

# **The Prosocial Roots of Children's Developing Morality**

Inaugural-Dissertation

zur Erlangung des Doktorgrades der Philosophie

an der Ludwig-Maximilians-Universität München

vorgelegt von

Anja Kaßecker

aus München

2020

Referent/in: Prof. Dr. Beate Sodian

Korreferent/in: Prof. Dr. Marco F. H. Schmidt

Tag der mündlichen Prüfung: 18.05.2020

## Acknowledgments

There are a number of people to whom I want to express my deepest gratitude. First, my gratitude goes to my supervisor Prof. Dr. Marco Schmidt, who shared his immense intellectual knowledge and enthusiasm for the research ideas with me and supported me in every aspect of my scientific development. I would like to thank Prof. Dr. Beate Sodian for agreeing to supervise my thesis, and for her great guidance and thoughtful comments on my projects.

My deepest thanks go to Veronika Eisenschmid, Emmily Fedra, Julia Friedrich and Stephan Verschoor for the shared experiences and the unforgettable time together. I gratefully acknowledge Mitho Müller, Alexandra Patin and Katrin Preckel for their helpful comments and valuable input to my work. Furthermore, I would like to thank the great people from the research colloquium at the Department of Developmental Psychology of the LMU and especially Prof. Dr. Markus Paulus for the opportunity to present my work and give me helpful feedback.

Next, I thank my former students Laura Dietrich, Ramona Engelhardt, Justine Lazar, Helena Marx and Teresa Tänzer for their great work and help in collecting data for my studies. Without their support, the project would certainly not have gone so well. My appreciation also goes to the caregivers, parents and children who participated in my studies and without whom it would not have been possible to conduct this research.

My special thanks go to my beloved family and my friends, especially Ingrid, Rainer, Monika and Renate for their everlasting faith in me and their unconditional love and support. I am very grateful for that! Finally, my deepest gratitude goes to my husband Mathias and my little daughter Karla for their enduring patience, love and support, especially during the final stages, when they visited all the museums in Munich so that I could focus on my writing at home.

# Table of Contents

<b>Abstract.....</b>	<b>6</b>
<b>1. General Introduction .....</b>	<b>10</b>
1.1. Prosociality .....	14
1.2. Morality .....	17
1.3. The Relationship between Prosociality and Morality .....	20
1.4. The Early Development of Human Morality: An Empirical Overview .....	25
1.5. Focus of the Dissertation and Methodological Approach .....	32
<b>2. Study 1: 18-month-old Infants Differentiate between Moral and Conventional Violations .....</b>	<b>36</b>
2.1. Introduction .....	36
2.2. Method.....	40
2.3. Results .....	44
2.4. Discussion.....	47
<b>3. Study Set 2: The Moral Roots of Fairness: Young Children’s Enforcement of Fairness Norms Is Related to Their Prosocial Behavior .....</b>	<b>50</b>
3.1. Introduction .....	50
3.2. Experiment 1.....	54
3.3. Experiment 2.....	62
3.4. Experiment 3.....	68
3.5. General Discussion .....	75

<b>4. Study 3: Decategorization Leads Preschoolers to Treat In-Group and Out-Group Members Equally.....</b>	<b>79</b>
4.1. Introduction .....	79
4.2. Method.....	83
4.3. Results .....	87
4.4. Discussion.....	88
<b>5. General Discussion.....</b>	<b>92</b>
5.1. Summary.....	93
5.2. Developmental Origins of Morality.....	95
5.3. The Relationship between Prosociality and Morality .....	98
5.4. Scope of Morality .....	100
5.5. Limitations and Implications for Future Research .....	102
5.6. Conclusion .....	107
<b>References.....</b>	<b>108</b>
<b>List of Figures.....</b>	<b>129</b>

## Abstract

According to many scholars, prosociality, in particular altruism and empathic concern, is considered an important motivational factor both in adulthood and in the development of morality (Batson & Shaw, 1991; Jensen et al., 2014; Nichols, 2004; Roughley & Bayertz, 2019). So far, a large number of studies have addressed the development of children's first-party prosociality and their third-party understanding of moral norms separately. In particular, there is much evidence that during the second year of life, young children develop empathic concern and sympathy for others in need in prosocial situations (Hepach, 2017; Hepach et al., 2012). Moreover, recent findings suggest that 18-month-old infants already show some rudimentary forms of norm understanding in at least dyadic conventional situations. This rudimentary norm understanding is interpreted as second-personal normative expectations (Schmidt et al., 2019). Finally, 3-year-old children not only have descriptive expectations about morality, but also normative ones as suggested by their enforcement of moral norms as unaffected bystanders (Rakoczy et al., 2016; Rossano et al., 2011; Vaish et al., 2011). However, the interrelation between prosociality and morality, in particular the prosocial motivational source of the early sense of morality remains unclear.

This thesis aimed to investigate the developmental origins of morality in young children. In particular, it examines the relation between the two main aspects of uniquely human cooperation – prosociality and morality – from a developmental perspective. These two aspects are of particular importance, not only because they each play a key role in maintaining the unique human capacity for large-scale cooperation (Tomasello, 2016, 2018) but also because of their close relation to each other (Batson, 2010; Batson & Shaw, 1991; Nichols, 2004). The present thesis therefore focused on three guiding questions that are essential for the ontogeny of morality and its relation to

young children's prosocial (altruistic) motivation to understand, adhere to, and enforce moral norms: (1) What are the developmental origins of morality? (2) What is the underlying prosocial motivation for children's normative appreciation of morality? (3) What is the scope of morality?

Study 1 investigated the developmental origins of morality in 18-month-old infants. A novel eye-tracking paradigm (anticipatory looking, pupil dilation) was used to examine whether infants differentiate between prototypical moral (harmful) and conventional (harmless) violations. In a between-subjects-design, children watched the same video clip whose audio stream differed according to condition. In the first two (conventional) conditions, an instructor told an observer to destroy a picture with a particular tool chosen from two available tools (tool A: conventional violation condition; tool B: no violation condition). In the moral violation condition, the instructor forbade the observer to destroy the picture at all. In all three conditions, the observer then grasped tool B and destroyed the picture, which led to three different (violation) situations. Infants differentiated between two types of conventional norm situations in their anticipatory looking. Moreover, they showed a larger relative increase in pupil dilation in response to a moral violation than to a conventional violation. These findings suggest that 18-month-old infants have third-party descriptive expectations about the distinction between conventional and moral violation situations. Moreover, they provide the first evidence that empathic concern may be a decisive capacity for the distinction between these two violation situations.

Study set 2 looked at the underlying prosocial motivation for the appreciation of morality as a normative notion in 3-year-old children. In three experiments, children were given a third-party fairness task (which varied across experiments) and two different prosocial tasks. To investigate whether the children have a proper norm understanding of fairness by looking not only at norm adherence, but also at norm

enforcement, a spontaneous protest paradigm was used. In Experiment 1, children protested and corrected unequal (but not equal) allocations, suggesting a normative understanding of third-party fairness. Experiment 2 assessed whether children's normative expectations about fairness have a moral (authority-independent) dimension. To do so, children observed a distributor who followed (unequal condition) or violated (equal condition) an authority's command to allocate resources unequally. Again, despite the authority's dictate to act unequally, children protested more against unequal versus equal allocations. In Experiment 3, results show that children enforced fairness norms by altruistic punishment in the sense of restorative justice. While in Experiment 1 and 2 I found a positive relation of protest behavior and emotional sharing (empathic concern), in Experiment 3 children's altruistic punishment was associated with their own costly sharing behavior (altruism).

Finally, in Study 3, I explored the scope of morality (looking at equal treatment) in 5-to 6-year-old children in a typical intergroup context. Here, I investigated whether decategorization – a candidate mechanism to overcome in-group bias by emphasizing the individual person – would lead preschoolers to treat in-group and out-group members equally when sharing resources in a dictator game. I found that preschoolers shared more resources with an in-group than with an out-group recipient when social category membership was emphasized. When individuating information was emphasized (decategorization), however, children shared the same with in-group and out-group individuals.

Taken together, the empirical studies of this dissertation provide a novel overview of the prosocial roots of children's developing morality. In particular, the present findings suggest that (1) the ability to feel sympathy may be critical for the development of the moral-conventional distinction and that 18-month-old infants, at minimum, have third-party descriptive expectations about that distinction. (2) The



ontogeny of fairness norms can be characterized as moral in that it is associated with 3-year-old children's developing concern for the welfare of others in different contexts.

(3) The presentation of out-group members as individuals may be a powerful tool to reduce in-group bias and to foster equal treatment (an important moral category) of in-group and out-group members in 5- to 6-year-old preschool children.

# 1. General Introduction

I would like to begin with the following story that once happened to me: my husband had broken his arm and could not help with our weekly groceries. He could carry little things, but lifting a crate of water was impossible. When we walked back from the grocery store to our apartment, he had a small carrier bag with him and I carried the crate of water. It was so heavy that I doubted whether I could carry it all the way home. I almost wanted to leave the heavy load behind. Then, suddenly out of nowhere a helping hand grabbed the other side of the crate from behind. I wondered what was going on and turned around. A woman looked at me compassionately and carried the purchase together with me without saying much until we arrived at the entrance of our house. She briefly said goodbye and continued on her way. Something that amazed me and which I will remember forever is her great helpfulness towards a stranger. We did not know each other, and we knew that we would not meet again after the event, thus she could not expect me to “pay her back” in any way. Therefore, I wondered, what was the motivation behind her helpful behavior? Was it a conventional reason, because we have learned to be helpful, in meeting some normative standards? Was it for the sake of maintaining a good reputation, in the hope that other people would notice her as a helpful person? Or because she just wanted to feel better and didn't have to endure the suffering? Did she help because she cared about my well-being and therefore it was really about morality? This leads me to another question: why should I care about her motivation at all, if she helped me in the end?

From a philosophical point of view, there are various assumptions as to whether the outcome of an action (consequentialist ethics) or an action itself (deontological ethics) is considered morally relevant, as per Kant, Bentham or Mill, to name a few (Alexander

& Moore, 2016). In contrast to the consequentialist ethics (maximization of the "good": e.g., helping leads to good feelings for both the helper and the person in need), deontologists believe that an action must be consistent with moral principles, regardless of what is the best outcome for the most people. Therefore, moral norms must be followed by every moral agent and not simply maximized. In this sense, "the right" should always take precedence over "the good". Thus, if helping is a moral norm, it should be applied to all moral agents, regardless of whether it brings the greatest benefit to everyone involved.

In this thesis, however, I will focus more on the underlying psychological motivation behind moral agency. Daniel Batson, who studied altruistic helping, argued as follows: being a moral person depends on the interest the other person pursues to help. Consequently, his main interest was to find out whether the help was out of selfish or altruistic motives. In his empathy-altruism hypothesis, he stated:

Reducing the need of a person for whom one feels empathy is likely to enable the helper to gain social and self-rewards, avoid social and self-punishments, and reduce feelings of personal distress. But [...] feeling empathy for the person in need evokes motivation to help in which these benefits to self are not the ultimate goal of helping; they are unintended consequences (Batson & Shaw 1991, p. 114).

In line with these considerations, the ultimate goal, pure help, cannot be interpreted as morally right or wrong, but the motivation behind the helping behavior can. In this context, a morally relevant (altruistic) act means that the ultimate goal is to increase another's well-being. Selfish motives, on the other hand, should be subordinated or excluded. Importantly, other-regarding concerns, which the moral sentimentalists David Hume and Adam Smith described as sympathy for the other in need, play a decisive role

in acting out of the right moral reasons (Kauppinen, 2018). To get back to the woman who helped me: she looked compassionate and there were no other obvious selfish motives. So perhaps her help was motivated by an altruistic (moral) motivation.

What about children? Do they – compared to other animal species – possess the capacity for prosocial motivation and moral agency? Many animal species are capable of cooperative (prosocial) behavior, such as fish, ants and bees. Even our closest living primate relatives, the chimpanzees, show some restricted forms of collaboration. Could one assume that their behavior is moral? Their main interest to cooperate is to maintain their own genes, therefore they cooperate with genetically related others, but cooperation with unrelated others is unlikely to occur (Kelly & Thibaut, 1979; Nowak, 2006; Trivers, 1971). With nonrelatives, collaboration is especially useful when individuals are interdependent, as happens to humans when maintaining a peaceful group life in the cultural community. Indeed, human cooperation is unique in the animal kingdom. From an evolutionary point of view, compared to other animal species, humans are ultra-social beings who even cooperate with nonkin (de Waal, 2008; Hoffman, 2000; Tomasello, 2016, 2018). According to the cooperation theory of moral development (Tomasello, 2016, 2018), there are two ontogenetic steps in children's early moral development: beginning with second-personal prosociality in the second year of life which turns to norm-based morality around the age of three. This suggests that even our children can already engage in moral behavior. But there is much evidence that children are not only able to act cooperatively in a second-personal context (joint action, joint intentionality), but also understand morally relevant situations from a third-person perspective, at least in a descriptive sense earlier than is currently assumed, even starting in their second year of life (Schmidt et al., 2019). Thus, this thesis examines specifically the relation of the two main aspects of the uniquely human cooperation – prosociality and morality – from a developmental point of view. In particular, I

investigate what constitutes the prosocial (altruistic) motivation of children to understand, adhere to, and enforce moral norms.

In Chapter 1, I introduce prosociality, by focusing on altruism and empathic concern. I then give an overview of the characteristics and the nature of morality, regarding morality as a social normative phenomenon. Next, I describe the relation between prosociality and morality, in particular the prosocial motivational aspects of morality. Thereafter, I give an empirical overview of the development of both aspects. Chapters 2-4 present my empirical studies. Finally, Chapter 5 closes with a general discussion of the current findings and provides some potential limitations and future directions.

## 1.1. Prosociality

Prosocial behavior is defined as voluntary behavior that is intended to benefit others, such as sharing, helping, or comforting another (Eisenberg et al., 2015). Prosociality is not only important for interactions between members of a particular group, but also a decisive motivation in intergroup contexts, in which the ultimate goal is not one's own welfare or the welfare of other specific group members who benefit, but instead the welfare of the group as a whole (Batson, 2010; Dawes et al., 1988; Tajfel, 1982; Turner et al., 1994).

In addition to situational factors, scientists have always been interested in whether specific personality traits are responsible for human prosociality (Batson, 2010; Batson & Shaw, 1991; Cialdini et al., 1987; Eisenberg et al., 2015; Hoffman, 2000; Jensen et al., 2014; Nichols, 2004). In summary, these different explanations can be traced back to two separate considerations: (1) people act prosocially on the basis of a voluntary intrinsic motivation, out of the genuine interest in the well-being of others and through internalized values, goals and rewards, or (2) people need external incentives and act predominantly based on hidden selfish strategies such as avoiding punishment and social exclusion. In the following paragraph, the first consideration, defined as psychological altruism, is discussed (Eisenberg & Mussen, 1989).

According to Batson (Batson, 2010; Batson & Shaw, 1991), psychological altruism is a motivational state with the ultimate goal of increasing another's welfare, and is therefore contrary to psychological egoism, a motivational state with the ultimate goal of increasing one's own welfare. Following the assumption that altruism is primarily about other-regarding concerns, many psychologists (e.g., Andreoni, 1990) also describe the motivation as (impurely) altruistic when the prosocial act is completed because a person is either intrinsically rewarded (the warm glow effect), avoids feeling

guilt, or feels less aversive aroused caused by witnessing someone else's suffering. In line with such considerations of both egotistical and altruistic motivations behind prosociality, many philosophers such as Thomas Hobbes, Immanuel Kant, Friedrich Nietzsche or Jeremy Bentham argued against the possibility of a pure form of altruism (Batson, 2010; Waldmann et al., 2012). Altruism strongly contradicts the principle of homo economicus, which assumes that people only engage in prosocial behavior if they receive something in return (Cialdini et al., 1987). This consideration does not necessarily have to be material; it can also refer to affective (experiencing pleasant emotions after receiving help) or social (receiving thanks or respect from other people) considerations. According to Batson (Batson, 2010; Batson & Shaw, 1991), impure altruistic motivation is classified as a relatively subtle form of egoism and therefore is not seen as an altruistic act. However, he does not deny the existence of pure altruistic motives, like many other philosophers do, such as David Hume, Adam Smith or Jean-Jacques Rousseau (Waldmann et al., 2012).

An important role for prosocial behavior is given by empathy. Empathy-related processes sometimes motivate altruism but, depending on their nature, they can also motivate self-focused behavior (Batson & Shaw, 1991; Eisenberg et al., 2015; Eisenberg & Fabes, 1990; Eisenberg & Mussen, 1989; Hoffman, 2000). While empathy is defined as the mere ability to cognitively and emotionally represent other thoughts and feelings (e.g., correctly identifying another person's internal state or feeling what the other person feels), empathic concern and personal distress often result from empathy. The cognitive ability to correctly perceive another person's internal state is defined as empathic accuracy (Ickes, 1993). Feeling as another person feels is referred to as emotional contagion, or affective resonance (de Waal, 2008; Hatfield et al., 1994). Empathic concern is associated with altruistic, other-oriented emotions, in that it involves feeling for the other person in need. It includes emotions such as sympathy,

compassion or tenderness. Contrary to this, personal distress is defined as a self-focused, aversive emotional response to another's distress and is associated with egotistic motivation. Three possible egotistic prosocial motivations have been identified: (1) aversive-arousal reduction (reducing negative emotions caused by witnessing another in need), (2) punishment avoidance and (3) reward seeking. Taking two different emotional states together, the empathy-altruism hypothesis (Batson & Shaw, 1991) claims that only concern felt for a person in need produces altruistic motivation to relieve that need. Therefore, personal distress may also be positively related to prosocial behavior, but primarily when there is no easier way to reduce one's own distress other than helping another.

Finally, there are other definitions of altruism which differ from Batson's psychological-motivational notion. First, from an economic point of view (Kelly & Thibaut, 1979; Nowak, 2006; Trivers, 1971), (behavioral) altruism is equated with costly sharing or helping behavior (actions that benefit others and not oneself) regardless of underlying motives. Secondly, evolutionary altruism differs from psychological altruism, in that evolutionary altruism exhibits behavior that reduces reproductive fitness (Buss, 2016; Dawkins, 1976).



## 1.2. Morality

Before starting to describe morality as special social normative phenomena, certain key features of normativity should be defined (for an overview, see Schmidt & Rakoczy, 2018, 2019). Normative phenomena can usually belong to either practical or epistemic normativity (Engel, 2011; Schmidt & Rakoczy, 2018). Practical norms refer to and give reason for human actions. They are therefore part of human cultural practices and values. Epistemic norms describe human beliefs. They give reasons to believe certain things and are thus fundamental to our theoretical reasoning, cultural knowledge and understanding of truth. The focus here is on practical norms and in particular moral norms as social normative phenomena.

