low SES and Control Low SES) is regressed on a treatment dummy, the two parental background variables prosociality of mothers and intensity of social interaction, and the respective interactions. The results reveal that the treatment is most effective in children whose mothers score low on prosociality. The respective interaction coefficient is significant ($p < 0.05$) and indicates that the treatment effect increases by 19.0% of a standard deviation of prosociality if the mother scores one standard deviation lower on prosociality. A similar effect is found for the interaction of treatment and intensity of mother-child interaction. The treatment effect increases by 17.8% of a standard deviation of prosociality ($p < 0.1$) if the child experiences one standard deviation less intensive social interaction in his or her home environment.

¹⁷See, e.g., Sutter and Kocher (2007), Fehr et al. (2008, 2013) and Bauer et al. (2014).

	Prosociality	
	(1)	(2)
Treatment Dummy	0.307*** [0.092]	0.254*** [0.092]
Prosociality of mother	0.257*** [0.055]	
Prosociality of mother x Treat	-0.190** [0.088]	
Social interaction		0.205*** [0.056]
Social interaction x Treat		-0.178* [0.096]
Constant	-0.134** [0.057]	-0.102* [0.057]
Observations	484	491
Adjusted R^2	0.059	0.036

Table 2: Interaction effects of treatment and parental background. The prosociality measures of mother and child are constructed as equally weighted scores of standardized measures of altruism, trust and other-regarding behavior, respectively. The social interaction measure is estimated as confirmatory factor analysis using items on the frequency of interactive mother-child activities. Prosociality and social interaction measures are standardized. The significant negative interaction terms in columns (1) and (2), “Prosociality of mother x Treat” and “Social interaction x Treat”, indicate that the treatment benefits those children most who experience little social interaction and whose mothers score low on prosociality. Displayed coefficients are OLS estimates with robust standard errors in brackets. ***, **, * indicate significance at 1-, 5-, and 10-percent level, respectively.

These findings show that the mentoring program benefits children whose mothers display relatively low levels of prosociality and therefore lack a prosocial role model, and who experience relatively little intensive social interaction and therefore lack social reinforcement of prosocial behavior. Mentors provide exactly these two resources: First, the main concern of the mentoring program is to foster personal development through encouraging social interactions between child and mentor. These interactions range from visiting the zoo or museum to playing board games or cooking. Second, mentors are potential prosocial role models. The mere fact that they voluntarily spend a considerable amount of time to serve in a mentoring program devoted to support children reveals prosocial motivation. In fact, volunteering in this program is a costly altruistic activity. Further support for the fact that mentors are prosocial role models comes from our survey data. It turns out that mentors score

very high on our prosociality measure. Compared to low SES mothers they score 34.3% of a standard deviation higher on prosociality ($p < 0.01$, $N = 404$, two-sided t -test). Taken these findings together, the mentoring program benefits children by providing resources and stimuli that are both critical and relatively scarce in the particular family environment.

3.4 Robustness checks and alternative treatment estimates

Two potential limitations for the interpretation of randomized intervention studies concern baseline imbalance and selective attrition. An important feature of our study design is that we have collected the measure of interest (prosociality) not only after but also before treatment assignment. We are therefore able to directly address both potential concerns. With respect to baseline balance we can show that the randomization procedure was successful: the pre-treatment measure of prosociality does not differ by treatment status ($p = 0.662$, $N = 590$, two-sided t -test). Similarly, we can rule out selective attrition. The lost to follow-up rates in waves 2 and 3 are neither related to pre-treatment prosociality, treatment assignment nor their interaction (see Table A2).

We also performed various robustness checks regarding our results. In particular, we show treatment effect estimates using inverse probability weighting (IPW) and panel difference-in-differences (diff-in-diff). In Table A3, we estimate treatment effects based on re-weighting the observed data using inverse probabilities of participation in wave 2 and wave 3 interviews. The predicted probabilities result from a Probit model of a binary selection indicator (indicating whether or not wave 2 or wave 3 interviews were conducted) as a function of pre-treatment prosociality and treatment assignment. In Table A4, we present difference-in-differences panel estimates. In sum and consistent with baseline balance and non-selectivity of attrition, these results confirm our main findings, both in terms of effect size and significance.