First, when phenomena in social interactions are classified as normative, then typically the observed actions are compared with an ideal standard of that action. This leads to standards of correctness, with which the concrete action can be assessed as right or wrong (Hechter & Opp, 2001; Schmidt & Rakoczy, 2018). Norms are valid in both general and agent-independent ways (Nagel, 1986). This implies that everyone involved in the social practice (including oneself) is expected to adhere to these norms under the same conditions. Norms have a binding force and authority over all members of a particular community (Korsgaard, 1996; Roughley & Bayertz, 2019; Schmidt & Rakoczy, 2018). Thus, norms – as learned behavioral standards – are applied, regulated and created among members of our cultural community, which ultimately leads to peaceful group life and large-scale cooperation even with nonrelated others. Important for a proper norm understanding is not only the acceptance of certain rules, but also the development of strategies to enforce the rules by normative agents, for example through sanctions or rewards (Brandom, 1994; Schmidt & Rakoczy, 2018, 2019; Searle, 2010). At the least, third-party norm enforcement is considered a key mechanism in the

evolution of human cooperation: by following a deontic modality, humans have normative expectations about how people “ought” to act in certain situations (Chudek & Henrich, 2011; Fehr & Fischbacher, 2004). Normative expectations are distinct from descriptive expectations about how people will behave. Descriptive expectations are thought to have a mind-to-world direction of fit (similar to epistemic states describing how reality is), while normative expectations are thought to describe a world-to-mind direction of fit (similar to volitional states and desires; Christen & Glock, 2012; Schmidt & Rakoczy, 2018, 2019; Searle, 2010). Finally, norms typically apply in one context but not in another. Examples of context-relativity are conventional agreements, such as traffic rules (driving on the right or left side of the road) or dress codes (wearing black at a funeral). However, moral norms are also context-dependent in the sense that there is a hierarchical order as to which of these norms appears to be most important in a particular situation (Schmidt & Rakoczy, 2018). Now that the key features of social normative phenomena have been described, I will turn my focus to moral norms.

Moral norms are usually described as prototypical for social norms: they are defined as the understanding of others’ welfare, justice and rights, and thus have a non-arbitrary character due to the serious consequences of the norm violations (e.g., harm of another person ; Nichols, 2004; Schmidt & Rakoczy, 2019; Turiel, 1983). According to social domain theory (Turiel, 1983, 2006), people distinguish moral norms from arbitrary conventional norms (e.g., agreements, customs or rules) using the following criteria: both norms are generally permissible, but only moral norms should be adhered to under all circumstances (generalizability), they are valid even if an authority orders a violation of the norm (authority contingency) and violations are considered to be more severe (seriousness) than when disregarding conventions. Thus, moral norms carry the most normative weight and violations are seen as more deserving of punishment than when disregarding conventions.

Fairness norms as part of moral norms are, from many points of view, an important factor in the maintenance of cooperation and morality (Chudek & Henrich, 2011; Fehr et al., 2008; Nowak, 2006; Tomasello, 2016). On the psychological level, acting out of fairness principles is contrary to acting out of selfish motives, as it is potentially motivated by concern for the well-being of others (psychological altruism). At the evolutionary level, fairness as a cost-benefit calculation is often associated with costs to oneself and benefits to others (biological altruism). A number of different forms of fairness are distinguished: distributive justice, procedural justice and interactive justice (Feinberg, 2017; Kane, 1996). Depending on the situation, different principles of distributive justice are regarded as fair (Feinberg, 2017; Kane, 1996; Rawls, 1999). According to the principle of merit, a fair distribution should be based on individual contributions. For example, the person who has shown more effort will receive more goods. The principle of equality means that all relevant persons receive an equal distribution. A good example of this is Rawl's thought experiment, "Veils of Ignorance" (Rawls, 1999), in which individual status (including one's own identity) and contributions to the community are difficult to identify because all group members are considered anonymous. As a consequence, people tend to treat everyone equally in the end because nothing is known about the other (or oneself). Finally, the need principle explains that fairness is measured by whether a person's specific needs are taken into account when making a decision. For example, the person who is poor or hungry receives more of the goods.

### 1.3. The Relationship between Prosociality and Morality

There are a variety of different theories about the relationship between prosociality and morality (for an overview, see Roughley, 2018). Due to the thematic focus of my work I will address only three explanatory approaches. In the following I will first summarize the cooperation theory of morality (Tomasello, 2016, 2018). Then I will compare it with theories that are in line with moral sentimentalism (Charles Darwin, Adam Smith, David Hume, Shaun Nichols, Daniel Batson), which focus on sympathy as an important prosocial motivation for morality, as well as theories that focus more on the broader emotional components of moral judgments as alternative explanations (Jonathan Haidt, Joshua Greene).

Tomasello believes that the development of morality must take into account not only that altruistic behavior acts as a driving force, but that morality primarily requires "mutual respect" between individuals, which is based on self-other equivalence regarding role fulfillment in collaborative activities (Tomasello, 2016). Most importantly, he sees this capacity as the genesis of the key normative concept of morality, which develops in three steps (Roughley, 2018; Tomasello, 2016). The first step is the "you > me" formula, which represents early humans' (and great apes') morality of sympathy. That is, the pre-moral capacity of prosociality and in particular altruism and empathic concern for the welfare of others between kin and friends, where (biological) costs are not too great. The second step is the "you = me" formula, which represents, in particular, the fundamental capacity of early humans' morality of fairness. That is, a kind of impartiality, deriving from the recognition of self-other equivalence, in which every partner deserves equal rights. However, the capacity of mutual respect and deservingness is limited to a partner in second-personal, interdependent, collaborative interactions. Finally, the third step is the "we > me" formula, which

represents the full-blown capacity for morality in an agent-neutral sense. In line with that, the key new attitudes of joint agency are “respect”, “resentment” and “guilt”, and joint agency is transformed from small-scale to large-scale interactions, from local “responsibilities” and “deservingness” into fully fledged moral obligations (Roughley, 2018; Tomasello, 2016, 2018). When looking at the ontogeny of morality in humans, children are not only passive recipients of social information, but rather they are shaped by the social world in which they live. They take on an active role in learning when and how to cooperate through interaction with parents, peers and communities. Tomasello claims that the capacity of shared intentionality (which only humans seem to possess) is the key ability not only for the development of social cognition but also for morality (Tomasello, 2016, 2018; Tomasello et al., 2005, 2012). In his interdependence hypothesis (Tomasello et al., 2012) he defines shared intentionality as the capacity to share intentions in a “we”-mode (to think and act as “we” rather than just “I”), which develops in two steps from second-personal joint intentionality to group-based collective intentionality. According to the cooperation theory of moral development (Tomasello, 2018), morality is a special form of cooperation that also develops in two ontogenetic steps: from second-personal (pre-moral) cooperative and collaborative behavior to fully fledged objective moral capacities. Beginning with second-personal actions, collaborative behavior takes place in the direct interaction between two persons, such as a child and an adult acting together towards a shared goal (e.g., Warneken & Tomasello, 2007). At around the age of three, children’s cooperative behavior takes a normative turn, from “[...] what the child wants to do and wants others to do to what she and others ought, should, or must do” (Tomasello, 2018, p. 249). At this age, children begin to understand that norm agreements are not only directly linked to a particular person in second-personal actions, but should be applied independently to each member of a community (agent-neutral norm understanding). Children are now

able to understand that norms have a binding force for all members of the group and that we as a group are responsible for these norms (Tomasello, 2016).

Another approach of looking at prosociality and morality is offered by evolutionary theories influenced by moral sentimentalists, such as Adam Smith and David Hume (Kauppinen, 2018; Roughley, 2018). In contrast to Tomasello's point of view, David Hume and Adam Smith argued that these other-oriented motivations (empathic concern, sympathy) especially are the core capacity for altruistic motivation and morality (Batson, 2010; Kauppinen, 2018; Maibom, 2017; Waldmann et al., 2012). Clearly influenced by Adam Smith's theory of morality, Darwin assumes three "social instincts", which are decisive for the development of human morality: sympathy, the development of the ability to think, and language (Darwin, 1871; Roughley, 2018). Thinking and language enable comparisons over time and across individuals, but Darwin saw sympathy as the crucial source of morality. On the one hand, this "other-oriented" motivation enables early humans to be receptive to the well-being (and suffering) of their conspecifics, and on the other hand, it enables them to judge whether the moral judgments of others are to be regarded as right or wrong. Darwin associated sympathy with an initial immature moral capacity. Only the later component "reasoning" led to the understanding that moral values are not only valid within the group, but universally (for other nations/animals). With regard to Darwin, however, the empirical verification of his hypothesis was largely unexplored. Nevertheless, there is much evidence today that the ability to distinguish between moral and conventional norms depends not only on a normative agreement determining which actions are prohibited, but also on an affective mechanism (Nichols, 2002, 2004; Turiel, 1983, 2006). In line with that, most theories of moral reasoning acknowledge that emotions are an important part of moral judgments (Waldmann et al., 2012). Even Kant in his more rationalist theory of morality, claims that moral judgments are typically

accompanied by moral feelings (Kant, 1959; Maibom, 2017; Waldmann et al., 2012). There is much empirical evidence that affects are deeply intertwined with moral reasoning and lead to strong emotional reactions as part of moral judgments (Decety & Howard, 2013; Eisenberg et al., 2014; Jensen et al., 2014; Killen & Smetana, 2015; Nichols, 2002; Smetana et al., 2014). Not only mere empathy (feeling what the other person feels), but especially other-orientated feelings (sympathy, compassion, tenderness, softheartness) play a decisive role in judging harmful transgressions as wrong (Nichols, 2002).

However, according to some scholars, other negative emotions such as disgust or personal distress are also associated with moral norm violations (Batson & Shaw, 1991, 1991; Graham et al., 2009; Haidt et al., 1993; Nichols, 2002, 2004). For example, Joshua Greene (e.g., Greene, 2001) describes emotional and cognitive dispositions as indicators of moral obligations. Greene's model is therefore based on the dual-process theory, in which automatic emotional processes (System 1) interact with conscious and slow reflective processes (System 2). As a consequence of early human cooperation in small-scale interactions, specific emotional dispositions initially developed that lead one to react strongly to violent or uncooperative behavior (Roughley, 2018). Moreover, cognitive reflective abilities developed later in the evolution of human morality. These innate human emotional dispositions belong to System 1, which are initially activated automatically in moral situations. Somewhat later, in the reflective examination of moral situations, System 1 is replaced by System 2. Similar to Joshua Greene's theory is Jonathan Haidt's theory of moral foundations (Haidt & Joseph, 2004). As a social intuitionist approach to morality, he describes that moral judgments are formed by fast emotional moral intuitions (such as disgust), whereas moral thinking is formed as rationalization according to previously formed judgments (Haidt, 2001; Haidt et al., 1993).

Taking all findings together, the above approaches contrast with the rationalist work of Piaget and Kohlberg on moral development, which they regard as a primarily cognitive activity (Killen & Smetana, 2015; Roughley & Bayertz, 2019; Waldmann et al., 2012). Most importantly, however, prosociality and morality are not only two separate ontogenetic steps in the evolution of human morality, but rather they are also heavily interrelated, with prosocial motivation and, in particular, sympathy playing a crucial role in the development of a fully fledged understanding of morality.



## **1.4. The Early Development of Human Morality: An Empirical**

### **Overview**

There is much evidence that babies are already attuned to the social world around them and attend to the needs, emotions and mental states of others (Grossmann & Johnson, 2007). For example, even right after birth, newborns prefer real faces to other facial objects (Leo et al., 2018; Morton & Johnson, 1991). In the first year of life, infants acquire more social skills through the exchange with their environment. However, this exchange is mainly limited to simple dyadic interactions, in which the infant either interacts with an object, such as a ball, or refers primarily to its caregiver, for example when the mother smiles at the child and the child smiles back (Bakeman & Adamson, 1984).

A milestone in social-cognitive development is reached between the ninth and twelfth months of life. Infants begin to understand more and more that other people act based on goals, intentions and plans (Tomasello et al., 2005). As soon as the child realizes that other people are actors with mental states (such as intentions), the first triadic interactions begin. Triadic interactions take place between the child and an adult who jointly focuses his or her attention on an object, person or event. This moment is seen as an entry into social community life and is the basis for cultural learning – now things in the world can be experienced together. For the first time, children have a “we-intentionality” and understand that other people can be used as interaction partners with whom mental states can be shared. Thus, children increasingly use pointing gestures, such as declarative pointing, as a pre-linguistic means of communication to point out events to other people and to initiate joint attention. In one study, Liszkowski and colleagues (Liszkowski et al., 2004) observed that 12-month-old babies repeatedly drew

the experimenter's attention to a certain object via pointing gestures (e.g. a doll suddenly appearing from behind a curtain) in order to experience an event together.

During the second year of life, children begin to cooperate with other people through the acquired ability of joint attention, the setting of joint goals and the ability to share intentions. Understanding and sharing intentions is essential not only in joint actions, but also in the development of prosocial behavior (Tomasello, 2016). Parents and educators increasingly observe how their children no longer only take care of their own interests, but increasingly of the well-being of others as well. This special ability brings enormous advantages, not only for the community life, but also for the individual development. Prosocial behavior is an important component for acceptance among peers, self-confidence and emotion regulation (Paulus & Moore, 2012). The study of children's developing prosocial behavior starts long before children actually begin to act prosocially themselves. A plethora of studies show that 6- and 9-month-old infants demonstrate that they also prefer those who have helped others and selectively avoid agents with harmful intentions (Hamlin et al., 2011). The first signs of children's own helping behavior is evident from about 14 to 18 months of age, as numerous studies on instrumental help have shown (e.g., Warneken & Tomasello, 2006, 2007). The researchers designed various out of reach situations in which an investigator needed help to reach a specific goal. For example, the children watched an experimenter drop a clothespin while hanging up laundry, or he stood in front of a closed cabinet door and could not open it to put a pile of books in. This sensitivity towards others' needs extends into the second year of life, in which children's prosocial attention has transformed into an intrinsic motivation to see others helped. Findings on the overjustification effect give further evidence of the prosocial motivation theory, whereby initially intrinsically motivated children were less willing to help in the future with the prospect of material reward (Warneken & Tomasello, 2008). Moreover, further findings revealed that

children are happier and smile more when others' needs are fulfilled, even when this involves a material cost to themselves (Aknin et al., 2015). Not only is the child's own behavior decisive, but parenting style also plays an important role in the development of children's prosociality. Dahl (2015) showed that parents increasingly responded with praise to their children's helping behavior in order to encourage the children to help. In particular, those children who were encouraged more at the beginning of their second year of life helped more often at the end of their second year. In addition, Brownell, Svetlova, Anderson, Nichols, and Drummond (2013) investigated parenting behavior in parents of 18- to 30-month-old children and found that parents of more helpful children also tended to encourage them to name and explain people's emotions in picture books. Interestingly, it was not the explanations of the parents that correlated with the children's helping behavior, but the questions that encouraged them to think and empathize.

From early on, humans are at least rudimentarily able to put themselves in other people's shoes, which is also evident on an affective level, such as in emotional contagion. For example, babies cry when other babies cry – but not when they hear their own crying or other loud sounds (Dondi et al., 1999; Simner, 1971). This automatic and involuntary mood transmission has evolutionary significance; for example, to synchronize emotions and behavior within a group and to react quickly to danger. However, emotional contagion does not necessarily motivate prosocial behavior, rather empathic concern seems to be decisive (Batson & Shaw, 1991). In line with the empathy-altruism hypothesis, various studies give evidence that altruistic motivations (empathic concern) rather than egotistic motives are associated with moral issues, using different experimental methods like heart rate, facial expressions and self-report (Eisenberg et al., 1990; Eisenberg & Fabes, 1990), neurobiological studies (Hastings, Miller, Kahle, & Zahn-Waxler, 2014; Knafo, Zahn-Waxler, Van Hulle, Robinson, &

Rhee, 2008) or behavioral measurements (Hepach et al., 2012; Vaish et al., 2009).

Recently, pupillometry has proven to be particularly suitable for measuring emotional arousal in young children, by showing that the pupil dilate more in prosocial situations (Hepach, 2017; Hepach et al., 2017). Finally, individual differences in dispositional trait sympathy but not in personal distress also tend to positively correlate with prosocial behavior (Eisenberg et al., 2014, 2015; Malti et al., 2012).

According to many scholars, a necessary prerequisite for empathic concern is children's own self-awareness, measured by the rouge test (Amsterdam, 1972). From about 18 months, the children are able to recognize themselves in the mirror by touching their face – not the mirror – to remove a dot on their nose. Self-awareness leads to the knowledge of being able to separate oneself from one's environment. This self-other distinction allows us to empathize with others while at the same time differentiating our own feelings from those of others. Findings show that there is a connection between empathy and self-recognition in 15- to 24-month-olds, measured by the so-called Teddy Test (Bischof-Köhler, 1993). Results showed that the children react very differently to the experimenter because his teddy was damaged: from empathic helpers, who even comforted him, to egocentrically emotionally infected children, who cried out of personal distress, to children who were confused or uninvolved, apathetic to the events and who did not react at all. Self-awareness thus seems to have an important influence on the development of empathic concern and compassion, as well as on helpfulness and consolation. On the other hand, recent evidence suggests an early capacity for empathic concern in the first year of life, which is more associated with implicit forms of self-awareness (Davidov et al., 2013; Roth-Hanania et al., 2011). In particular, the researchers observed that 8- to 10-month-old children were concerned about the discomfort of their mother, who had suffered pain due to an unintentional hammer blow to her fingers. Results revealed individual differences in the ability to

empathize among the 10-month-old babies that was associated with later prosocial behavior at the ages of 12 and 14 months.

There is much evidence for an early rudimentary form of general norm understanding in infants during their second year of life. A recent study (Schmidt et al., 2019) provides evidence not only for descriptive expectations of appropriate social behavior but also for a more concrete but simpler form of norm understanding, at least in dyadic interactions. Results revealed that even 18-month-old infants intervened, corrected, and directed a puppet more in the normative than in the non-normative conditions. Thus, infants showed so-called second-personal normative expectations (an interpersonal “we” that regulates both “you” and “me”) about their partner’s behavior in a triadic interaction (“You should do X!”). Importantly, these simple normative expectations will later become group-minded impersonal and abstract forms of a social norm understanding (Nagel, 1986; Tomasello, 2016; Turiel, 1983). At the age of 3, children then take a normative turn in which social second-personal interactions are transformed by a developing understanding of normative standards of particular groups (Tomasello, 2018). Social (moral and conventional) norms are then not only adequately followed but also defended by children reacting to norm violations of others. For example, such studies show that 3-year olds intervene and protest against moral norm violations in a variety of contexts such as harm, fairness or property rights (Rakoczy et al., 2016; Rossano et al., 2011; Schmidt et al., 2013; Vaish et al., 2011). In addition to behavioral studies, a number of interview studies support the results in which children of different age groups were asked about their judgments of moral and conventional violations (Smetana, 1981, 1984; Smetana et al., 1993; Smetana & Braeges, 1990). The results revealed that children under the age of 3 could not yet distinguish between morality and convention. Children who were more advanced in their language skills, however, were more able to do so. Against this background, a recent study investigated

the evaluations of moral and conventional violations of 2- to 5-year-old children by adding a behavioral puppet task to the language-driven explicit judgment tasks (Smetana et al., 2018). The results show that even the two-year-olds differentiated correctly between the moral and conventional situations and that they acted less prosocial to moral norm violators than to conventional norm violators. The younger children could not pass all tasks correctly.

The development of fairness norms has long been a topic of interest. Many empirical findings have shown that even infants have at minimum descriptive expectations about, and preferences for equality in resource allocation (Geraci & Surian, 2011; Meristo et al., 2016), and that third-party expectations of fairness are closely interrelated with infants' own (first-party) morally relevant sharing behavior (Schmidt & Sommerville, 2011). In a recent study, Rakoczy and colleagues (Rakoczy et al., 2016) found that 3- and 5-year-olds protested against unfair, merit-based resource allocation decisions in a paradigm involving collaboration among the children, both when children are affected and when they are disinterested observers. Furthermore, there are a great number of developmental studies that choose costly punishment to investigate children's reactions to unfair behavior (Kenward & Öst, 2012; Krasnow et al., 2016; McAuliffe et al., 2015; Robbins & Rochat, 2011; Salali et al., 2015). Such studies show that even preschool children use costly punishment and sacrifice their own resources to punish unfair behavior. This suggests that by 3 years of age, children have some normative expectations, in particular about equality in resource allocation. Although 3-year-old children have a certain sense of justice, only older children between the ages of 5 to 8 understand more complex relationships, such as legitimate reasons for an unequal distribution, such as merit, need or agreements (Schmidt, Svetlova, et al., 2016).

Finally, the development of morality is not limited to the particular in-group, but is also an important factor in intergroup contexts. Children show explicit and implicit preferences for members of their own language group, gender (Shutts, 2015) and race (Dunham et al., 2008). A plethora of developmental investigations focuses on social categorization processes. Such studies show that intergroup-bias already occurs at the age of 3 and more robustly at the age of 5 in minimal group settings (Dunham et al., 2011), manifested in various prosocial forms like sharing or helping behavior (Benozio & Diesendruck, 2015; Fehr et al., 2008; Over, 2018; Sierksma et al., 2018) or in trust and loyalty (Misch et al., 2016; Rutland et al., 2015). Recent studies even pointed to an interplay between morality and group identity (Chalik & Dunham, 2018; Rutland et al., 2010).

## **1.5. Focus of the Dissertation and Methodological Approach**

The general aim of this thesis was to examine young children's developmental origins of morality. Thus, two aspects – prosociality and morality – are of particular importance, not only because they are thought to play a key role in maintaining the unique human capacity for large-scale cooperation (even with nonrelatives) but also because of their close relation to each other. A large number of studies have addressed the development of children's first-party prosociality and their third-party understanding of moral norms separately, but little is known about their interrelation to each other. Therefore, I examined three research questions that are essential for understanding the ontogeny of morality and its interrelation with young children's prosocial (altruistic) motivation to understand, adhere to, and enforce moral norms within their own and between different groups.

(1) What are the developmental origins of morality? There is much evidence that children in their second year of life develop empathic concern and sympathy for others in need in both prosocial and morally relevant situations. Moreover, recent findings suggest that 18-month-old children already show some rudimentary forms of norm understanding in at least dyadic conventional situations, which are interpreted as second-personal normative expectations about how a partner in a dyad should act rather than an abstract and group-wide impersonal understanding of norms that develops around the third year of life. Thus, I assume that both the capacity to feel sympathy and the ability for third-party descriptive expectations enable infants to differentiate between simple conventional and moral violations without having a full-blown normative and agent-neutral understanding of norms yet.

(2) What is the underlying motivation for children's normative appreciation of morality? In particular, I am interested in the role of altruism and sympathy as important



motivational factors in the development of fairness norms by investigating the relation between young children's first-party prosociality and their third-party normative expectations about fairness.

(3) What is the scope of morality (i.e., children's sharing behavior in an intergroup context)? There is much evidence that preschool children show in-group favoritism. That is, preferring members of their own group over members of other groups which is particularly evident in their prosocial (sharing) behavior. However, the extent to which specific mechanisms leads to reversing the effect and ultimately to equal treatment of both groups (equality as a major moral category) has not yet been clarified.

All three studies were conducted with different methods and measures to allow for a broader picture of the ontogeny of morality and the role of prosociality in this process. In Study 1, a novel eye-tracking paradigm was designed using anticipatory looking as a non-verbal measure to examine infants' third-party descriptive expectations and pupil dilation as a measure of their physiological arousal in response to moral and conventional violation situations. In Study Set 2, in Experiments 1 and 2, fairness violations were committed in a third-party protest paradigm with hand-puppets to allow for the assessment of young children's spontaneous verbal utterances and behavioral interventions as a standardized measurement for children's (agent-neutral, moral) norm understanding. In Experiment 3, I used a costly punishment paradigm with stickers. Moreover, in all three experiments, children's first-party prosocial behavior was measured with a simplified costly sharing behavior (altruism) paradigm and an emotion sharing (empathic concern) paradigm. In Study 3, I used a classical dictator game with stickers to measure children's costly sharing behavior in an intergroup context.

The first two studies addressed the developmental origins of morality within the children's own cultural group (i.e., there were no particular cues of group membership),

whereas the third study dealt with the scope of morality in an intergroup context (in-group/ out-group comparison).

In Study 1, I investigated the developmental origins of morality in 18-month-old infants. The question was whether 18-month-old infants differentiate between prototypical moral (harm) and conventional (harmless) violations using Turiel's (1983, 2006) social domain theory approach. In a between-subjects-design, children watched the very same video clip in three different conditions. In the first two conditions, an instructor told an actor to destroy a picture with an upper tool A (conventional violation condition) or with a lower tool B (no violation condition), whereas in the third (moral violation) condition, the instructor forbade the actor to destroy the picture at all. The actor then said he would destroy the picture, grasped the lower tool B, and destroyed the picture, resulting either in no norm violation, a conventional violation, or a moral violation. I hypothesized that infants would show a larger relative increase in pupil dilation in response to a moral violation than to a conventional violation. Moreover, I expected infants to differentiate between two types of conventional norm situations in their anticipatory looking based on prescribed actions.

Study set 2 looked at the underlying prosocial motivation for the normative appreciation of morality in 3-year-old children. In three experiments, I investigated whether the origin of fairness (understood as a normative notion) is mainly seen as a convention (agreements, regularities, habits) or whether it is related to moral issues (based on concern for the welfare of others), in particular with prosocial tendencies and other-regard. Three-year-old children were given a third-party fairness task that varied across the experiments and different prosocial tasks. In Experiment 1, I examined whether children protested and corrected unequal but not equal allocations without given them any cues to act fairly. Experiment 2 assessed whether children's normative expectations about fairness have a moral (authority-independent) dimension. Here,

children observed a distributor who followed (unequal condition) or violated (equal condition) an authority's command to allocate resources unequally. In Experiment 3, I was interested in whether children enforced fairness norms using altruistic punishment in the sense of restorative justice. Moreover, in all three experiments I examined the extent to which children's norm understanding of fairness was associated with their own prosocial motivation, measuring children's costly sharing behavior and emotional sharing behavior as a function of their intrinsic motivation to act prosocially.

Finally, in Study 3, I explored the scope of morality in 5-to 6-year-old children in a typical intergroup context. In today's times of heterogeneous societies and attempts to overcome prejudice and xenophobia, it is important to examine not only intergroup-bias, but also ways to overcome group thinking and foster equal treatment of individuals regardless of group membership. Thus, I investigated whether decategorization – a candidate mechanism to overcome intergroup bias by emphasizing the individual person – would lead preschoolers to treat in-group and out-group members equally when sharing resources.

## **2. Study 1: 18-month-old Infants Differentiate between Moral and Conventional Violations**

### **2.1. Introduction**

In all human societies, people follow, create and enforce norms. But not only that, even children adhere norms and punish other community members who do not abide by the agreed upon rules (Göckeritz et al., 2014; Rakoczy & Schmidt, 2013; Schmidt, Butler, et al., 2016; Schmidt & Tomasello, 2012). Not all norm violations are treated alike. According to social domain theory (Turiel, 1983, 2006), people differentiate between two types of norms. Moral norms are defined as the understanding of others' welfare, fairness and rights (e.g., against harming others) which are differentiated from arbitrary conventional norms that provide consensually determined expectations for appropriate behavior, such as agreements, customs or rules (e.g. dress codes, classroom regularities). A wealth of studies has pointed to an early ability to distinguish between moral and conventional norm violations (e.g., Killen & Smetana, 2015; Smetana, 1984; Smetana & Braeges, 1990), but it is not yet clear exactly when children will be able to do so.

There is much evidence for an early rudimentary form of norm understanding in infancy. In the first two age years, children are capable to use normative evaluations (e.g., descriptive third-party fairness expectations; Geraci & Surian, 2011; Meristo, Strid, & Surian, 2016) and that these normative expectations were interrelated with children's own altruistic behavior (Schmidt & Sommerville, 2011). Moreover, a recent study (Schmidt et al., 2019) provide evidence not only for descriptive expectations of appropriate social behavior but also for a more concrete but simpler form of norm understanding in at least dyadic interactions. Results revealed that even 18-month-old

infants intervened, corrected, and directed a puppet more in the normative than in the non-normative conditions. Thus, infants showed so-called second-personal normative expectations (an interpersonal “we” that regulates both “you” and “me”) about their partner’s behavior in a triadic interaction (“You should do X!”). Importantly, these simple normative expectations will later be scaled up to group-minded impersonal and abstract forms of a social norm understanding (Nagel, 1986; Tomasello, 2016; Turiel, 1983).

In line with social domain theory (Turiel, 1983, 2006), the ability to distinguish between moral and conventional norms depends on a normative agreement that determines which actions are prohibited and on an affective mechanism (Nichols, 2002, 2004). Moral norm violations are seen as more severe and more deserving of punishment than disregarding conventions. Even at preschool age moral and conventional norms are not only adequately followed but also defended by children reacting to norm violations of others. For example, such studies show that around three years of age, young children intervene and protest against moral norm violations in a variety of contexts such as harm, fairness or property rights (Rakoczy et al., 2016; Rossano et al., 2011; Schmidt et al., 2013; Vaish et al., 2011). In addition to the behavioral studies, a number of interview studies supported the results in which children of different age groups were asked about their judgments of moral and conventional violations (Smetana, 1981, 1984; Smetana et al., 1993; Smetana & Braeges, 1990). The results revealed that children under the age of three could not yet distinguish between morality and convention. Although children who were more advanced in their language skills, were more able to do so. Thus, a recent study investigated the evaluations of moral and conventional violations of 2- to 5-year-old children by adding a behavioral puppet task to the language-driven explicit judgment tasks (Smetana et al., 2018). The results show that even the 2-year-olds differentiated correctly between moral and

conventional situations and that they act less prosocial to moral than to conventional norm violators. But still the younger children could not pass all tasks correctly. Hence, there is much evidence that younger children may already be able to distinguish between morality and convention, for example when using non-linguistic tasks.

Studies with children and adults show that affects are deeply intertwined with moral reasoning and lead to strong emotional reactions as part of moral judgments (Decety & Howard, 2013; Eisenberg et al., 2014; Jensen et al., 2014; Killen & Smetana, 2015; Nichols, 2002; Smetana et al., 2014; Turiel, 1983). Moreover, there is an ongoing debate about whether empathy, sympathy or personal distress lead to these strong emotional reactions in moral situations (Batson & Shaw, 1991; de Waal, 2008; Eisenberg & Fabes, 1990). While empathy is defined as the mere ability to cognitively and emotionally representing other thoughts and feelings (e.g., feeling what the other person feels), sympathy and personal distress often results from empathy (Eisenberg & Fabes, 1990). According to Batson and Shaw (1991) both personal distress and sympathy end in prosocial behavior, but the motivations behind the two constructs are different. Sympathy is associated with altruistic, other-oriented motives and involves emotions such as concern. Personal distress is defined as a self-focused, aversive emotional response to another's distress and associated with egoistic motivation. Therefore, personal distress is positively related to prosocial behavior primarily when there is no easier way than helping to reduce one's own distress. Various studies give evidence that altruistic motivations (empathic concern, sympathy) are associated with moral issues rather than egoistic motives, using different experimental methods like heart rate, facial expressions and self-report (Eisenberg et al., 1990; Eisenberg & Fabes, 1990), neurobiological studies (Hastings, Miller, Kahle, & Zahn-Waxler, 2014; Knafo, Zahn-Waxler, Van Hulle, Robinson, & Rhee, 2008) or behavioral measurements (Hepach et al., 2012; Vaish et al., 2009).

Recently, pupillometry has proven to be particularly suitable for measuring emotional arousal in young children, by showing that pupil dilated more in morally relevant situations (Hepach, 2017; Hepach et al., 2017). Importantly, physiological changes may reflect discrete emotional responses, but on the other hand they also are influenced by a wide range of non-emotional factors such as attention, cognition, physical activity, or extraneous stimuli (Eisenberg & Fabes, 1990). For instance, pupil size dilated more as sensory response to dark than to bright light, or in cognitive tasks, such as increased mental effort (Kahnemann & Beatty, 1967; Sirois & Brisson, 2014; Verschoor et al., 2015). Therefore, it is important to exclude and control alternative results in pupillometry. Taken together, (i) infants have an early rudimentary form of norm understanding at least in dyadic interactions, (ii) methods that rely on linguistic abilities may lead to a systematic underestimation of children's understanding of the distinction between morality and convention below three years of age, (iii) and previous findings suggest that early third-party descriptive expectations might be the basis for the development of understanding conventional norms. Nevertheless, these studies leave open, whether infants distinct between conventional and moral violation situations at least as descriptive expectations, and moreover whether infants also understand the normative force of moral norms on an affective level (empathic concern, sympathy).

For the present study, we therefore developed a novel eye-tracking paradigm, implementing both anticipatory gaze and pupil dilation as measures of affective arousal (Hepach et al., 2012, 2017). By using these non-verbal measures, we investigate whether 18-month-old infants differentiate between moral (harm) and conventional (harmless) violations.

Since we concurrently measure pupillometry and looking behavior in response to movies, we used the exact same visual presentation across three conditions. The type of violation varies across conditions by alterations in the speech content of the actor who

explained a game. To be sure those infants understand the normative moral and conventional directives we decided to use imperatives. In line with our considerations, a recent study show, that toddlers are already able to appropriately understand the normative structure of imperative speech acts and criticized actors, who disregard the instruction of a speaker who explained a directive (Rakoczy & Tomasello, 2009). Depending on the speech content, the exact same action performed thereafter by the observer constituted a conventional norm violation, a moral norm violation or a no violation. In the no violation situation, the observer follows the instruction to destroy a drawing. In the conventional norm violation situation, the observer reaches the goal (the destruction of a drawing) but violates the game instructions by using the wrong tool to accomplish it. The moral norm violation situation constitutes a moral norm violation in that the observer destroys the instructor's property while explicitly instructed not to. We expected infants to differentiate between the two types of violations in their anticipatory looking. Moreover, we hypothesized that infants in the no violation situation would predict that the actor grasps the tool she actually grasps thereafter, while infants in the conventional violation situation should predict the other tool to be used. In the moral violation situation the infants should not have a specific expectation. Second, we predicted that infants would show a larger relative increase in pupil dilation in response to a moral violation than to the conventional violation situation.

## **2.2. Method**

### **2.2.1. Participants**

Seventy-two 18-month-old children ( $M = 18$  months 6 days,  $Min = 17$  month 15 days,  $Max = 19$  month 12 days,  $SD = 0.42$ , 34 girls) participated in the study. Additionally, 2 children were excluded due to fussiness and another 9 due to experimenter error.

Children were recruited through the municipality and received small gifts as



compensation. An informed consent and a questionnaire regarding general health and development were obtained. The children were all healthy full-term and without pre- or perinatal complications.

### **2.2.2. Design**

We designed a between-subject eye-tracking experiment in which one of three game conditions is presented: a no violation (control-) condition (NVC), a conventional violation condition (CVC) and a moral violation condition (MVC).

### **2.2.3. Materials**

#### **2.2.3.1. Test Environment and Apparatus**

During the experiment the children sat in a stimulus-poor booth on the lap of their caretaker, who was seated in front of the eye tracker apparatus. The distance between eyes and apparatus was approximately 65 centimeters (the screen's viewing angle was  $43.5^\circ$  by  $28.0^\circ$ ). The behavior of the children was monitored online by the experimenter from a separate control room by means of a camera located above the eye tracker. A 24 inch TFT-screen (HP Elite Display E242, 1920 x 1200 pixels, 16:10), equipped with a Tobii Pro X3-120 eye tracker (Tobii Technology, Stockholm, Schweden) was used for visual and auditory data presentation. The Tobii external processing unit recorded gaze data at 120 Hz and pupillary data at 40 Hz. The Tobii X3 has an average accuracy of  $.4^\circ$  and allows for head movements by the subjects (50x40x40cm). Stimulus presentation was controlled by a PC running Tobii Studio® software (version 3.4.8). Prior to the eye-tracking experiment the caretakers were instructed not to move after calibration and gently hold the child in order to maintain eye tracker alignment, and to entertain the infant during the (approximately 1-min) break between calibration and the experiment. The caretakers were instructed to look at the children's head (not at the screen) during

calibration and testing. The eye tracker was calibrated with the standard 9-point calibration, when this failed to elicit enough attention we used one with a small Tobii animation instead. The calibration was accepted with a minimum of seven points acquired. Lighting conditions were kept constant across participants.