Mainly due to an unexpected shortage of mentors, not all ITT children have been matched to a mentor (see above). Nevertheless, we focus on ITT estimates since they only rely on the random treatment assignment rather than on specific assumptions about the mentor-mentee matching process implemented by “Balu und Du”. ITT estimates provide conservative treatment estimates and represent a lower bound of the actual treatment effect since not all children selected into the treatment group were actually treated. To ease comparability with other studies and to provide policy relevant information, we also provide local average treatment effects (LATE). In Table A5, we present two-stage least squares (2SLS) estimates using the random

assignment as an instrument for actual treatment. Note that by design control group children could not participate in the program (no always-takers) and therefore the LATE is the average effect on the treated. Given that the matching rate was 73.9%, the LATE estimates exceed the ITT effects by about 35%.

4 Conclusion

This study provides several important insights for understanding the formation of prosociality. First, we document the role of parental environment in terms of SES, social interaction patterns and intergenerational transmission. Second, our panel data allow us to document a general within-subject increase in prosociality for elementary school-age children. Using the panel dimension we also show that the effect of parental background is persistent over time. Third, our main result provides causal evidence on the effect of social environment on prosocial attitudes. The effect is significant both statistically and economically, and remarkably robust over time: The positive treatment effect of the enriched social environment is enduring and also observed two years after the end of the intervention. In terms of heterogeneous treatment effects, we provide evidence that the mentoring program is most effective for children who experience relatively little intensive social interaction in their home environment and whose mothers score relatively low on prosociality. In combination with the fact that the mentoring program provides intensive social interactions and that mentors are particularly prosocial, our findings suggest that the program serves as a substitute for prosocial stimuli in the family environment. All findings are based on a comprehensive measure of prosociality, consisting of the three facets altruism, trust and other-regarding behavior, using different elicitation methods, and statements from different sources.

We believe that our findings are of broad significance. Prosociality pervades human societies and is of fundamental importance at all levels of social interaction. Our findings show that prosociality is malleable and provide insights concerning the effectiveness of early childhood interventions. Investments, such as the mentoring program under study, have the potential to systematically affect character formation and to close developmental gaps arising from differences in the socio-economic family background. This is important in light of increasing social inequalities and the intergenerational persistence of life outcomes (Putnam, 2015; Aizer and Currie, 2014; Currie and Moretti, 2003; Case et al., 2002).

In addition, our study contributes to the debate on cultural evolution. Our results provide evidence in favor of theories of human social behaviors that do not

exclusively rely on genetic causes but on patterns of social interaction such as teaching, imitation and learning in social environments (Boyd et al., 2011; Henrich et al., 2004; Fehr and Fischbacher, 2003; Boyd and Richerson, 1985). Previous support for theories of cultural evolution primarily relies on cross society comparisons. We contribute complementary evidence by showing substantial within-society heterogeneities in prosociality arising from randomly assigned social environments.

References

- Aizer, A. and Currie, J. (2014), ‘The intergenerational transmission of inequality: Maternal disadvantage and health at birth’, *Science* **344**(6186), 856–861.
- Alan, S., Baydar, N., Boneva, T., Crossley, T. F. and Ertac, S. (2013), ‘Parental socialisation effort and the intergenerational transmission of risk preferences’, *IFS Working Paper* **W13/12**.
- Alan, S. and Ertac, S. (2014), ‘Good things come to those who (are taught how to) wait: Results from a randomized educational intervention on time preference’, *Discussion Paper* .
- Algan, Y., Beasley, E., Vitaro, F. and Tremblay, R. E. (2013), ‘The long-term impact of social skills training at school entry: A randomized controlled trial’, *Unpublished manuscript* .
- Almås, I., Cappelen, A. W., Sørensen, E. Ø. and Tungodden, B. (2010), ‘Fairness and the development of inequality acceptance’, *Science* **328**(5982), 1176–1178.
- Bandura, A. (1986), *Social foundations of thought and action*, Englewood Cliffs, NJ Prentice Hall.
- Bandura, A. and Mischel, W. (1965), ‘Modifications of self-imposed delay of reward through exposure to live and symbolic models.’, *Journal of Personality and Social Psychology* **2**(5), 698.
- Bauer, M., Chytilová, J. and Pertold-Gebicka, B. (2014), ‘Parental background and other-regarding preferences in children’, *Experimental Economics* **17**(1), 24–46.
- Becker, A., Deckers, T., Dohmen, T., Falk, A. and Kosse, F. (2012), ‘The relationship between economic preferences and psychological personality measures’, *Annual Review of Economics* **4**(1), 453–478.
- Benenson, J. F., Pascoe, J. and Radmore, N. (2007), ‘Children’s altruistic behavior in the dictator game’, *Evolution and Human Behavior* **28**(3), 168–175.
- Bettinger, E. and Slonim, R. (2006), ‘Using experimental economics to measure the effects of a natural educational experiment on altruism’, *Journal of Public Economics* **90**(8), 1625–1648.
- Bisin, A. and Verdier, T. (2001), ‘The economics of cultural transmission and the dynamics of preferences’, *Journal of Economic Theory* **97**(2), 298–319.
- Boyd, R. and Richerson, P. J. (1985), *Culture and the evolutionary process*, University of Chicago Press, Chicago.
- Boyd, R., Richerson, P. J. and Henrich, J. (2011), ‘The cultural niche: Why social learning is essential for human adaptation’, *Proceedings of the National Academy of Sciences* **108**, 10918–10925.