#### **2.2.3.2. Stimuli**

The video material consisted of 4 consecutive movies (43.5° x 25.2°, 30 frames/S). To standardize luminance across conditions, all three conditions (NVC, CVC and MVC) contained exactly the same visual material. Only the voiceover of movie 1 was changed across conditions. The number of syllables was kept constant across conditions. The volume of the soundtracks was normalized and mastered for equal volume across conditions. The Timing of speech was synchronized across conditions.

#### **2.2.4. Procedure**

Children were tested at a time when they were likely to be alert. After the instruction of the caretaker and the calibration, the test phase started with the visual presentation of one of the three conditions (Movie 1 - 4). The introduction phase (Movie 1) functioned to explain the rules of the game. First, two actors (instructor, observer) introduced themselves by name and looked directly in front of the camera to familiarize the participants with the actors faces and voices. In the first two conditions, the instructor explained to the observer to destroy a picture with an upper tool A (CVC) or with a lower tool B (NVC), whereas in the third (MVC) condition, the instructor asked the observer not to destroy the picture at all. Pupil dilation was measured during specially designed movies. During this phase (Movie 2) the participants saw a red dot that was placed in the center of the screen (5 sec animated, 5 sec stationary). Movie 2 was used as baseline measurement. In the anticipatory gaze phase (Movie 3) the actor first

announced to destroy the picture, followed by a still frame (length 6 sec). The still frame was used to test for anticipatory gaze. The test phase ended with the observer grasping the lower tool B, and destroyed the picture, resulting either in NVC, CVC, or MVC. Finally, a second pupil animation phase (Movie 4) was used as post measurement for the pupil diameter. Importantly, Movie 2 und 4 were exactly the same.

### **2.2.5. Eye-Tracking Data Handling**

For the predictive looking analysis, we used Tobii Studio ® software (version 3.4.8) to aggregate looking times per subject per Areas Of Interest (AOIs) of the first three seconds of the still frame. The standard Tobii I-VT fixation algorithm was used to define fixations. The AOIs were the same for the time windows tested across all three conditions (see supplemental material). Both pincer AOIs were  $10.7^\circ$  by  $9.3^\circ$ ,  $4.1^\circ$  above and  $5.5^\circ$  below the vertical midline.

For the pupillary analysis we exported the 40 Hz raw pupillary data per eye from Tobii Studio ®. Pupillary data were recorded during the stationary part of movie 2 and 4. The red fixation dot's size was  $1.4^\circ$ . We then applied a number of processing steps using R. Firstly, we applied an outlier rejection based on minimum and maximum pupil size: min = 1mm, max = 6mm (Verschoor et al., 2013). Then we applied an outlier rejection based on the maximum allowed change in pupil size in 25ms, defined as 0.5 mm in 25 ms (Verschoor et al., 2013). Next we interpolated both the left and the right eye according to Hepach et al. (2012) (maximum gap: 3 missing data points). Then data from left and right eye were combined by averaging them into one value. If only one of those was present, then the present data point was used. Thereafter the combined data were interpolated once more using the Hepach procedure (Hepach et al., 2012) (maximum gap 3 missing data points), any missing values were removed and a last outlier rejection was performed based on the Standard Deviation (*SD*) of the segment

(any values that deviated more than 3 *SDs* were removed). Lastly, we calculated the change in pupil size from movie 2 to 4 by subtracting the mean pupil size of movie 2 from movie 4.

### **2.2.6. Statistical Analysis**

Statistical Analysis were run in R, version 3.6.1. (R Core Team, 2019). Since the two dependent measures (anticipatory looking and pupil diameter) violated assumptions of standard linear models (i.e., normally distributed errors), we each calculated a Generalized Linear Model (GLM) with gaussian error structure. For all GLMs likelihood ratio tests (LRT; Dobson, 2002) were used to test for the main effect of condition by each comparing the full model with a null model without the predictor variable condition. Unstandardized parameter estimates (*b*), 95% confidence intervals of parameter estimates (CIs) and standard errors (*SE*) were obtained from the full model. Eta Squared ( $\eta^2$ ) was computed to estimate effect sizes for main effects of condition.

## **2.3. Results**

For anticipatory looking, we found a significant main effect of condition,  $F(2) = 3.42$ ,  $p = .038$ ,  $\eta^2 = 0.09$ , which is a medium effect size according to Cohen's convention (1992). As predicted, planned comparisons revealed that children showed anticipatory looks significantly more towards tool A versus tool B in the CVC ( $M_{CVC} = 0.2$ ,  $SD_{CVC} = 0.57$ ) than in the NVC ( $M_{NVC} = -0.06$ ,  $SD_{NVC} = 0.21$ ),  $b = 0.26$ ,  $SE = 0.11$ ,  $t = 2.44$ ,  $p = .018$ , CI [0.05, 0.47] (see Figure 1).

For pupil diameter, we first examined whether the participants differed in pupil size across conditions during baseline measurement (Movie 2), before the different conditions were presented. We did not find any differences across conditions,  $F(2) =$

1.19,  $p = .309$  (see Figure 2). Thus, in all conditions the same prerequisites for the relative change of the pupil sizes were met ( $M_{NVC} = 3.38$ ,  $SD_{NVC} = 0.43$ ;  $M_{CVC} = 3.23$ ,  $SD_{CVC} = 0.33$ ;  $M_{MVC} = 3.41$ ,  $SD_{MVC} = 0.54$ ). Second, we found a significant main effect (with medium effect size) on relative change in pupil diameter (Movie 4 minus Movie 2) depending on condition,  $F(2) = 3.32$ ,  $p = .042$ ,  $\eta^2 = 0.09$ . As predicted, planned comparisons revealed, that children in the MVC ( $M_{MVC} = 0.16$ ,  $SD_{MVC} = 0.19$ ) reacted with a significantly higher change in pupil diameter as compared to the CVC ( $M_{CVC} = 0.04$ ,  $SD_{CVC} = 0.14$ ),  $b = 0.12$ ,  $SE = 0.05$ ,  $t = 2.49$ ,  $p = .015$ , CI [0.02, 0.21]. Furthermore, we did not find any difference between CVC and NVC ( $M_{NVC} = 0.07$ ,  $SD_{NVC} = 0.16$ ),  $b = -0.03$ ,  $SE = 0.05$ ,  $t = -0.68$ ,  $p = .500$ , CI [-0.01, 0.18.] (see Figure 3).

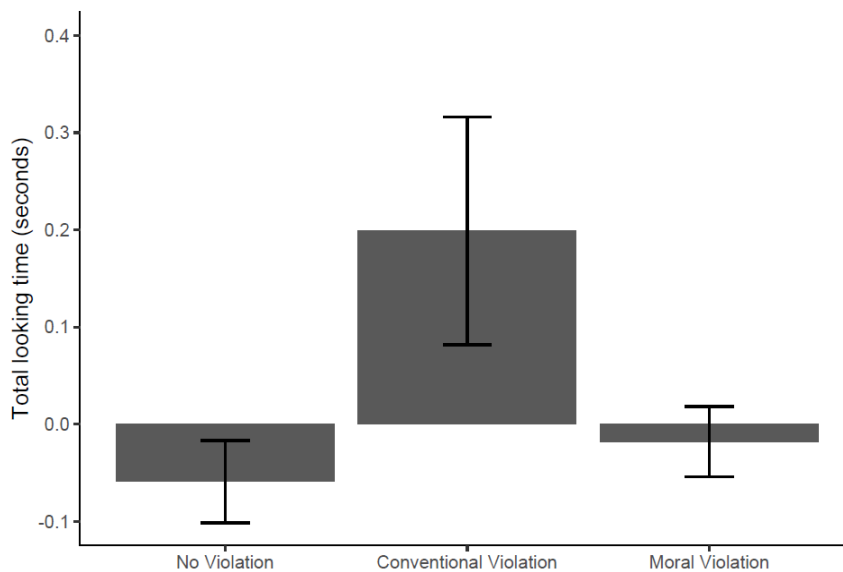


Figure 1. Differences in anticipatory looking between tools A and B as a function of condition. Error bars indicate standard errors of the mean.

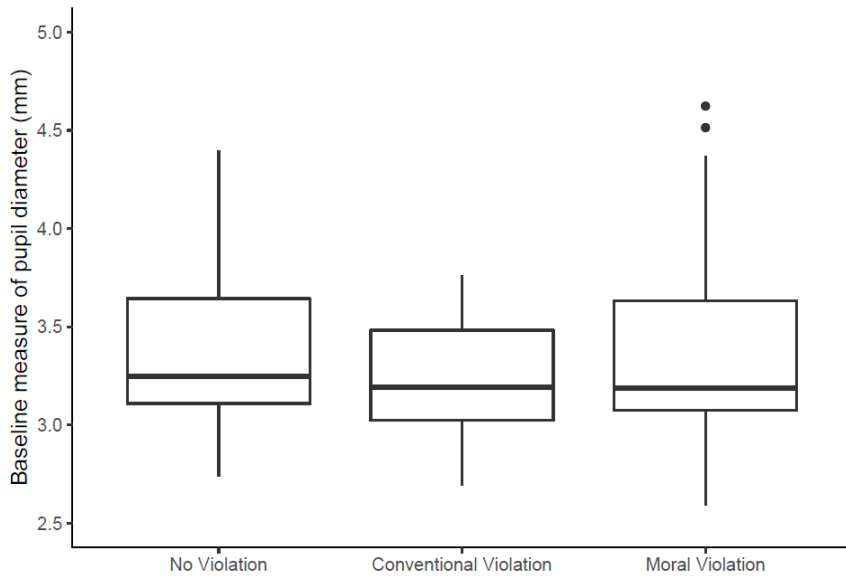


Figure 2. Box-and-whisker plots of pupil diameter from the baseline measure (Movie 2) as a function of condition.

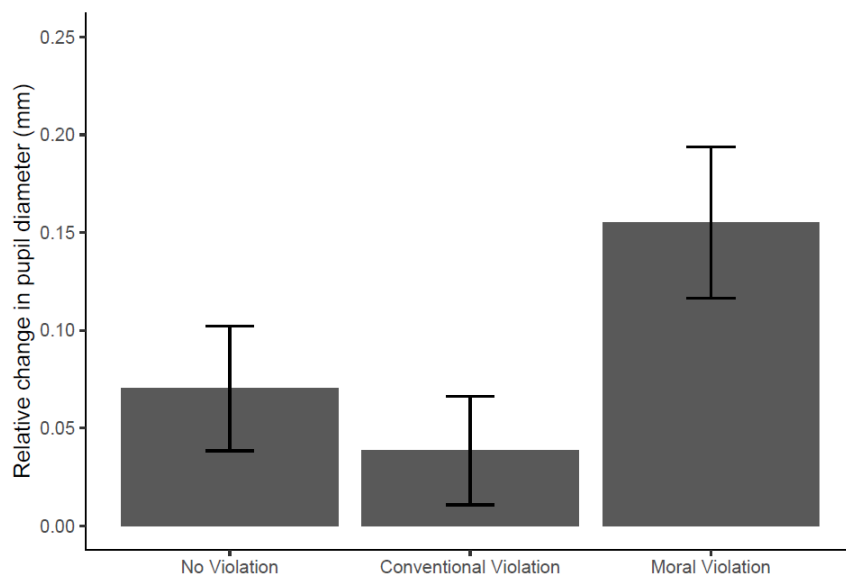


Figure 3. Relative change in pupil diameter (Movie 4 minus Movie 2) as a function of condition. Error bars indicate standard errors of the mean.

## 2.4. Discussion

Our findings suggest that the distinction between situations that are considered prototypical moral and conventional violations begins much earlier in development than previously thought. Using anticipatory looking, we first found that 18-month-old children distinguished between two types of conventional situations, having descriptive expectations about an agent's actions depending on prior information about the violation situations. More precisely, these situations consisted of two mutually exclusive game rules which were explained by the instructor, namely, that the observer should tear apart a picture with a certain tool, while not using the other tool (A or B depending on condition), which – later in context – leads to a no violation or a violation situation. As we predicted, in each conventional situation, the children exactly expected the observer to use the prescribed tool. This finding suggests that infants possess a third-party descriptive understanding of conventional rule agreements. Second, in line with social domain theory (Smetana & Braeges, 1990; Turiel, 1983, 2006), we were interested in whether infants were capable to differentiate between moral and conventional violation situations. Thus, we predicted a larger physiological arousal to the violation of a moral situation (destroying the property of the owner) than to the violation of a conventional situations (using the wrong tool), since moral violations are associated with emotional reactions that lead to an enhanced physiological arousal. As predicted, we found a larger relative increase in pupil dilation in response to a moral violation than to a conventional violation situation. Furthermore no difference was found between the other two conditions, in particular the conventional violation and no violation situation.

However, there are further causes why pupil dilation interacts with physiological arousal (Eisenberg et al., 2014; Pletti et al., 2017). We controlled for cognitive or

sensory effort to exclude such alternative interpretations. First, we used the same video material across all conditions controlling for luminance. Second, we controlled for cognitive effort caused by speech comprehension by using the same sentence structure with the same number of syllables, speech volume and timing across all three conditions. Moreover in our view, the finding that infants had increased physiological arousal only in the moral violation suggests that we measured increased affective arousal (potentially based on empathic concern), and not alternative candidates, such as increased cognitive effort or attention. In the latter case, we would have expected a relative increase in pupil size in both the moral and conventional violation situation, since a norm violation also occurred in the conventional situation, and there is no a priori reason to assume that the moral situation is cognitively more complex or harder to process than the conventional violation situation. In fact, one may even construe the moral situation as simpler because everything is forbidden and no action should be performed at all. Importantly, if relative pupil size differences were driven by cognitive effort or the like, we should have found a pupil dilation difference between conventional violation and no violation, because it is cognitively more demanding to process an expected vs. unexpected action. However, the question arises, how this affective arousal in the moral violation condition should be interpreted. A large number of experimental studies have shown, that such affective arousal is associated with an empathic concern for the victim and leads to an altruistic motivation to show prosocial behavior (Batson & Shaw, 1991; Eisenberg et al., 1990; Eisenberg & Fabes, 1990; Hepach, 2017; Hepach et al., 2017; Jensen et al., 2014; Nichols, 2002, 2004; Vaish et al., 2009). Moreover, our findings are consistent with psychological, physiological and philosophical explanations of morality (Killen & de Waal, 2000; Killen & Smetana, 2015; Nichols, 2002; Smetana, 1984; Smetana et al., 2014; Tomasello, 2016; Turiel, 1983, 2006). Violations of moral norms are considered more serious and punishable



than conventional violations, since they take the welfare of others into account. A plethora of studies have shown such interrelations between morality and other-regarding concerns (empathic concern, sympathy) in children and adults (Decety et al., 2012; Decety & Yoder, 2016; Eisenberg et al., 2014; Hepach, 2017; Hepach et al., 2012; Jensen et al., 2014; Lamm et al., 2019; Nichols, 2002; Roth-Hanania et al., 2011; Svetlova et al., 2010; Vaish et al., 2009).

Taken together, our findings go beyond previous language-driven work in interview studies as well as behavioral studies, which suggest that children at the age of three reliably distinguish between non-arbitrary moral norms (e.g., against harming others) and more arbitrary conventional norms (e.g., Schmidt, Rakoczy, & Tomasello, 2012; Smetana et al., 2018; Smetana & Braeges, 1990). Our results are in line with recent findings, who suggest an early rudimentary norm understanding in the second year of life (Schmidt et al., 2019). We found that already 18-month-old children not only have distinguished (in a descriptive sense) between different types of conventional situations, but they were also more physiological aroused when witnessing a moral violation which is interpreted as affective arousal due to the severity of the moral violation (presumably associated with empathic concern for the victim). Hence, both the capacity to feel sympathy and the ability for third-party descriptive expectations enable infants to differentiate between simple conventional and moral violations without having a full-blown normative and agent-neutral understanding of norms yet.

Further research is needed to disentangle the phenomenon of infants' capacity to distinct moral and conventional situations. A combination of eye-tracking and behavioral task may be a suitable paradigm for studying the interaction between the children's first-personal behavior and the evaluation of third-parties in morally relevant situations. Investigating across these lines will contribute to a richer account of when and why children are able to distinct between morally relevant situations from others.

### **3. Study Set 2: The Moral Roots of Fairness: Young Children's Enforcement of Fairness Norms Is Related to Their Prosocial Behavior**

#### **3.1. Introduction**

Third-party norm enforcement is considered a key mechanism in the evolution of human cooperation (Chudek & Henrich, 2011; Fehr & Fischbacher, 2004): we have normative expectations about how people “ought” to act in certain situations, for instance simple codes of conduct, such as dress codes or stronger prohibitions, such as not hitting, stealing or cheating each other. Thus, social norms – as learned behavioral standards – are shared and enforced among members of our cultural community, which ultimately leads to a peaceful group life and a large-scale cooperation. According to social domain theory (Turiel, 1983, 2008) social norms are subdivided into conventional and moral norms. While conventional norms are based on habits, agreements, traditions and customs (e.g., driving the car on the right side of the road), moral norms carry more normative weight and are based on a concern for the welfare of others (e.g., not pulling someone's hair). When people are asked to distinguish moral issues from conventions, they commonly use the following criteria: both norms are generally permissible, but only moral norms should be adhered to under all circumstances (generalizability), they are valid even if an authority prescribes to violate the norm (authority contingency) and violations are considered more severe (seriousness) than disregarding conventions. For example, stealing is generally forbidden, even if a teacher allows the children to do so. In contrast, raising your hand to talk to the teacher (as regularity in the classroom) should only be valid at school, but not when you meet him on the street.

Empirical findings suggest, that 26-month-olds don't, but 34-month-olds do, make the moral-conventional distinction, at least regarding to some of Turiel's criteria (Smetana et al., 2012; Smetana & Braeges, 1990). Moreover, 2- and 3-year-olds protest simple conventional game violations, albeit the younger age group uses more imperative, not normative language (Rakoczy et al., 2008), and both age groups also learn to enforce norms especially in dyadic contexts (Hardecker & Tomasello, 2017). Behavioral studies also show that 3-year-olds, but not 2-year-olds, protest against moral transgressions, such as violations of property rights, even if they are not affected (Rossano et al., 2011). The development of fairness as distributive justice (equality in resource allocation) has long been a topic of interest. Many empirical findings have shown that even infants have at minimum descriptive expectations about, and preferences for equality in resource allocation (Geraci & Surian, 2011; Meristo et al., 2016), and that third-party expectations of fairness are closely interrelated with infants' own (first-party) morally relevant sharing behavior (Schmidt & Sommerville, 2011). In a recent study, Rakoczy and colleagues (Rakoczy et al., 2016) have found that 3- and 5-year-olds protest against unfair resource allocation decisions in a paradigm closely linked to merit after having collaborated both when children are affected and when they are a disinterested observer. This suggests that by 3 years of age, children have some normative expectations and in particular about equality in resource allocation.