- Burks, S. V., Nosenzo, D., Anderson, J., Bombyk, M., Ganzhorn, D., Goette, L. and Rustichini, A. (2016), ‘Lab measures of other-regarding preferences can predict some related on-the-job behavior: Evidence from a large scale field experiment’, *IZA Discussion Paper Series* **9767**.
- Carpenter, J. and Seki, E. (2011), ‘Do social preferences increase productivity? Field experimental evidence from fishermen in Toyama Bay’, *Economic Inquiry* **49**(2), 612–630.
- Case, A., Lubotsky, D. and Paxson, C. (2002), ‘Economic status and health in childhood: The origins of the gradient’, *American Economic Review* **92**(5), 1308–1334.
- Cesarini, D., Dawes, C. T., Fowler, J. H., Johannesson, M., Lichtenstein, P. and Wallace, B. (2008), ‘Heritability of cooperative behavior in the trust game’, *Proceedings of the National Academy of Sciences* **105**(10), 3721–3726.
- Cesarini, D., Dawes, C. T., Johannesson, M., Lichtenstein, P. and Wallace, B. (2009), ‘Genetic variation in preferences for giving and risk taking’, *The Quarterly Journal of Economics* pp. 809–842.
- Chetty, R., Hendren, N. and Katz, L. F. (2016), ‘The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment’, *American Economic Review* **106** (4).
- Cooper, D. and Kagel, J. H. (2009), ‘Other regarding preferences: a selective survey of experimental results’, *Handbook of Experimental Economics* **2**.
- Cunha, F. and Heckman, J. (2007), ‘The technology of skill formation’, *American Economic Review* **97**(2), 31–47.
- Currie, J. and Moretti, E. (2003), ‘Mother’s education and the intergenerational transmission of human capital: Evidence from college openings’, *The Quarterly Journal of Economics* **118**(4), 1495–1532.
- Dohmen, T., Falk, A., Huffman, D. and Sunde, U. (2012), ‘The intergenerational transmission of risk and trust attitudes’, *The Review of Economic Studies* **79**(2), 645–677.
- Falk, A., Becker, A., Dohmen, T., Huffman, D. and Sunde, U. (2016), ‘The preference survey module: A validated instrument for measuring risk, time, and social preferences’, *IZA Discussion Paper Series* **9674**.
- Falk, A., Becker, A., Dohmen, T. J., Enke, B., Huffman, D. and Sunde, U. (2015), ‘The nature and predictive power of preferences: Global evidence’, *IZA Discussion Paper Series* **9504**.
- Fehr, E., Bernhard, H. and Rockenbach, B. (2008), ‘Egalitarianism in young children’, *Nature* **454**(7208), 1079–1083.