Some researchers have pointed to different notions of justice (retributive and restorative justice) as responses to rule breaking in resource allocation tasks (Heffner & FeldmanHall, 2019; Riedl et al., 2015; Wenzel et al., 2008). Retributive justice refers to the repair of justice through punishment of the perpetrator, whereas restorative justice focuses more on collective solutions between the perpetrator and the victim, and thus repairs justice under a normative aspect through reaffirming a shared value-consensus among all affected group members. Interestingly, most developmental studies choose

costly punishment (as retributive justice) to investigate children's reactions to unfair behavior (Kenward & Östh, 2012; Krasnow et al., 2016; McAuliffe et al., 2015; Robbins & Rochat, 2011; Salali et al., 2015). Such studies show that even preschool children use costly punishment and sacrifice their own resources to punish unfair behavior. Nevertheless, it is not yet clear whether children's response to unfair behavior can be separated into restorative and retributive justice reactions and especially whether 3-year-old children use restorative justice in the sense of a normative (shared value-consensus) notion.

Taken all findings together, it is not known so far whether children's normative expectations about fairness are more moral (based on concern for the welfare of others) or more conventional (based on habits, rules, agreement, etc.), and what mechanisms underlie young children's motivation to defend (third-party) fairness norms. A major question therefore pertains to the origins of our sense of fairness understood as a normative notion: is it mainly conventional, habitual (everyone typically gets his/her share) or is it interrelated with moral development, in particular with prosocial tendencies and other-regard, especially sympathy and altruism – that is, being interested in and concerned about others' well-being and motivated to act accordingly (Batson & Shaw, 1991; Hepach et al., 2017; Jensen et al., 2014; Nichols, 2002, 2004; Schmidt & Sommerville, 2011; Vaish et al., 2009). Furthermore, a second major question arises whether children are willing to sacrifice their own resources in order to repair justice in the sense of a restorative (normative) aspect (Riedl et al., 2015).

We therefore investigated to what extent children's sense of fairness is moral or conventional and with which prosocial motivations (empathic concern, altruism) it is interrelated. We hypothesize, that if an early concern for fairness as a normative concept is moral, children's tendency to react against fairness violations should be correlated with their own prosocial behavior. If not, children's behavior should be unrelated to

their first-party prosocial behavior. We predict that early normative expectations about third-party fairness are moral, because during development infants learn (1st-, 2nd-, and 3rd-party) about, and experience, fair sharing and allocation, and are interested in others' well-being. Thus, we also suggest – following social domain theory (Killen & Smetana, 2015; Smetana et al., 2012; Smetana & Braeges, 1990; Turiel, 1983, 2008) – that children's moral concern for fairness cannot be easily changed by authorities legitimizing unfairness (authority-independency). That is, children who have a moral concern for third-party fairness won't react against someone who does not follow an authority's order to be unfair (i.e., to allocate resources unequally). Children, however, who have not yet developed a moral concern for third-party fairness, will protest if a distributor does not follow the authority's dictate to be unfair. Furthermore, the relationship between one's own prosocial tendencies and third-party punishment for restoring justice should also be linked when children sacrifice their own resources to restore fairness as a norm.

Thus, we conducted three experiments to investigate whether children show distinct prosocial motivations (empathic concern, altruism) in different fairness contexts. In Experiment 1, we investigated whether three-year-old children protest against unequal distributions without giving them any indication to act fairly. In Experiment 2, we were interested in whether 3-year-old children have a moral understanding of fairness that is authority-independent. Finally, Experiment 3 examined whether children punish unfair behavior by contributing own resources in a restorative sense of justice.

## **3.2. Experiment 1**

### **3.2.1. Method**

#### **3.2.1.1. Participants**

Twenty-eight 3-year-old children ( $M = 38$  months; 34 – 45 months; 16 girls) participated in the study. Children were native German speakers, came from mixed socio-economic backgrounds from a large German city and were recruited via urban daycare centers (in which testing took place). Parents provided written informed consent. Six additional children were tested, but excluded from the final sample due to experimenter error (3), language and comprehension difficulties (1), and uncooperativeness (2).

#### **3.2.1.2. Design**

In a within-participant-design, children received a short warm-up session, followed by a third party fairness task and two prosocial tasks. The two prosocial tasks were presented either before or after the fairness tasks. Additionally, the order of the two prosocial tasks was systematically varied. Finally, the children performed a vocabulary test at the end of the session.

#### **3.2.1.3. Procedure**

Two experimenters conducted the study, which lasted roughly 20 minutes. At the start of the session, the first experimenter (E1) introduced two hand puppets (owl and bear), that were animated by the second experimenter (E2). The child, E1, and E2 sat at a table, the child to the left of E1 and E1 vis-à-vis to E2.

### *Third-party Fairness Task*

The third-party fairness task was based on a previous study (Rakoczy et al., 2016). E1 played the role of the coordinator of the situation, brought out four resources (different fruit types: strawberry, apple, orange, and pear, in counterbalanced trial order) and instructed the distributor puppet to allocate the items between himself and the recipient puppet (owl and bear in both roles in counterbalanced order). The child as unaffected bystander witnessed the allocation between the two hungry puppets, which consisted of four distribution trials: one baseline trial with equal distribution (2:2) and three test trials with unequal distribution (3:1). Neither E1 nor both puppets used any fairness words (e.g., distribute, allocate, fair, divide etc.) or any emotional cues to find out, if the child already held a normative fairness concept and to avoid any influence to the child's protest behavior in a direction to act fair. E1 announced that he had important paperwork to do, and turned away. According to an agreement between the two puppets, the recipient asked the distributor to take care of the fruits and went into his bed to sleep, which stood beside the distribution situation on the table. The other puppet began to distribute the four items slowly, always started with the recipients' plate. When he finished the distribution, he looked at the two plates for 5 s. After this short pause, he moved the two plates closer to the child and asked: "Like this?" Hence, the child had the opportunity to protest or correct against the outcome of the resource allocation. The situation always ended up in the same way: the recipient woke up and regarded the distribution; E1 turned back and cleared up the items.

### *Emotional Sharing Task*

The emotional sharing task was based on a former prosocial situation task (Vaish et al., 2009) and was adapted in the following way. E2 and the child sat on the table, E2 to the

left of the child. E1 came to the table with three mice in the hand and said, “Look what I found, mice!” Then at the same time she handed two mice to the child and one mouse to E2. Directly after that, E1 placed a box on the table in front of E2 and disappeared out of the focus of the child. E2 played happily with her mouse and the box, without engaging with the child. About 10 s later, E2 accidentally lost her mouse through a hole in the box, pointed to the hole, and spoke in a sad tone, “Oh no, my mouse!” She then tried to grasp it out of the hole. Once her effort remained unsuccessful, she leant over the box and sighed. From this moment on, she was vocally and facially obviously sad. E2 looked slightly at the direction of the child, but never focused directly the face, child’s hands or its mice to prevent giving her hints or pressuring her to help. During the next 2 min (from the moment E2 leant over the box and sighed), children’s behavior was coded (see below). After the 2 min, E2 opened her box, got her mouse back, and was again obviously happy. The prosocial situation did not last the full 2 min (a) if children became very upset, in which case the study was cut short and E2 opened the box to get back her mouse, or (b) if children handed one or both of their mice to E2, in which case E2 gratefully took and then handed back mouse/mice before opened the box and got back her own mouse.

### *Costly Sharing Task*

The costly sharing task was adapted from a previous sharing task (Schmidt & Sommerville, 2011). E2 waited silently at the back. E1 sat vis-à-vis to the child at the table. E1 placed two toys (Lego brick, stuffed cat, both same sizes) on a wooden tray 100 cm apart (position counterbalanced) and asked the child to pick one of the two toys (labeled as preferred toy). Directly after the decision, E1 gave the second toy (non-preferred) to the child and disappeared with the tray in the back. After that E2 came back, sat in front of the infant, looked directly at the child’s face and asked her for a toy



in an alternating manner between “can I have one?” and “can I have one, please?” every five seconds for up to 25 seconds.

### *Spontaneous Intervention*

During a short warm-up phase, E1, the distributor puppet, and the child played first with a ball and then with two instrumental tasks (hammer task, disk-and-peg task) in order to familiarize the child with the puppet. In the instrumental tasks, children had the opportunity to intervene spontaneously and correct the protagonist who made a mistake with conventional toys (trying but failing to push a ball into cuboid with his nose; putting a disk vertically onto a peg, which would not fit).

### *Verbal Intelligence*

For measuring verbal intelligence level, children were administered the vocabulary subtest of the Kaufmann Assessment Battery for Children - Second Edition (KABC-II, Kaufmann, 2015) for children and adolescents aged 3 to 18 years. Children had to name the objects on a series of pictures shown to them.

#### **3.2.1.4. Coding and Reliability**

All sessions were recorded, transcribed, and coded from videotape by a single observer. A second independent observer, blind to the hypotheses and conditions of the study, transcribed and coded a random sample of 25% of all sessions for reliability, except for verbal intelligence score.

### *Protest*

Children’s spontaneous behavioral interventions and verbal utterances were transcribed and given one of the following codes. A given response was coded as “protest” if the

child clearly intervened by using normative vocabulary (e.g., “No! This is not fair!”; “You have to do it differently!”; “This is not right!”), if the child expressed an imperative to the distributor without using normative vocabulary (“No, he needs one more!”; “Look, there is one missing here!”), if the child corrected the distribution to equal or to unequal to the disadvantage of the distributor and if the intervention was clearly addressed to the distributor puppet, or if the child pointed to the fruits or plates. All other indifferent reactions were coded as “none”. Overall, each child received a summed score consisted of each protest behavior over the three test trials (0 - 3) and over the baseline trial (0 - 1). Interrater reliability was good, Cohen’s  $\kappa = .80$ .

### *Emotional Sharing*

Codes were used in hierarchical categories, with their associated scores in parentheses. A given response was coded as “helps/shares (3)” if the child fully approached E2 and clearly offered her one or both mice, put the mouse near E2, made help suggestions (e.g., “I help you to get the mouse out of the box”) or by indirect helping directed to E1 in an effort to draw E1’s attention to the situation. A given response was coded as “describes situation (2)” if the child described the situation verbally (e.g., “The mouse is lost/in the box”) or with gestures (e.g. pointing to box), if the child points out that he had a mouse or with gestural communication (e.g., pointing to own mouse), while looking not to E1 but to situation or E2. A given response was coded as “attends to the situation (1)” if the child watched E2 and the situation in a serious way, stopped play, went to E1 or moved away but continued watching E2. A given response was coded as “ignore situation (0)”, if the child showed no involvement or interest in the situation. In a different way from Vaish, Carpenter, and Tomasello (2009), since no child showed distress, this category was not included in analyses. Although children could show any

or all of these prosocial behaviors, for analyses, children's prosocial score consisted of each child's highest score. Interrater reliability was very good, Cohen's  $\kappa = 1$ .

### *Costly Sharing*

Codes were used in hierarchical categories with their associated scores in parentheses. If the child shared both toys with E1, then the response was coded as "both (3)". If the child shared the preferred toy with E1, then it was coded as "pref (2)". The sharing of the unpreferred toy was coded as "unpref (1)", whereas the code "none (0)" was given if the child shared nothing. Interrater reliability was very good, Cohen's  $\kappa = 1$ .

### *Spontaneous Intervention*

We controlled for children's overall tendency to react spontaneously in social interactions and its possible influence on the protest behavior in the fairness tasks, by measuring children's behavior during the warm up tasks. A sum score over the two warm up tasks was calculated, by counting how often the children reacted spontaneously (0-2). Interrater reliability was very good, Cohen's  $\kappa = .91$ .

### *Verbal Intelligence*

We calculated the standard deviation ( $SD = 15$ ) from the mean ( $M = 100$ ), to determine the verbal intelligence level. In the lower age range of the test (6 years and younger), the reliability measures for the subtests are calculated between .70 and .97. The internal consistency of the K-ABC-II core subtests also shows sufficient reliability.

#### **3.2.1.5. Statistical Analysis**

Statistical Analysis were run in R, version 3.6.1. (R Core Team, 2019). We used a Wilcoxon signed rank test (Wilcoxon, 1945) to compare the protest measures between

the one baseline trial (equal distribution) and the three test trials (unequal distribution). Since the protest measure violated assumptions of standard linear models (i.e., normally distributed errors), we used a GLM with poisson error structure for the protest measure over the three test trials. The full model included the two predictor variables emotional sharing and costly sharing (both z-transformed), three control variables gender, verbal intelligence, and spontaneous intervention (all z-transformed, except gender), and an offset term (log-transformed total valid number of test trials per child) to adjust for the number of opportunities children had to perform protest (i.e., response variables were treated as rates). Effects of interest and planned comparisons were tested using likelihood ratio tests (LRT; Dobson, 2002) by each comparing the fit of the full model with a fit of a reduced model without the predictor variables. Preliminary analyses found no effects of gender, verbal intelligence, and spontaneous intervention. Unstandardized parameter estimates (*b*), standard errors (*SE*), 95% confidence intervals (CIs) and odds ratios (*ORs*) were obtained from the full model. For non-parametric tests, we computed the effect size *r* (Rosenthal, 1994).

### **3.2.2. Results**

As predicted, children protested significantly more in the unequal (test) trials than in the equal (baseline) trial,  $IM_{\text{unequal}} = 1.33$ ,  $IM_{\text{equal}} = 0$ ,  $z = -3.62$ ,  $p = .000$ ,  $r = -0.68$ , which is a large effect size according to Cohen's convention (1992). More precisely, none of the children ever protested in the baseline trial. Thus, we built a subset including the protest measures as dependent variable over the three test trials. For protest behavior, there was a significant main effect of emotional sharing,  $\chi^2(1) = 4.76$ ,  $p = .029$ ,  $b = 0.50$ ,  $SE = 0.25$ ,  $CI [0.01, 0.97]$ ,  $OR = 1.64$ . However, no main effect of costly sharing was found,  $\chi^2(1) = 0.82$ ,  $p = .365$ ,  $b = 0.15$ ,  $SE = 0.16$ ,  $CI [-0.17, 0.47]$ ,  $OR = 1.17$  (see Figure 4).

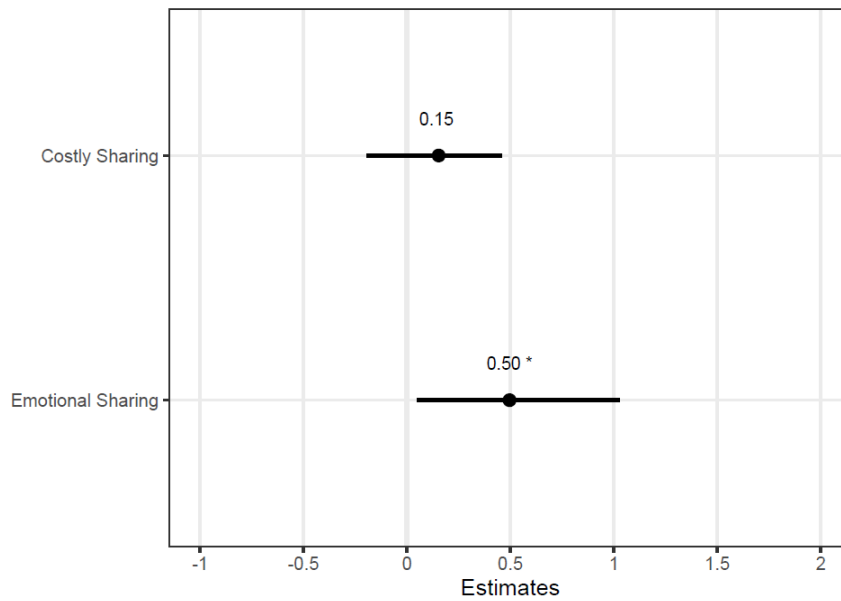


Figure 4. Experiment 1 revealed that emotional sharing, but not costly sharing, had a significant effect on children’s protest behavior against unequal distributions. Results are represented using estimates (beta coefficients) and their CIs from GLM (poisson), \*  $p < .05$ .

### 3.2.3. Discussion

In Experiment 1, we were interested in whether 3-year-old children understand fairness as a normative notion. Our results revealed that although not giving them any fairness cues, children, who observed a third-party distribution between two hand puppets, protested and intervened against unequal distributions, but interestingly none of the children ever protested against equal distributions. Moreover, we also asked which motivational factors (empathic concern, altruism) are associated with children’s responses against unequal distributions. Results show, that emotional sharing, but not costly sharing is related to children’s verbal and behavioral utterances, indicating that compassionate children are more willing to react against unfair allocations. Our results thus are in line with previous findings (Hepach, 2017; Jensen et al., 2014; Rakoczy et al., 2016; Vaish et al., 2009), and provide evidence that other-regarding concerns such

as sympathy are key motivations of children's normative fairness expectations. In sum, since we used Experiment 1 as a baseline measure, we investigated the normative notion of children's fairness expectations more systematically in Experiment 2. We were therefore interested whether children's normative expectations of fairness are simply based on conventions such as simple regularities and habits or whether children have a deeper moral understanding of fairness, which take the welfare of others into account. In this context, we use Turiels criteria "authority independency". Assuming that 3-year-olds already have a normative understanding of fairness, we expected that children would protest and intervene against unequal distributions despite the authority's demand to act unequally. Secondly, we expected that children's reactions to unequal distributions would be related to empathic concern (for the victim) and not to general altruistic tendencies.

### **3.3. Experiment 2**

#### **3.3.1. Method**

##### **3.3.1.1. Participants**

Twenty-eight 3-year-old children ( $M = 41$  months; 36 – 47 months; 14 girls) participated in the study and were recruited and tested as in Experiment 1. Three additional children were tested, but excluded from the final sample due to uncooperativeness (3).

##### **3.3.1.2. Design**

The number and order of tasks in a within-participants-design (warm-up session, fairness task, and prosocial tasks) was identical to Experiment 1.