- Fehr, E. and Fischbacher, U. (2003), ‘The nature of human altruism’, *Nature* **425**(6960), 785–791.
- Fehr, E., Fischbacher, U., von Rosenblatt, B., Schupp, J. and Wagner, G. G. (2002), ‘A nation-wide laboratory’, *Journal of Applied Social Science Studies* **122**, 519–542.
- Fehr, E. and Gächter, S. (2002), ‘Altruistic punishment in humans’, *Nature* **415**(6868), 137–140.
- Fehr, E., Glätzle-Rützler, D. and Sutter, M. (2013), ‘The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence’, *European Economic Review* **64**, 369–383.
- Goodman, R. (1997), ‘The strengths and difficulties questionnaire: a research note’, *Journal of Child Psychology and Psychiatry* **38**(5), 581–586.
- Guiso, L., Sapienza, P. and Zingales, L. (2009), ‘Cultural biases in economic exchange?’, *The Quarterly Journal of Economics* **124**(3), 1095–1131.
- Harbaugh, W. T. and Krause, K. (2000), ‘Children’s altruism in public good and dictator experiments’, *Economic Inquiry* **38**(1), 95–109.
- Heckman, J. J. (2008), ‘Schools, skills, and synapses’, *Economic Inquiry* **46**(3), 289–324.
- Heckman, J. J. (2013), *Giving kids a fair chance*, MIT Press, Cambridge.
- Heckman, J. J., Pinto, R. and Savelyev, P. A. (2013), ‘Understanding the mechanisms through which an influential early childhood program boosted adult outcomes’, *American Economic Review* **103**, 2052–2086.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E. and Gintis, H. (2004), *Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies*, Oxford University Press.
- Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B. and Borghans, L. (2014), ‘Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success’, *OECD Education Working Paper* **110**.
- Knack, S. and Keefer, P. (1997), ‘Does social capital have an economic payoff? A cross-country investigation’, *The Quarterly Journal of Economics* pp. 1251–1288.
- Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U. and Fehr, E. (2005), ‘Oxytocin increases trust in humans’, *Nature* **435**(7042), 673–676.
- Kosse, F. and Pfeiffer, F. (2012), ‘Impatience among preschool children and their mothers’, *Economics Letters* **115**(3), 493–495.
- La Porta, R., Lopez-De-Silanes, F., Shleifer, A. and Vishny, R. W. (1997), ‘Trust in large organizations’, *American Economic Review* **87**, 333–338.

- Lang, F. R., John, D., Lüdtke, O., Schupp, J. and Wagner, G. G. (2011), ‘Short assessment of the big five: robust across survey methods except telephone interviewing’, *Behavior Research Methods* **43**(2), 548–567.
- Müller-Kohlenberg, H. and Drexler, S. (2013), *Mentoring: Practices, Potential Challenges and Benefits*, Nova Science Publishers, New York.
- Ostrom, E., Dietz, T., Dolsak, N., Stern, P. C., Stonich, S. and Weber, E. U. (2002), *The Drama of the Commons*, National Academies Press, Washington D.C.
- Putnam, R. D. (2015), *Our kids: The American dream in crisis*, Simon and Schuster, New York.
- Rao, G. (2013), ‘Familiarity does not breed contempt: Diversity, discrimination and generosity in delhi schools’, *Discussion Paper* .
- Rodríguez-Planas, N. (2012), ‘Mentoring, educational services, and incentives to learn: What do we know about them?’, *Evaluation and program planning* **35**(4), 481–490.
- Skinner, B. F. (1953), *Science and human behavior*, Macmillan, New York.
- Sutter, M. and Kocher, M. G. (2007), ‘Trust and trustworthiness across different age groups’, *Games and Economic Behavior* **59**(2), 364–382.
- Wagner, G. G., Frick, J. R. and Schupp, J. (2007), ‘The German Socio-Economic Panel Study (SOEP)–Scope, Evolution and Enhancements’, *Journal of Applied Social Science Studies* **127**(1), 139–169.
- Zumbuehl, M., Dohmen, T. and Pfann, G. (2013), ‘Parental investment and the intergenerational transmission of economic preferences and attitudes’, *IZA Discussion Paper Series* **7476**.

Appendix

Additional Tables and Figures

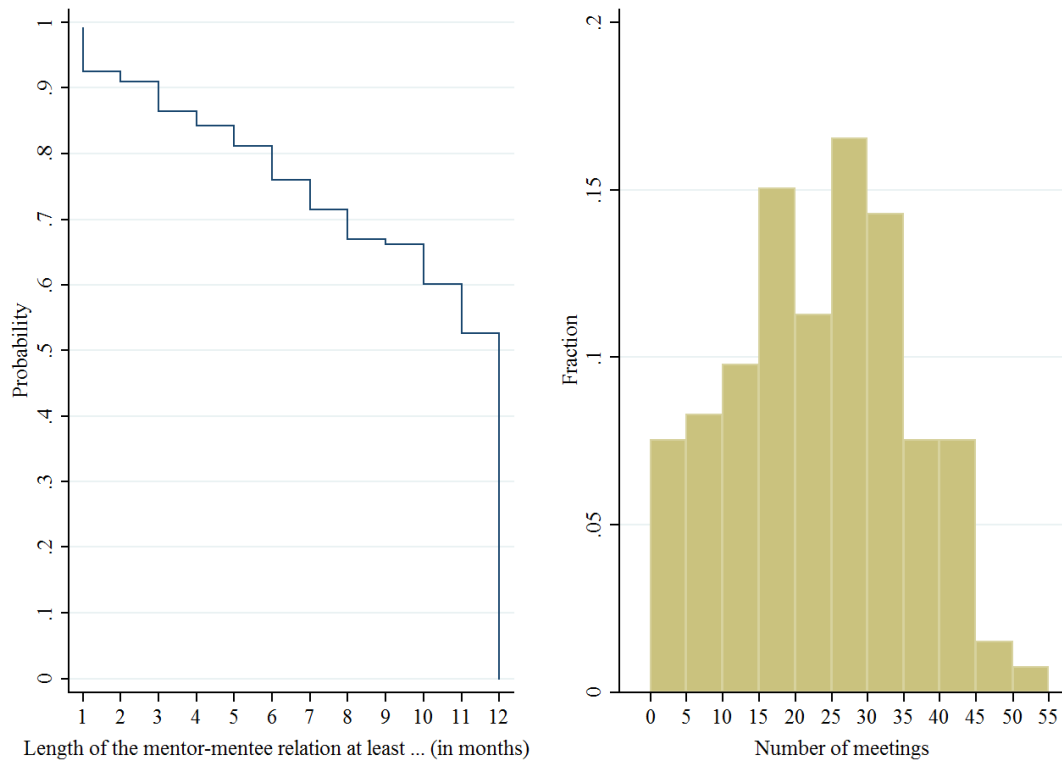


Figure A1: Length of mentor-mentee relations and numbers of meetings. The left panel shows a reverse cumulative probability plot. It indicates that more than 50% of the pairs met for 12 months, more than 80% of the pairs met at least for six months. The mean is 9.3 months. The right panel shows a histogram of the number of meetings. The number of meetings for most of the pairs is centred around the mean of 22.8.



Figure A2: Toys arranged in four categories (example).

How much do you agree to this statement?

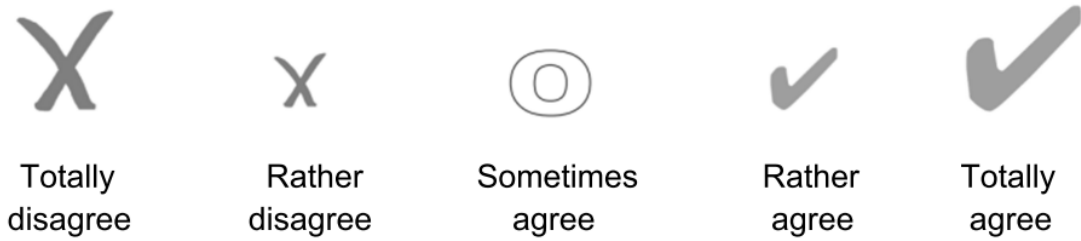


Figure A3: Rating scale used in the children survey.