### **3.3.1.3. Procedure**

#### *Third-Party Fairness Task*

The general procedure was almost the same as in Experiment 1 (e.g., roles and seating arrangements of E1, E2 and hand puppets; 4 different fruit types; 4 distribution trials). In addition to Experiment 1, we were interested, if children continue to protest even when an authority prompts the puppets to follow an unequal distribution. Thus, E1 no longer played the role of a neutral coordinator, but rather established a dictate in its role as authority to distribute resources unequally. Hence, in the first sub-phase, E1 brought out four fruits and stipulated the distributor puppet to allocate the items unequally between him and the recipient puppet to distributor's advantage. Afterwards he turned away. In response, both puppets were surprised by the authority's demand to act unequal. Supplementary, the recipient puppet was obviously sad. Like in Experiment 1, the recipient asked the distributor to allocate the four resources before he went to bed. The second sub-phase consisted of four distribution test trials, two trial with equal distribution (2:2) and two trials with unequal distribution (3:1). Thus, the child as unaffected bystander had the opportunity to intervene and/or protest, directly to the authority in the sub-phase 1, and/or, against the distributor puppet in the sub-phase 2, who followed or violated the rule of inequality. The situation ended up in the same way as in Experiment 1.

#### *Emotional Sharing task*

The general procedure was identical to Experiment 1 with the following exceptions. At the beginning, E2 brought out his favorite toy (a stuffed bear) and explained that his toy always comforts him, when he is sad. Afterwards he cuddled the bear and placed it between himself and the child. Then E1 came back and handed one mouse to E2 and

one mouse to the child. The procedure than followed the same way as in Experiment 1, where E2 lost his mouse in the box and the reactions of the child were coded. Thus, the child had the opportunity to act prosocial by helping to get the mouse of the box, comforting with the teddy, or sharing his own mouse with E2.

#### *Costly Sharing Task*

The general procedure was identical to Experiment 1 with following transformations. First, we replaced the stuffed cat with a same sized plastic figure, to get the same texture as the Lego brick. Second, E2 asked for the toys with the words, “Oh, can I have that?”, and “Can I have that, please?”

#### *Control Measures*

As in Experiment 1, spontaneous interventions were conducted during the warm up session. Verbal intelligence level was assessed with the Kaufmann Assessment Battery for Children - Second Edition (KABC-II) (Kaufmann, 2015).

#### **3.3.1.4. Coding**

Coding was the same as in Experiment 1. Interrater reliability was very good (Cohen’s  $\kappa = .88$ , protest;  $\kappa = 1$ , emotional sharing;  $\kappa = 1$ , costly sharing;  $\kappa = .83$ , spontaneous intervention).

#### **3.3.1.5. Statistical Analysis**

Analyses were carried out as in Experiment 1. To account for non-independence of observations due to our within-participants design (each child participated in each condition), we used generalized linear mixed models (GLMMs) with binomial error structure that allow for the inclusion of both fixed and random effects (Baayen et al.,



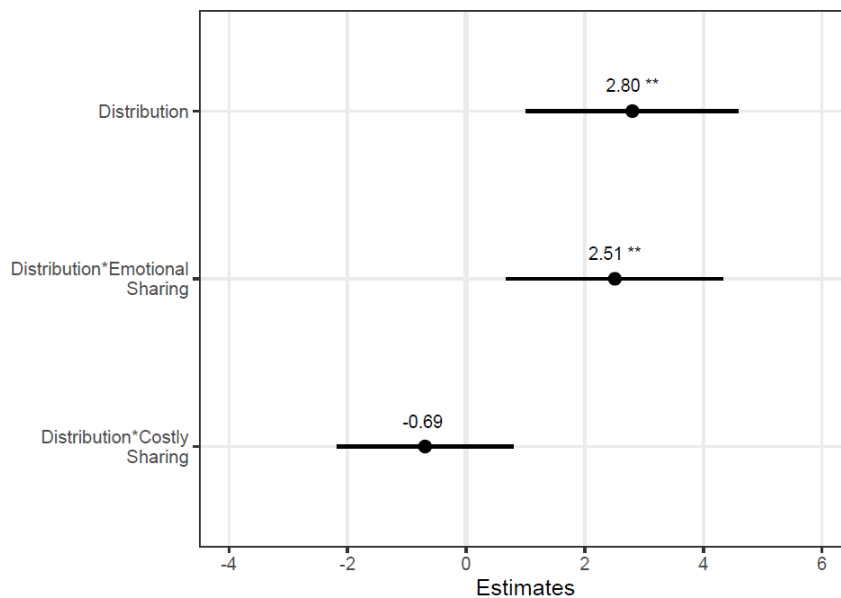
2008; Bates et al., 2014). The full model included the predictor variables distribution (equal vs. unequal), emotional sharing and costly sharing (both z-transformed), the interaction of emotional sharing x distribution, the interaction of costly sharing x distribution, the three control variables gender, verbal intelligence, and spontaneous intervention (all z-transformed, except gender) as fixed effects and participant ID as a random effect. First, the combined significance of the predictor variables was tested by comparing the fit of the full model (including the interaction effects, predictor variables and control variables) with the fit of a null model that only contained the control variables using LRTs. The approach of testing the overall full model against a null model helps to protect against Type I error inflation arising from models comprising more than one predictor variable (Forstmeier & Schielzeth, 2011). This full-null model comparison was significant,  $\chi^2(5) = 31.42, p < .001$ . Next, effects of interest and planned comparisons were tested using likelihood ratio tests (LRT; Dobson, 2002) by each comparing the fit of the full model with a fit of a reduced model without the predictor variables. Preliminary analyses found no effects of gender, verbal intelligence, and spontaneous intervention. In order to test for the main effect of emotional sharing and costly sharing on direct protest against the authority, we used LRT in a GLM with poisson error structure. Unstandardized parameter estimates ( $b$ ), standard errors ( $SE$ ), 95% confidence intervals ( $CI$ s), and odds ratios ( $OR$ s) were obtained from the full models.

### **3.3.2. Results**

#### **3.3.2.1. Overall Protest**

Overall, children protested significantly more in the unequal distribution condition than in the equal distribution condition,  $\chi^2(1) = 18.96, p < .001, b = 2.80, SE = 0.91,$

CI [1.30, 5.16],  $OR = 16.49$  (see Figure 5). Moreover, we found a significant interaction effect of distribution and emotional sharing  $\chi^2(1) = 10.70, p = .001, b = 2.51, SE = 0.93$ , CI [0.92, 4.81],  $OR = 12.27$  (see Figure 5), suggesting that the higher children scored in emotional sharing, the more they tended to protest in the unequal distribution condition as compared with the equal distribution condition (see Figure 6, for an overview of the protest behavior of the four different emotional sharing groups). Furthermore, no significant interaction effect between costly sharing and distribution was found,  $\chi^2(1) = 0.87, p = .35, b = -0.69, SE = 0.76, CI [-2.31, 0.76], OR = 0.50$  (see Figure 5).



*Figure 5.* Experiment 2 indicated that the children protested significantly more against the unequal versus equal distribution. Moreover, despite an authority's command to act unequally, emotional but not costly sharing had a significant effect on children's protest behavior against unequal distributions. Results are represented using estimates (beta coefficients) and their CIs from GLMM (binomial), \*\* $p < .01$ .

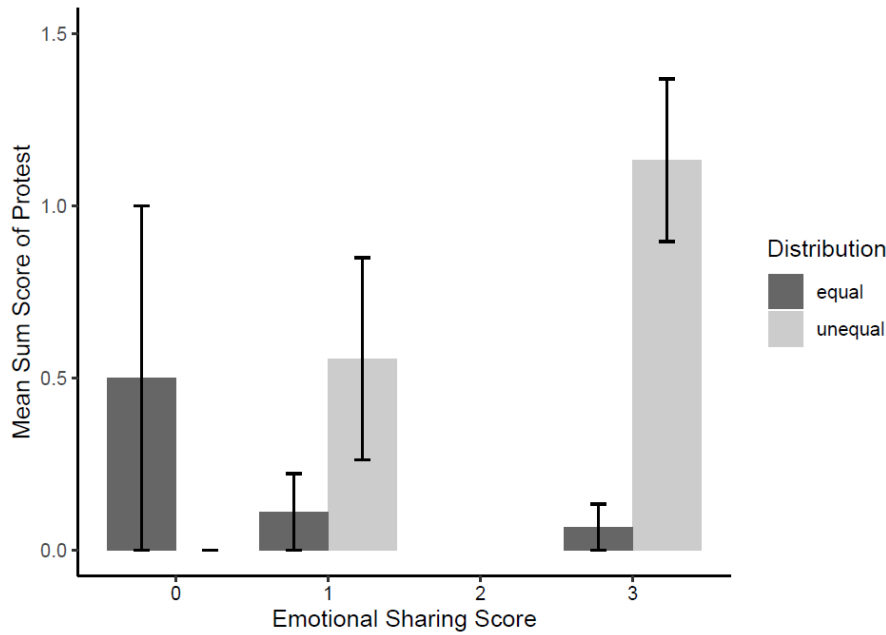


Figure 6. Mean sum score of protest for the four different emotional sharing groups (0 = ignore; 1 = attend; 2 = describe; 3 = help/share/comfort) and the distribution condition (equal versus unequal distribution).

### 3.3.2.2. Protest Directed at Authority

In order to test if emotional sharing or costly sharing is related to the protest measure directed at the authority, we ran further analysis. Once again, emotional sharing had a significant effect on protest behavior against the authority,  $\chi^2(1) = 4.93, p = .036, b = 0.59, SE = 0.28, CI [0.04, 1.14], OR = 1.80$ . Again, there was no main effect of costly sharing,  $\chi^2(1) = 0.52, p = .470, b = -0.17, SE = 0.23, CI [-0.62, 0.29], OR = 0.85$ .

### 3.3.3. Discussion

In Experiment 2, we were interested whether children's normative expectations about fairness have a moral dimension by using Turiel's authority-independency criteria (Turiel, 1983, 2008). Results indicating, that despite the authorities demands to act unequally, children protested and corrected against the norm follower who acted

unequally, but not against the norm violator who acted equally. In addition, we have also examined the direct protest against authority and found that children also showed protest in this context, which is a strong reaction because it might cost a lot of effort for the children, especially when it is necessary to protest against an adult (experimenter) whom they do not know. In sum, these findings give further evidence that 3-year-old children regarded fairness as a moral component rather than a simple behavioral regularity. Moreover, we replicated our findings from Experiment 1 and found that children's prosocial motivation in response to an unfair offers was associated with emotional sharing, but not with costly sharing. In Experiment 3, we were particularly interested in the extent to which children act against inequality by measuring the costly punishment behavior of children as restorative justice. In other words, if children are willing to sacrifice their own resources, they would restore justice between the victim and the perpetrator. In addition, we have suggested that altruistic tendencies measured by children's first-party costly sharing behavior (and not the emotional sharing, empathic concern) should be related to the third-party costly punishment behavior.

### **3.4. Experiment 3**

#### **3.4.1. Method**

##### **3.4.1.1. Participants**

Twenty-eight 3-year-olds ( $M = 44$  months; 36 – 49 months; 14 girls) participated in the study and were recruited and tested as in Experiment 1 and 2. Three additional children were tested, but excluded from the final sample due to difficulties in language comprehension (1), uncooperativeness (1), or technical error (1).

### **3.4.1.2. Design**

The number and order of tasks in a within-participants-design (warm-up session, fairness task, and prosocial tasks) was identical to Experiment 1 and 2.

### **3.4.1.3. Procedure**

#### *Third-Party Fairness Task*

The general procedure was identical to Experiment 1 with the following exceptions. As in Experiment 1, E1 played the role of coordinator in the situation. In the introduction phase, he took out five stickers and explained that these now belong to the child. Then E1 placed the stickers in a box and put the box right beside the child. Afterwards, he brought out six fruits and instructed the distributor puppet to allocate the items between himself and the recipient puppet. As in Experiment 1, the distributor allocated the fruits in four trials, one baseline equal distribution trial (3:3) and three unequal distribution tests trials (5:1). In contrast to the test trial in Experiment 1, the distributor intended to act unfairly.

Immediately after the distribution, E1 turned back. In the baseline trial, E1 asked the child, "Well, look, the bear has exactly as many fruits as the owl. Both have the same number. Is that good or bad?" In the test trial, E1 described, "Well, look, the bear has more fruits than the owl. Is that good or bad?" In this context, it is important to mention that only the "bad" answers of the child in the test phase led to the opportunity to utilize own stickers to punish the unfair distributor by redistribute the same number of fruits from the distributor to the recipient. Thus, E1 opened the box and asked the child to give away up to five stickers. Here, the child had the chance to decide by putting as many stickers as it liked in the lid of the box. Along these lines, E1 removed the same number of fruits from the distributor to hand them to the recipient. At the end

of each trial, E1 took out another wooden box and told "Then we put your remaining stickers here in your treasure chest."

#### *Emotional Sharing task*

The procedure was identical to Experiment 2.

#### *Costly Sharing Task*

The procedure was identical to Experiment 2.

#### *Control Measures*

As in Experiment 1 and 2, spontaneous interventions were conducted during the warm up session. Verbal intelligence level was assessed with the Kaufmann Assessment Battery for Children - Second Edition (KABC-II) (Kaufmann, 2015).

#### **3.4.1.4. Coding**

All sessions were recorded, transcribed, and coded like in Experiment 1 and 2. In the third-party fairness task, we counted the number of stickers the child was willing to give away to punish the distributor. Furthermore, we coded the answer of the children to the explicit question, if the distributor acted good or bad. Interrater reliability was very good ( $\kappa = 1$ , stickers;  $\kappa = 1$ , explicit question). For the prosocial tasks as well as the spontaneous intervention, coding was the same as in Experiment 1 and 2. Interrater reliability was good to very good ( $\kappa = 0.78$ , emotional sharing;  $\kappa = 1$ , costly sharing;  $\kappa = 1$ , spontaneous intervention).

### **3.4.1.5. Statistical Analysis**

Analyses were carried out as in Experiment 1 and 2. First, we were interested in whether the proportion of children who answered both explicit questions correctly deviated significantly from chance (50%). Thus, we conducted an exact binomial test and a planned exact one-sample Wilcoxon signed-rank test. For costly punishment, we used GLMMs with poisson error structure. The full model included as fixed effects the predictor variables emotional sharing and costly sharing (both z-transformed), three control variables gender, verbal intelligence, and spontaneous intervention (all z-transformed, except gender), trial (z-transformed) as random slope and subject ID as random intercept. Effects of interest and planned comparisons were tested using likelihood ratio tests (LRT; Dobson, 2002) by each comparing the fit of the full model with a fit of a reduced model without the predictor variables. Preliminary analyses found no effects of gender and spontaneous intervention. Unstandardized parameter estimates (*b*), standard errors (*SE*), 95% confidence intervals (*CI*s), and odds ratios (*OR*s) were obtained from the full models. For non-parametric tests, we computed the effect size *r* (Rosenthal, 1994).

## **3.4.2. Results**

### **3.4.2.1. Explicit Question**

An exact binomial test indicated that the proportion of children of .96, *CI* [0.84, 1], who correctly answered “good” in the equal distribution trial was significantly higher than the expected chance level .50,  $p < .001$  (one-sided). Moreover, a planned exact one-sample Wilcoxon signed-rank test revealed, that the proportion of children (.83), who correctly answered “bad” in the unequal distribution trial was not higher than chance level .50,  $V = 239$ ,  $p = .193$  (one-sided). In order to examine the latter result, we

subdivided children's responding to the explicit questions in three groups: a correct answers group (all unequal distribution trials were judged as bad), a mixed answers group (not all unequal trials were judged as bad), and an incorrect answers group (all unequal trials were judged as good) and ran further control measures (see below).

#### **3.4.2.2. Costly Punishment**

Unlike in Experiment 1 und 2, we found a significant main effect of costly sharing on children's third-party costly punishment behavior,  $\chi^2(1) = 4.43, p = .035, b = 0.57, SE = 0.24, CI [0.09, 1.05], OR = 1.77$ , and did not find a significant main effect of emotional sharing,  $\chi^2(1) = 0.13, p = .718, b = -0.09, SE = 0.26, CI [-0.61, 0.42], OR = 0.91$  (Figure 7). Furthermore, we found a significant main effect of the vocabulary test,  $\chi^2(1) = 5.99, p = .014, b = -0.73, SE = 0.21, CI [-1.13, -0.32], OR = 0.48$ , which indicated that children, who higher scored in the vocabulary test, are less willing to costly punish. We therefore ran further control analysis to exclude the influence of language competence and used the three explicit question groups.



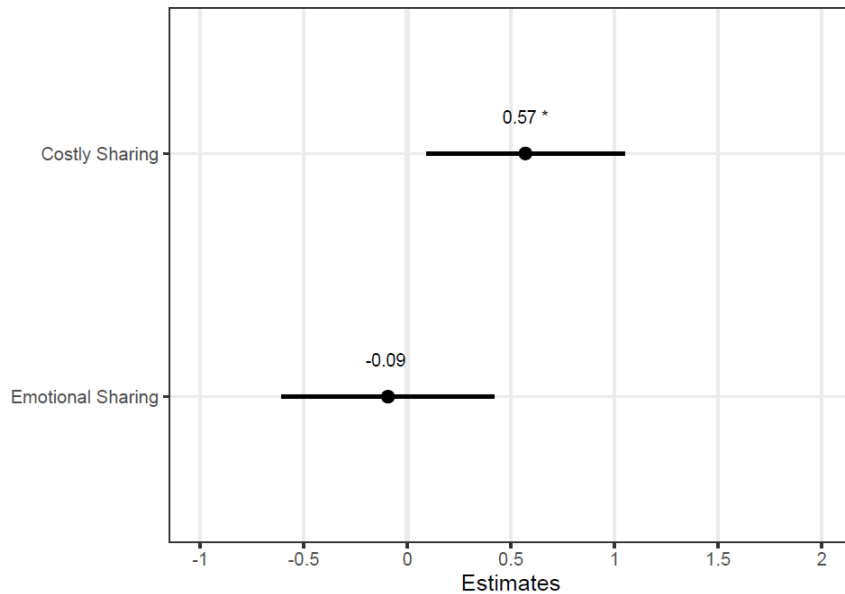


Figure 7. Experiment 3 showed that children's third party costly punishment behavior of an unfair distributor is associated with children's first party costly sharing behavior. Results are represented using estimates (beta coefficients) and its CIs from GLMM (poisson), \* $p < .05$ .

### 3.4.2.3. Control Analysis

Our first control question was whether the three different groups (correct, mixed, incorrect responding to the explicit question) differed in their language competence. We therefore used a GLM with gaussian error structure with the predictor group and tested the main effect group on the dependent variable vocabulary test using LRT. Results revealed no significant main effect on group,  $F(1) = 1.15$ ,  $p = .294$ . These findings indicate that the three groups did not differ in their language comprehension skills ( $M_{correct} = 96.54$ ,  $SD_{correct} = 13.45$ ;  $M_{mixed} = 100.83$ ,  $SD_{mixed} = 16.93$ ;  $M_{incorrect} = 102.11$ ,  $SD_{incorrect} = 6.412$ ). Second, for children's costly punishment behavior, we used a GLM with gaussian error structure with the predictor group and found a significant main effect,  $F(1) = 7.83$ ,  $p = .010$ . The Bonferroni post hoc t-tests revealed, that at least the correct responding group ( $M = 0.35$ ,  $SD = 0.34$ ) differed significantly from the incorrect responding group ( $M = 0.00$ ,  $SD = 0.00$ ,  $t(12) = -3.75$ ,  $p = .027$ ). The other

comparisons to the mixed responding group ( $M = 0.28$ ,  $SD = 0.37$ ) did not yield any significant results. Third, we were interested whether the competent children in the explicit question group also acted more altruistically. Kruskal-Wallis rank sum test indicated a significant main effect of group,  $H(2) = 10.86$ ,  $p = .004$ . The Bonferroni post hoc tests (Dunn's-test) revealed that the children from the correct answer group ( $Mdn = 19.50$ ) significantly differs from the mixed answer group ( $Mdn = 10.67$ ,  $p = .050$ ) and from the incorrect answer group ( $Mdn = 9.83$ ,  $p = .008$ ). Overall, the results suggest that (i) the responding to the explicit question is not dependent on language skills, and (ii) that the three different groups reflect individual differences of children's fairness expectations and costly sharing behavior.

### **3.4.3. Discussion**

In Experiment 3, we investigated whether children are willing to sacrifice their own resources to punish the perpetrator and restore justice between the victim and the perpetrator. Our results indicated that children indeed responded in line with our considerations. Interestingly, in that costly restorative justice situation, children's motivations is positively related with their own costly sharing behavior, which can be interpreted as general altruistic tendencies (cf. Jensen et al., 2014; Schmidt & Sommerville, 2011). Moreover, competent children, who answered all questions correctly (judging equal distributions as good and unequal distributions as bad) are also those who showed more punitive behavior and were more altruistic than the incompetent children (answering all questions wrong). However, this does not seem to depend on linguistic abilities.

### 3.5. General Discussion

Our results provide clear evidence that 3-year-old children have a moral concept of fairness in terms of distributive justice. Three experiments revealed that children enforce third-party fairness norms as unaffected bystanders, even without giving them indications of acting fairly (Experiment 1: normative understanding of fairness), that they protest and intervene against injustice even when an authority dictates inequality (Experiment 2: moral norm understanding: authority-independency), and finally that they are even willing to sacrifice their own resources in order to punish unfair behavior in the sense of restorative justice (Experiment 3: costly punishment). Moreover, children's normative expectations of fairness are closely linked to their own prosocial behavior, particularly to the concern for the welfare of others (Experiment 1 & 2: sympathy, empathic concern) and to an altruistic interest in punishing norm violators (Experiment 3).

Thus, our results go beyond prior research on children's norm understanding about fairness in many ways. Previous findings suggested that preschool children have descriptive expectations about equal resource allocation (Geraci & Surian, 2011; Meristo et al., 2016; Schmidt & Sommerville, 2011). A recent study also provided evidence about fairness as normative expectations in 3-year-old children, but these results were linked with joined merit outcomes (Rakoczy et al., 2016). Moreover, from a moral-motivational perspective (Killen & Smetana, 2015; Smetana & Braeges, 1990; Turiel, 1983, 2008), 3-year-old children are capable to distinct between conventions and moral issues at least to some of Turiel's criteria. And 3-year old children protest and react against a norm violator in conventional as well as moral contexts as unaffected observers (Rakoczy et al., 2008; Rossano et al., 2011; Schmidt & Tomasello, 2012; Vaish et al., 2011). We replicated earlier findings (Rakoczy et al., 2016) and gave

additional evidence (Experiment 1) by showing that children protest even without giving them any other influences or indications to act fair. Experiment 2 showed that children's expectations of fairness not only have a conventional dimension, e.g. as a simple behavioral standard, but also a moral dimension in many respects. First, the children protested against a norm follower who adhered to a prescribed rule. Secondly, the children protested directly against an authority who had established the rule. Thirdly, the children did not protest against a norm violator, who equally shared. If there was only an interest in maintaining fairness as a conventional rule, the protest pattern should have been oppositely by showing that children simple protest a norm violator irrespectively what rule is followed. Thus, our findings are in line with social domain theory (Smetana & Braeges, 1990; Turiel, 1983, 2008), which claims that adults and children explicitly judge moral from conventional transgressions along various criteria, in particular concerning authority-independency. Moreover, a plethora of studies showed that 3-year-old children are generally capable to intervene and protest against norm violations, for instance against conventional game rule violations (Rakoczy et al., 2008) as well as against moral transgressions (Rossano et al., 2011; Vaish et al., 2011). Furthermore, Experiment 3 indicated the strength of moral norm understanding, by investigating how far the children would go to punish unfair behavior and to restore fairness. The results showed that the 3-year-olds were willing to sacrifice their own resources to punish and restore unfair behavior. These results are consistent with previous findings showing third-party costly punishment behavior in preschoolers against moral violations (Kenward & Östh, 2012; Krasnow et al., 2016; McAuliffe et al., 2015; Robbins & Rochat, 2011). This is a new aspect, as our findings are in line with considerations on restorative justice, which focus more on the collectivist aspect of fairness. In these considerations, repairing fairness is used as a normative term that involves all affected parties (perpetrator and victim), not simply punishing the

perpetrator (Heffner & FeldmanHall, 2019; Riedl et al., 2015; Wenzel et al., 2008). We created a paradigm that not only offered the possibility of simply punishing the perpetrator, but also of restoring justice for the victim by giving the opportunity to transfer resources, not just taking them away. Here we showed that competent children, which were better in answering the explicit question correctly, acted also more altruistically and were those with stronger punishment and restorative justice tendencies. However, results from Experiment 3 are ambiguous and give rise to further investigations. The motivation which underlies the behavior of those children who answered the explicit question (whether the unfair distribution is bad) in a mixed fashion is still unclear. It might be possible, that they played strategically due to egotistic motives, but possess normative fairness considerations, and simply were not willing to sacrifice own resources to repair equal treatment. It would be interesting to investigate what their motivations are related to. Furthermore, we also conducted a mixed paradigm with retributive as well as restorative justice considerations. It might be interesting to examine children's behavior only in a restorative justice situation. Furthermore, more research is needed to disentangle the phenomena of children's early sense of fairness. An interesting question for future research might be, when exactly infants' develop genuine normative expectations about fairness. A suitable paradigm using eye-tracking may show that children younger than 3 years of age are capable to differentiate between moral and conventional norms. One could for example investigate the potential influence of intelligence, executive functions, inhibition or theory of mind capacities on prosocial behavior and fairness. Finally, our results give evidence about a specific moral norm, such as fairness. However, our results do not answer aspects to other types of social norms, for instance harm-related moral issues.

Overall, our findings may help to better understand to what extent the ontogeny of fairness norms can be characterized as moral and in how far it is associated with children's developing concern for the welfare of others in different contexts.

## **4. Study 3: Decategorization Leads Preschoolers to Treat In-Group and Out-Group Members Equally**

### **4.1. Introduction**

Humans are socio-cultural beings who depend on one another and who have a natural tendency to collaborate, cooperate, and to form groups with others, which are typically characterized by a common set of values, norms, and practices (Baumeister & Leary, 1995; Deci & Ryan, 2008; Fehr & Fischbacher, 2004; Schmidt et al., 2019; Tomasello, 2016). The motivational and normative forces behind that mutual altruism and cooperation are grounded evolutionarily, since a relatively small, interdependent community offers the most effective loyalty, trust and cooperation (Brewer & Caporael, 2006; Fiske, 2000; Tomasello, 2016). According to these considerations, if one is part of a cohesive group, the chances of exploitation are relatively low. In the end, the entire group, strengthened by mutual support and cooperation among community members, has the greatest chance of survival compared to other individuals. Moreover, the identification with the preferred (in-)group is at the same time coupled with a social comparison with other (out-)groups, to which one automatically does not belong. This typically leads to social categorization and inter-group distinction, accompanied by positive emotional attachment and consensus among members of the same group (Tajfel, 1982; Tajfel & Turner, 1979).

Group identification and categorization may lead to intergroup-bias, that is a systematic tendency, to evaluate the in-group more favourably than the opposing out-group (Hewstone et al., 2002). As the salience of category membership increases, not only in-group members are perceived as more similar to each other and thus become part of one's own self-concept, but out-group members are also be seen as a

homogeneous whole, which ultimately leads to depersonalization and dehumanization (Brewer, 1996, 1999; Haslam, 2004; Tajfel, 1969; Turner et al., 1994). Intergroup-bias takes place under various circumstances, even without social interaction, in anonymous situations, and in the absence of conflicts and competition (Abrams, 2010; Crisp & Turner, 2009; Mullen et al., 1992). However, there is much disagreement whether intergroup-bias leads exclusively to in-group favouritism or whether it goes along with active forms of out-group derogation. In the first case (in-group favouritism) the out-group is simply ignored as a by-product (Brewer, 1999; Hewstone et al., 2002; Park & Judd, 2005; Tajfel, 1969). Whereas the latter case (out-group derogation) includes intentional aggressive behaviour and negative emotions toward the other group, such as discrimination, prejudice and stereotyping (Hewstone et al., 2002; Miglietta et al., 2014; Park & Judd, 2005; Sumner, 1906). A plethora of developmental investigations focuses on social categorization processes. Such studies show that intergroup-bias robustly takes place in the late preschool years in minimal group settings (Dunham et al., 2011), manifests in various prosocial forms like sharing or helping behavior (Benozio & Diesendruck, 2015; Fehr et al., 2008; Over, 2018; Sierksma et al., 2018) or in trust and loyalty (Misch et al., 2016; Rutland et al., 2015). Recent studies even pointed to an interplay between morality and group identity (Chalik & Dunham, 2018; Rutland et al., 2010).

In today's times of more and more heterogeneous societies and attempts to overcome prejudice and xenophobia, it is important to examine not only intergroup-bias, but also ways to overcome group thinking and foster equal treatment of individuals regardless of group membership. Over the years a variety of individual and intergroup approaches have been developed to reduce intergroup-bias (for an overview see: Beelmann & Heinemann, 2014; Hewstone et al., 2002; Paluck & Green, 2009). The most frequently investigated method in developmental research is the extended contact



approach, in which intergroup contact occurred indirectly through other in-group members (e.g., Vezzali, Hewstone, Capozza, Giovannini, & Wölfer, 2014; Vezzali, Stathi, Giovannini, Capozza, & Visintin, 2015). These interventions promote cross-group friendships between school children by showing that indirect or vicarious contacts are more effective than direct contacts.

Another promising candidate to overcome group-thinking is decategorization (Bettencourt et al., 1992; Brewer, 1996, 1997), which focuses on social identity and categorization processes. Decategorization aims to eliminate intergroup boundaries by reducing the salience of category membership and stressing out individual identity over group identity (e.g. emphasizing personal attributes) which enable more personalized, less homogeneous perceptions of in-group and out-group members. As a result, the long-term positive effects of decategorized contact are considered less useful and therefore less frequently used (Brewer, 1996; Brewer & Miller, 1984; Ensari & Miller, 2001; Gaertner et al., 1993). Only a few developmental studies examine decategorization as a possible intervention method. For example, Cameron and Rutland (Cameron et al., 2006, 2011; Cameron & Rutland, 2006) used decategorization, dual identity and common in-group identity in hypothetical scenarios, but in combination with extended contact hypothesis in preschoolers as well as school children. Interestingly, dual identity intervention approach in combination with the extended contact approach was the most effective way.

Taken all findings together, little is known about intergroup-bias and approaches to reduce them in early childhood. In particular, the direct and sole influence of decategorization without the combination with any other interventions and its impact to intergroup-bias has not yet been well explored. In most experimental designs, intergroup-bias and out-group prejudice are studied interchangeably. But there is much evidence that in-group favouritism and out-group derogation are separable phenomena

and that the origin of identification and attachment to in-groups is independent of intergroup conflict (Brewer, 1999; Hewstone et al., 2002; Paluck & Green, 2009). Hence, we try to disentangle these two phenomena of in-group favouritism and out-group derogation.

In the current study, we thus investigated preschoolers' responsiveness to decategorization by measuring children's prosocial behavior towards in-group and out-group members with a dictator game. The dictator game has turned out to be an effective method as measurement of costly sharing behavior in resource allocation contexts with children (Benenson et al., 2007; Blake & Rand, 2010). Nevertheless, only a few studies used the dictator game in intergroup contexts, but they all showed robust results of in-group preference as a function of prosocial behavior (Benozio & Diesendruck, 2015; Bettencourt et al., 1992; Fehr et al., 2008; Over, 2018). We have decided to base the experimentally induced dimension of category differentiation on the stimulus "geographical proximity". Additionally, we intensified the categorization process by inserting positive social identity of the relevant in-group, whereby the individual's knowledge belonging to certain social groups was combined together with positive emotions and values (Tajfel, 1969, 1982). A second stimulus "color" was used to make the identity of the category clearly visible to the children throughout the session (Bettencourt et al., 1992; Marcus-Newhall et al., 1993). Moreover, there is evidence that the cohesion within the in-group might be reduced, when the child interacts with a real interaction-partner by providing information that is contrary to the child's own attitude. Therefore, the children in our study were confronted with a hypothetical scenario in which they were asked to imagine their interaction partners (Crisp & Turner, 2009; Vezzali, Stathi, Crisp, et al., 2015).

Overall, we predicted that when using a dictator game, preschool children would behave more prosocially towards the in-group than towards the out-group in a typical

intergroup context, but that in a decategorized context, their intergroup-bias would decrease, thus leading them to treat in-group and out-group individuals in a similar way.

## **4.2. Method**

### **4.2.1. Participants**

Eighty 5- to 6-year-old children ( $M = 72$  months; 60 – 83 months; 40 girls) participated in the study. The participants were randomly assigned to two between-subjects conditions, with gender being evenly split (20 girls in each condition). Children were native German speakers, came from mixed socio-economic backgrounds from a large German city and were recruited via urban daycare centers and a museum in which testing took place. Parents provided written informed consent. Two additional children were tested, but excluded from the final sample due to experimenter error.

### **4.2.2. Material**

A yellow and a purple pad were used for a preference test. Two pictures of a crowd of stick man in yellow and in purple visualized the two groups (in-group vs. out-group). Two ovals painted on transparent foils were used as stencils, one of them was marked with a small dot to clarify, which one belonged to the participant. Two purple and two yellow paper boxes were utilized to store the stickers. One box was labeled with a white sticker on which the name of the participant was written in order to clarify to whom the content belonged. Six smiley stickers in orange and six stickers in green were used in the dictator game. In the personal condition, eight small tags displayed the following information about the recipient: the name (in-group: Nils and Vanessa; out-group: Nisa and Modessa), the favorite sweet (pictures of chocolate and gummy bear) and the

favorite toy (pictures of teddy bear and stuffed cat). The experimental set-up included a table and two chairs.

#### **4.2.3. Design**

A two (recipient identity: personal vs. impersonal; between-participants) by two (group membership: in-group vs. out-group; within-participants) mixed-design was used. As dependent measure, we counted the number of stickers (out of six) the child shared with each recipient (in-group vs. out-group member). The recipients of the participants were gender matched. The participants received six smiley stickers of the same color in a constant sequence across two trials (green, orange). The order of colors in the preference test (yellow vs. purple first), the order in which groups were introduced (in-group vs. out-group first), the trial order (member of in-group vs. out-group first), their preferences for sweets (chocolate vs. gummy bear) and toys (teddy bear vs. stuffed cat) were systematically varied.

#### **4.2.4. Procedure**

One experimenter conducted the study, which lasted roughly 10 minutes. The child and the experimenter sat vis-à-vis to each other at a table. The experiment started with a short warm-up phase, where the experimenter and the child played a puzzle game together in order to familiarize the child with the situation.

All children were first given an introductory phase (color preference test and introduction of in-group and out-group) followed by a presentation and test phase across two trials. In the presentation phase, children were presented with the recipient (in-group or out-group, within-participants condition), either in personalized or in impersonalized manner (between-participants condition). The trial ended with a test

phase in which children had the opportunity to share resources with the recipient (dictator game). The second trial then began once again with the presentation phase.

In the preference test, the experimenter laid a yellow and a purple pad on the table and asked the child, "Look, which color do you prefer, purple or yellow?" After the child's decision, the experimenter acknowledged the child's decision, "Great, you've chosen a wonderful color!". Then she introduced the two different groups. Hence, two verbal cues, one referring to color (purple vs. yellow) and one referring to the group's geographical proximity (own kindergarten vs. "Upendi", far away) were used to establish the in-group/out-group distinction. To present the in-group, the experimenter placed a same-colored in-group picture on the child's preferred colored pad and said, "Imagine they're children from your kindergarten! Therefore, you are the X [child's color preference: yellow or purple] group. And show me who do you want to be?" The experimenter marked the chosen stick man with the stencil. Then she familiarized the child with the out-group, placed the other colored group picture on the other pad and said: "These are children from another country. They are not from your kindergarten. They are Upendis and they come from far away. So they're the Y group [non-preferred color]." Afterwards, she introduced the child to the recipient, either in impersonal manner (focus on group membership) or in personal manner (focus on individuating information, name and personal preferences). In the impersonal-in-group condition, the experimenter said, "Look and he/she is also from your kindergarten, just like you." In the impersonal-out-group condition the child was told, "Look, and he/she is from Upendi, far away". The experimenter then marked the stick man from the corresponding recipient with a second stencil. In the personal condition, the procedure was the same, except that the child received more information about its recipient. The experimenter told, "Look, and this is [in-group: Nils/Vanessa vs. out-group: Nisa/Modessa], he/she also comes from [your kindergarten vs. Upendi]. [Nils/Vanessa vs. Nisa/Modessa] loves

[gummy bears vs. chocolate]. His/her favorite toy is a [cat vs. teddy bear]. The experimenter then marked the stick man from the corresponding recipient with a second stencil and placed a nametag and pictures with the sweet and toy next to the stick man. Afterwards the dictator game began. The experimenter fetched a box in the appropriate color (yellow or purple depending on the color of the in-group) and explained, "Look, this is your box, we'll stick your name on it". The experimenter placed the box next to the child's stick man. "And this is the box of that child from [your kindergarten vs. Upendi]." Now the second box (in-group color vs. out-group color) was placed next to the recipient's stick man. The child received six smiley stickers as piles in the middle between the two pads. The experimenter explained: "Look, here are six stickers! These are your stickers. Now, you can decide how many stickers you keep to yourself or give it to the child from [your kindergarten/Upendi vs. Nils/Vanessa/Nisa/Modessa]. You can keep them all or give them away. In addition, the stickers you want to keep come in your box here. The stickers you want to give to the child from [your kindergarten/Upendi vs. Nils/Vanessa/Nisa/Modessa], are coming into this box here." In order to prevent social desirability, the experimenter pretended to write something down and turned away. As soon as the child completed the decision, the experimenter turned back and put aside both boxes and the recipient's material. After that, the second round began. If the recipient was currently part of the in-group, then the second recipient was the out-group member and vice versa. The experimenter introduced the second recipient and then played once again the dictator game with the child.