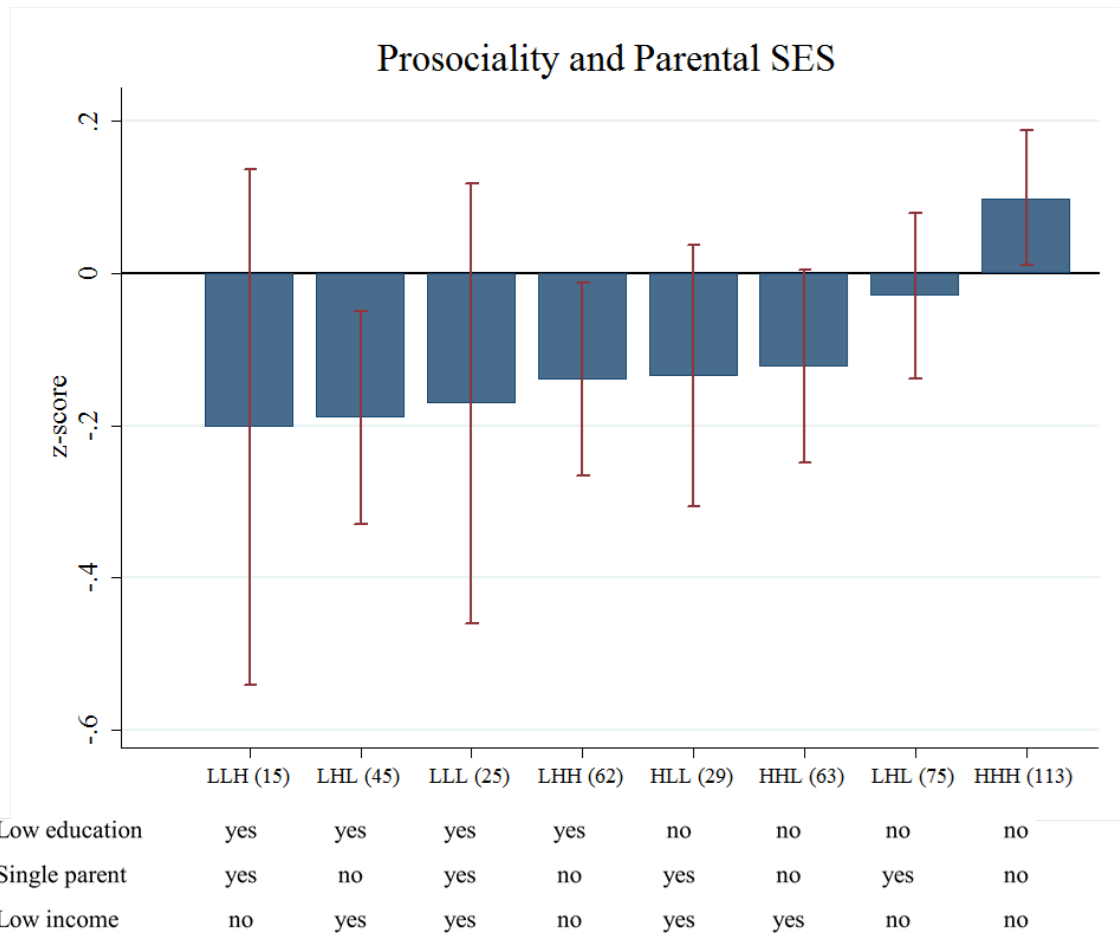


Figure A4: Subgroups of combinations of parental SES characteristics (ordered by means of children's prosociality). Number of observations for every subgroup is displayed in brackets (in sum $N = 427$, Control High SES and Control Low SES). The prosociality measure of children is constructed as equally weighted scores of standardized measures of altruism, trust and other-regarding behavior, respectively. The scale on the y -axis indicates z -scores (i.e., standardized measures) of children's prosociality. Error bars show standard errors of the means (SEM).

	Prosociality	Altruism	Trust	Other-regard. behavior
	(1)	(2)	(3)	(4)
Treatment Dummy	0.294*** [0.093]	0.199** [0.095]	0.235** [0.098]	0.108 [0.097]
Cologne Dummy	0.086 [0.103]	0.070 [0.105]	0.105 [0.104]	-0.016 [0.101]
Constant	-0.191* [0.100]	-0.147 [0.104]	-0.167* [0.101]	-0.039 [0.098]
Observations	494	493	494	492
Adjusted R^2	0.014	0.004	0.008	-0.001

Table A1: Treatment effects for the joint measure and facets of prosociality conditional on city of residence. Coefficients are OLS estimates based on post-treatment data (wave 2), with robust standard errors in brackets. The inclusion of a Cologne dummy (zero if living in Bonn) accounts for the fact that the assignment into treatment was random conditional on city of residence. The prosociality measure is constructed as equally weighted score of standardized measures of altruism, trust and other-regarding behavior. ***, **, * indicate significance at 1-, 5-, and 10-percent level, respectively. A joint F -test rejects the null-hypothesis of a zero treatment effect for the three facets of prosociality (altruism, trust, other-regarding behavior; $p = 0.010$, $F = 3.78$, $N = 491$).

	Assigned to treatment	Lost to follow-up	
	(1)	Wave 2 (2)	Wave 3 (3)
Prosociality (wave 1)	0.009 [0.020]	0.015 [0.020]	-0.007 [0.024]
Treatment Dummy		-0.022 [0.032]	-0.003 [0.040]
PS x Treatment Dummy		-0.021 [0.031]	-0.026 [0.039]
Constant	0.361*** [0.020]	0.172*** [0.020]	0.300*** [0.024]
Restricted to low SES	Yes	Yes	Yes
Observations	590	590	590
Adjusted R^2	-0.001	-0.003	-0.003

Table A2: Checks for baseline balance and selective attrition. In column (1) we test for baseline balance. The dependent variable is one if a child was selected into the Treatment Low SES group and zero if selected into the Control Low SES group. In columns (2) and (3) we test for selective attrition. The dependent variable is one if a child is lost to follow-up, i.e., did not take part in the wave 2 or wave 3 interviews, and zero otherwise. The explanatory variable prosociality is constructed as equally weighted score of standardized measures of altruism, trust and other-regarding behavior. Standardization is conducted using the distribution of wave 2. All measures were collected before the treatment assignment took place. Displayed coefficients are OLS estimates with robust standard errors in brackets. ***, **, * indicate significance at 1-, 5-, and 10-percent level, respectively.

	Prosociality	
	Wave 2 (1)	Wave 3 (2)
Treatment Dummy	0.273*** [0.092]	0.214** [0.104]
Constant	-0.118** [0.058]	0.332*** [0.060]
Observations	494	411
Adjusted R^2	0.015	0.008

Table A3: Treatment effects for prosociality in waves 2 and 3 using inverse probability weighting (IPW). Coefficients are weighted least-square (IPW) estimates with robust standard errors in brackets. The dependent variable in column (1) is the prosociality measure collected in wave 2. The dependent variable in column (2) is the prosociality measure collected in wave 3 (two-year follow-up). Weights are predicted inverse probabilities of not being lost to follow-up. Weights are estimated from a Probit model of a binary selection indicator (indicating whether wave 2 or wave 3 interview was conducted) regressed on baseline prosociality and treatment assignment. The prosociality measure is constructed as equally weighted score of standardized measures of altruism, trust and other-regarding behavior. Dependent variables are standardized. Standardization is conducted using the distribution of wave 2. ***, **, * indicate significance at 1-, 5-, and 10-percent level, respectively.

	Prosociality	
	(1)	(2)
Treatment Dummy	0.254*** [0.095]	0.202** [0.091]
Dummy Wave 2	0.092 [0.057]	0.111** [0.056]
Dummy Wave 3		0.467*** [0.065]
Constant	-0.179*** [0.021]	-0.175*** [0.026]
Individual Fixed Effects	Yes	Yes
Considered waves	1 and 2	1, 2 and 3
Observations	1,084	1,495

Table A4: Difference-in-differences treatment effects. Columns (1) and (2) show the coefficients of individual fixed-effects regressions with clustered standard errors (at the individual level) in brackets. The treatment dummy is zero for control group children in all waves, as well as for ITT children in wave 1. It is one for ITT children in waves 2 and 3. The dependent variable (prosociality) is constructed as equally weighted score of standardized measures of altruism, trust and other-regarding behavior, respectively. It is standardized using the distribution of wave 2. To account for the fact that the assignment into treatment was random conditional on city of residence, time-varying city effects (base: Cologne) are differenced out from the dependent variable, i.e., wave specific city fixed effects are estimated and subtracted. ***, **, * indicate significance at 1-, 5-, and 10-percent level, respectively.

	Prosociality	Altruism	Trust	Other-regard. behavior
	(1)	(2)	(3)	(4)
LATE	0.398*** [0.126]	0.270** [0.129]	0.318** [0.132]	0.146 [0.131]
Cologne Dummy	0.084 [0.103]	0.069 [0.106]	0.104 [0.103]	-0.017 [0.100]
Constant	-0.190* [0.100]	-0.147 [0.104]	-0.166* [0.101]	-0.038 [0.098]
Observations	494	493	494	492

Table A5: Local average treatment effect (LATE) analysis using random group assignment as instrument for actual treatment. 133 of the 180 children we intended to treat were actually matched with a mentor. The main reason for not matching all ITT children was a shortage of mentors. The prosociality measure is constructed as equally weighted score of standardized measures of altruism, trust and other-regarding behavior. All dependent variables are standardized. Coefficients are two-stage least-squares (2SLS) estimates using random assignment as an instrument for actual treatment, robust standard errors in brackets. ***, **, * indicate significance at 1-, 5-, and 10-percent level, respectively. A joint F -test rejects the null-hypothesis of a zero treatment effect for the three components of prosociality (altruism, trust, other-regarding behavior; $p = 0.010$, $F = 3.78$, $N = 491$).

Experimental protocols (translations from German)

Binary dictator game

“Have a look at these paper stars I’ve got here.”

⇒ Show the stars to the child.

“We now want to play a game in which you can win stars. Later on you can exchange these stars for a present. The more stars you win the better will the present be you are going to get afterwards.”

⇒ Take the child to the boxes with the presents; explain that the presents become better/bigger going from box 1 to box 4 and that one needs more stars to get a better/bigger present.

“You get some stars and you can give stars to another child as well. Here I have two paper bags. One is for you. Let’s write your name on it. The other bag is for another child that is of your age and also lives here in Bonn (Cologne). But both of us do not know exactly who that other child is. So we will not write a name on the second bag. The bags are for the stars that you and the other child are going to get in this game.”

“Here are two sheets of paper.”

⇒ Place both sheets of paper in front of the child (right and left).

“On each sheet you can see two circles with arrows. On both sheets one arrow is pointing to the bag of the other child and the other arrow is pointing to you and your bag. On the first sheet I place one star in the circle that is closer to the bag of the other child and I place one star in the circle that is closer to you.”

⇒ Place the stars in the circles now.

“On the second sheet I place two stars in the circle that is closer to you. In the other circle that is pointing to the bag of the other child I place no star. You can now choose one of those sheets. If you choose this one, this arrow is pointing at you. That means you are getting what is placed in this circle, one star.”

⇒ Point to the respective circles.

“The other circle is pointing to the bag of the other child. This means that the other child is getting what is placed in this circle, one star. If you choose this sheet, you get two stars and the other child gets no star.”

⇒ Point to the first sheet.

“If you choose this sheet, what will the other child get? And what will you get?”

⇒ Point to the second sheet.

“If you choose the second sheet, what will the other child get? And what will you get?”

Check of understanding of the rules

⇒ Repeat rules up to three times.

⇒ If the child has difficulties answering the control questions explain the rules again. Ask the questions again. If the child does not understand the rules at all, play the game nonetheless so that the child will not be disappointed. Do not play the game if the child is frustrated due to the lack of understanding and does not want to play.

“Okay, which sheet do you choose?”

⇒ After the decision, put the stars into the respective bags.

⇒ **Remove both bags.** Place the bag of the participating child nearby. Put the bag of the other child away.

“I will make sure that the other child gets something nice for the stars.”

Continuous dictator game A

“Now we will again play a game in which you can win stars. Later on you can exchange the stars for a present. The more stars you win the better your present will be. You will get some stars and you can give stars to another child as well. Here I have two paper bags. The first one is for you. Let’s write your name on it. The second bag is for another child that is of your age and lives in this area but not in Bonn (Cologne). Both of us do not know who that other child is. Therefore we do not write a name on the second bag. The bags are for the stars that you and the other child will get in this game.”

⇒ Put both bags side by side on the table in front of the child.

“Look! Here are 6 stars.”

⇒ Put the stars in front of the child between the two bags.

“Now you can decide how many stars you put on your bag and how many you put on the bag of the other child. The stars on your bag (point to the bag) are for you. The stars on the other bag are for the other child from another town nearby. I will make sure that the other child gets something nice for the stars.

You can decide how you want to divide the stars. You can split the stars or you can put all stars on one bag. How do you want to divide the stars? Now, please put all 6 stars on the bags in a way you would like to have it.”

⇒ Child puts the stars on the bags.

“Okay. How many stars do you get? And how many stars does the other child get?”

⇒ In case the answers are not correct: explain the correct answer and ask for new suggestion. In case both answers are correct:

“Fine. Let’s now put the stars in the bags.”

⇒ Remove the bags.

Continuous dictator game B

“We will now play a similar game in which you can win stars. You will get some stars and you can give stars to another child again. Here are two bags. We will write your name on the bag that is for you. The other bag is for another child that is about your age. This child lives in Africa which is very far away. The child cannot live with its parents, e.g., because they are too poor, ill or perhaps even dead. This child has only a few things to play with. We both do not know the name of the child therefore we do not write a name on the other bag. Again the bags are for the stars that you and the other child will receive.”

⇒ Put both bags side by side on the table in front of the child.

“Look! Here are 6 stars.”

⇒ Put the stars in front of the child between the two bags.

“Now you can decide how many stars you put on your bag and how many you put on the bag of the other child. The stars on your bag (point to the bag) are for you. The stars on the other bag are for the other child that lives in Africa without his or her parents and only with a few toys. I will make sure that the other child gets something nice for the stars.

You can decide how you want to divide the stars. You can split the stars or you can put all the stars on one bag. How do you want to divide the stars? Now, please put all 6 stars on the bags in a way you would like to have it.”

⇒ Child puts the stars on the bags.

“Okay. How many stars do you get? And how many stars does the other child get?”

⇒ In case the answers are not correct: explain the correct answer and ask for new suggestion. In case both answers are correct:

“Fine. Let’s now put the stars in the bags.”

⇒ Remove the bags