#### **4.2.5. Statistical Analysis**

Statistical Analysis were run in R, version 3.6.1. (R Core Team, 2019). To account for the non-independence of the data (i.e., repeated observations per child), we used

generalized linear mixed models (GLMMs) with gaussian error structure and identity link function (Baayen et al., 2008; Bates et al., 2014). Initial models included as fixed effects the predictor variables recipient identity (personal vs. impersonal), group membership (in-group vs. out-group), and their interaction, the control variables gender and trial order (z-transformed), and participant ID as a random intercept. First, the combined significance of the predictor variables (i.e., main and interaction effects) was tested by comparing the fit of the full model (including the predictor variables, control variables, and random effect) with the fit of a null model that only contained the control variables and random effect using a likelihood ratio test (LRT; Dobson, 2002). The approach of testing the overall full model against a null model helps to protect against Type I error inflation arising from models comprising more than one predictor variable (Forstmeier & Schielzeth, 2011). This full-null model comparison was significant,  $\chi^2(3) = 10.35, p = .016$ . Next, effects of interest and planned comparisons were tested using LRTs. Preliminary analyses found no effects of gender and trial order. Unstandardized parameter estimates ( $b$ ), standard errors ( $SE$ ), and 95% confidence intervals (CIs) were obtained from the respective full model. Cohen's  $d$  ( $d$ ) was computed to estimate effect sizes.

### 4.3. Results

As predicted, we found a significant interaction effect between recipient identity and group membership,  $\chi^2(1) = 4.87, p = .027, b = 0.55, SE = 0.26, CI [0.03, 1.07], d = 0.47$  which is a small effect size according to Cohen's conventions. Planned comparisons revealed that children shared significantly more stickers with the in-group member ( $M = 2.28, SD = 1.34$ ) than with the out-group member ( $M = 1.7, SD = 1.18$ ) in the impersonal condition,  $\chi^2(1) = 7.82, p = .005, b = 0.58, SE = 0.2, CI [0.18, 0.97]$ , but that

they shared equally (in-group:  $M = 1.75$ ,  $SD = 1.13$ ; out-group:  $M = 1.73$ ,  $SD = 1.2$ ) in the personal condition,  $\chi^2(1) = 0.03$ ,  $p = .866$ ,  $b = -0.03$ ,  $SE = 0.17$ ,  $CI [-0.37, 0.31]$  (see Figure 8).

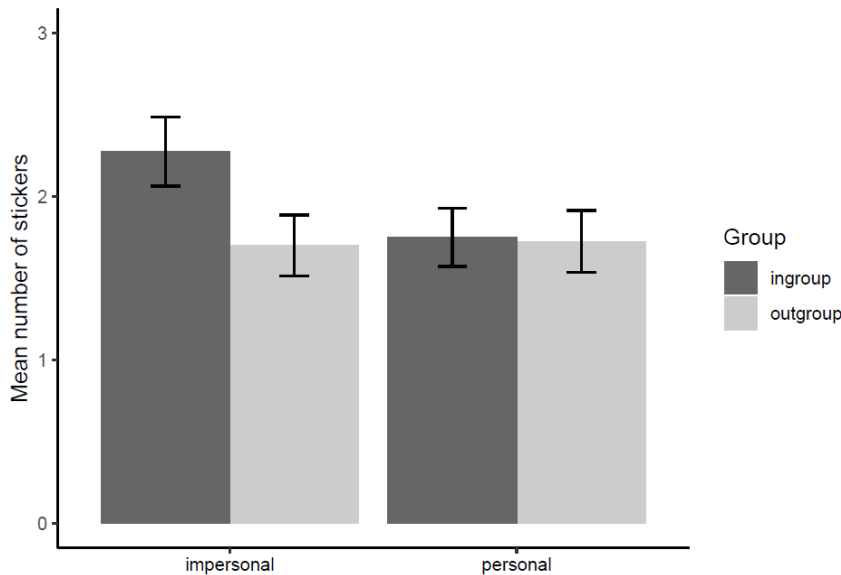


Figure 8. Mean number of stickers shared with in-group and out-group members as a function of condition. Error bars indicate standard errors of the mean.

#### 4.4. Discussion

Overall, the present study demonstrated the effectiveness of decategorization to reduce intergroup bias in preschool children. First, we show that in a typical intergroup context, children favor their own group over the out-group by sharing more stickers with in-group versus out-group members. Thus, we have replicated earlier results regarding intergroup bias in early childhood (e.g., Dunham et al., 2011). Second, when presenting group members in a decategorized manner, that is when the recipients were presented as individuals with certain preferences for toys and sweets, intergroup bias was diminished. In this case, children shared the same number of stickers with in-group as



well as out-group members. The results of our study are thus consistent with those from social psychology as well as developmental research (Bettencourt et al., 1992; Brewer, 1996, 1997; Cameron et al., 2006; Cameron & Rutland, 2006).

A further question arises in terms of the distribution pattern in the dictator game. Overall, the children acted rather selfishly than fair towards their recipients, regardless of group membership. As results revealed, children shared on average approximately one third of their resources. The only exception was children's preference of their own group in the impersonal intergroup situation, where they divided significantly more stickers to their group, but still remained under equal split. A possible explanation for this pattern is that preschool children simply show more egocentrism than adults do. On the other hand, various studies have shown that even preschool children are able to act fairly and punish even those who do not adhere to such norm (Fehr et al., 2008; Fehr & Fischbacher, 2004; Paulus & Moore, 2014). A better explanation is provided by game theoretical investigations in adulthood, where similar distribution pattern were found. Here, a meta-analysis of the dictator game revealed an average payout ratio of 28.3% to the recipient in first-person distributions (Engel, 2011). Moreover, in anonymous first person distributions is generally no willing to share resources equally for personal cost reasons. Indications from minimal group studies support this explanation, where participants who did third party rather than first party distributions, are more willing to act in the sense of fairness (Tajfel & Turner, 1979). In this context, the question arises how the children would have reacted to a third person distribution task, where they distribute stickers between in-group and out-group members without personal cost. This could be the subject of a possible follow-up study.

Interestingly, our results revealed, but at least on a descriptive level, that children shared only more resources with the anonymous in-group member as compared to all other conditions. These results give rise for further future directions. In particular, these

results might indicate that in-group favoritism and out-group derogation are separate phenomena and that the origin of identification and attachment to in-groups is independent of intergroup conflict. It would be interesting to investigate whether preschool children show in-group bias, but not active forms of out-group deviation. Moreover, in-group preference could rather be seen as a platform for possible later developing out-group derogation such as prejudice or discrimination, which under special circumstances could fully unfold in the course of development. These considerations also support the findings of Brewer's (Brewer, 1999) explanation of intergroup bias which is more interpreted as in-group bias but not active forms of out-group derogation. Social identity theory (Tajfel, 1982) also supports this theory, as in-group favoritism arises with positive reevaluation of the in-group as being part of the self-concept to heighten the own self-esteem. This process does not necessarily lead to the devaluation of the out-group. As a result, in-group members are then treated more positively as part of the self-concept regardless of the other groups. Hence, future investigations could try to disentangle these two phenomena of in-group favoritism and out-group derogation in early childhood.

A limitation of the study is a transfer of our findings in real life situations. First, decategorization is a suitable measure as a laboratory experiment, but further research is needed in the field. Our experiment was based on arbitrary and artificial categories, but we do not yet know what it is about deeply rooted prejudices, e.g. towards minorities as well as short lasting or long lasting effects. Second, it seems virtually impossible to completely deactivate category memberships in real life. Group formation arises automatically and switched on and off dependent of the situation, because it offers many advantages for the individual belonging to a certain small group (Brewer & Caporael, 2006; Park & Judd, 2005; Tajfel et al., 1971; Tomasello, 2016). On the long run, the aim should be to promote cooperation between groups without feeling threatened and to

recognize and appreciate diversity, tolerance and positive qualities of each group member being part of a greater community.

## 5. General Discussion

The present dissertation aimed to investigate the role of children's prosocial motivation in the ontogeny of morality. There is both theoretical reason and empirical evidence to propose that altruism and empathic concern are decisive motivational factors of morality, which the moral sentimentalists David Hume and Adam Smith have described as sympathy for the other in need, not only in adulthood, but also in the development of morality (Batson, 2010; Batson & Shaw, 1991; Jensen et al., 2014; Nichols, 2002, 2004; Roughley & Bayertz, 2019). However, past developmental research has mostly focused separately on the pre-moral development of children's first-party prosociality and their third-party understanding of moral norms (Tomasello, 2016, 2018). Accordingly, children in their second year of life develop empathic concern and sympathy for others in need in prosocial situations (Hepach, 2017; Hepach et al., 2012). Furthermore, recent findings suggest that 18-month-old children already show some rudimentary forms of norm understanding at least in dyadic conventional situations. This rudimentary norm understanding is interpreted as second-personal normative expectations (Schmidt et al., 2019). Moreover, 3-year-old children not only have descriptive expectations about morality, but also normative ones as suggested by their enforcement of moral norms as unaffected bystanders (Rakoczy et al., 2016; Rossano et al., 2011; Vaish et al., 2011). The relation between prosociality and morality, however, in particular, the prosocial motivational source of the early sense of morality, remains unclear. Thus in three studies, I addressed this topic by investigating the developmental origins of morality (Study 1), the underlying prosocial motivation for children's normative appreciation of morality (Study set 2), and by examining the scope of morality (Study 3). First, I will briefly summarize the main findings of the three empirical studies presented in the current thesis. Thereafter, a more general discussion of the results and their

contributions to the three stated questions will be outlined. After highlighting the theoretical impact of my work, some limitations and propositions for future research will be described and possible implications will be briefly touched upon before drawing some final conclusions.

## **5.1. Summary**

Study 1 investigated the developmental origins of morality in 18-month-old infants. In an eye-tracking paradigm, anticipatory looking and pupil dilation were used to examine whether infants differentiate between prototypical moral (harmful) and conventional (harmless) violations. In a between-subjects-design, children watched the very same video clip whose audio stream differed according to condition. In the first two conditions, an instructor told an observer to destroy a picture with tool A (conventional violation) or with tool B (no violation), whereas in the third (moral violation) condition, the instructor forbade the observer to destroy the picture at all. In all three conditions, the observer then grasped tool B and destroyed the picture, which led to the three different violation situations. Results demonstrated that infants showed a larger relative increase in pupil dilation in response to a moral violation than to a conventional violation. Moreover, infants differentiated between two types of conventional norm situations in their anticipatory looking based on prescribed actions.

These findings suggest that 18-month-old infants have third-party descriptive expectations about the distinction between conventional and moral violation situations. Moreover, they provide the first evidence that empathic concern may be a decisive capacity for the distinction between these two violation situations.

Study set 2 looked at the underlying prosocial motivation for the normative appreciation of morality in 3-year-old children. In three experiments, children were given a third-party fairness task (which varied across the experiments) and two different

prosocial tasks. A spontaneous protest paradigm was used to investigate, whether children enforce fairness norms (for a conceptual overview, see Schmidt & Rakoczy, 2018, 2019). In Experiment 1, children protested and corrected unequal (but not equal) allocations, suggesting a normative understanding of third-party fairness. Experiment 2 assessed whether children's normative expectations about fairness have a moral (authority-independent) dimension. Thus, children observed a distributor who followed (unequal condition) or violated (equal condition) an authority's command to allocate resources unequally. Again, despite the authority's dictate to act unequally, children protested more against unequal versus equal allocations. In Experiment 3, results show that children enforced fairness norms by altruistic punishment. While in Experiments 1 and 2 there was a positive relation of protest behavior and emotional sharing (empathic concern), in Experiment 3 children's third-party altruistic punishment was associated with their own costly sharing behavior (altruism).

Overall, results of Study set 2 provide evidence that young children's third-party normative expectations about fairness are closely related to their own prosocial behavior. More generally, this work may help better understand to what extent the ontogeny of fairness norms can be characterized as moral in that it is associated with children's developing concern for the welfare of others in different contexts.

In Study 3, I explored the scope of morality in 5- to 6-year-olds in a typical intergroup context. Here, I investigated whether decategorization – a candidate mechanism to overcome in-group bias by emphasizing the individual person – would lead preschoolers to treat in-group and out-group members equally when sharing resources in a dictator game. I found that preschoolers shared more resources with an in-group than with an out-group recipient when social category membership was emphasized. When individuating information was emphasized (decategorization), however, children shared the same with in-group and out-group individuals. These

findings suggest that the presentation of out-group members as individuals may be a powerful tool to reduce in-group bias and to foster equal treatment (an important moral category) of in-group and out-group members in preschool children.

## **5.2. Developmental Origins of Morality**

The main finding of the first study was that a distinction between situations that are considered prototypical moral and conventional violations begins much earlier in the development than previously thought. More specifically, results revealed that the ability to feel sympathy may be critical for the development of the moral-conventional distinction and that 18-month-old infants, at minimum, have third-party descriptive expectations about that distinction.

Using anticipatory looking, results show that infants distinguished between two types of conventional situations within a dyadic interaction (between an instructor and an observer). More precisely, these situations consisted of two mutually exclusive game rules that were explained by the instructor, namely that the observer should tear apart a picture with a certain tool, while not using the other tool, which – later in context – leads to a no violation or a violation situation. As predicted, in each conventional situation, the children expected the observer to use the prescribed tool. This finding suggests that infants possess a third-party descriptive understanding of conventional rule agreements. Furthermore, infants showed a larger relative increase in pupil dilation in response to a moral violation situation than in response to a conventional violation situation. No difference was found between a conventional violation and a no violation situation. These results demonstrate that children were more physiologically aroused when witnessing a moral violation, which is interpreted as affective arousal due to the severity of the moral violation (empathic concern, sympathy).

It is important to note that there may be other causes of pupil dilation correlating with physiological arousal (Eisenberg & Fabes, 1990; Pletti et al., 2017). To exclude such alternative interpretations, I sought to control for the various sensory and cognitive efforts. First, to control for luminance, the same video material across all conditions was used. Second, cognitive efforts caused, for instance, by different syntax structures were controlled for by using the same sentence structures with the same number of syllables, speech volume and timing across all three conditions. Furthermore, on a semantic level, there is no a priori reason to assume that the moral violation should be cognitively more complex than the other two conditions. On the contrary, one could argue that the moral violation situation should be easier to process than the other two situations (no violation and conventional violation), since simply the use of both tools is prohibited and no action should be performed at all. Another indication is that no differences were found in pupil dilation between conventional and no violation situations. This speaks against pupil size differences being due to cognitive effort, as pupil size would have varied based on the increased cognitive effort needed to process an unexpected as opposed to an expected action. However, that was not the case. This and the fact that the children had increased physiological arousal only in the moral violation condition seems good evidence for increased affective arousal and not for cognitive effort.

There are several reasons these findings contribute to the developmental origins of morality. First, the findings of the eye-tracking paradigm extend previous language-driven work in interview studies as well as behavioral studies, which suggest that children at the age of 3 reliably distinguish between non-arbitrary moral norms (e.g., against harming others) and more arbitrary conventional norms (e.g., game rules, dress code; e.g., Schmidt, Rakoczy, & Tomasello, 2012; Smetana et al., 2018; Smetana & Braeges, 1990). In line with social domain theory (Smetana & Braeges, 1990; Turiel,



1983, 2006), the results show that even 18-month-old infants were capable of distinguishing between moral and conventional violation situations.

Second, the findings are consistent with psychological, physiological and philosophical explanations of morality (Killen & de Waal, 2000; Killen & Smetana, 2015; Nichols, 2002; Smetana, 1984; Smetana et al., 2014; Tomasello, 2016; Turiel, 1983, 2006). Violations of moral norms are considered more serious and punishable than conventional norm violations because they take the welfare of others into account. A plethora of studies have shown such interrelations between morality and other-regarding concerns (empathic concern, sympathy) in children and adults (Decety et al., 2012; Decety & Yoder, 2016; Eisenberg et al., 2014; Hepach, 2017; Hepach et al., 2012; Jensen et al., 2014; Lamm et al., 2019; Nichols, 2002; Roth-Hanania et al., 2011; Svetlova et al., 2010; Vaish et al., 2009). Moreover, a large number of experimental studies have shown that such affective arousal is associated with an empathic concern for the victim and leads to an altruistic motivation to show prosocial behavior (Batson & Shaw, 1991; Eisenberg et al., 1990; Eisenberg & Fabes, 1990; Hepach, 2017; Hepach et al., 2017; Jensen et al., 2014; Nichols, 2002, 2004; Vaish et al., 2009).

Third, the current results complement the findings of a recent study that revealed that 18-month-old infants have an early rudimentary form of norm understanding at least in dyadic interactions, specifically of how "we" (a dyadic of "you" and "me") ought to do something, which is called second-personal normativity (Schmidt et al., 2019). There is further evidence that 18-month-old children are not fully capable of applying and regulating the self-another-distinction, in particular on the cognitive level, which is associated with the still undeveloped mentalizing, executive and inhibitory functions (Decety & Sommerville, 2003; Milward & Sebanz, 2016; Steinbeis, 2016).

Taken together, in accordance with the above-mentioned numerous theoretical reasons and empirical evidence on the important role of prosocial motivation in

understanding morality, Study 1 provides further information on Tomasello's model (Tomasello, 2016). In other words, not only might prosociality be a crucial precursor for the development of morality, but it also appears to be an important building block that is more closely related to the (at least descriptive) understanding of moral situations than previously thought. Thus, the ability to feel sympathy may be critical for the development of the moral-conventional distinction and 18-month-old infants, at minimum, have third-party descriptive expectations about that distinction. These findings can be interpreted as a precursor of morality rather than a full-fledged, agent-independent, group-minded social norm understanding, which develops at the age of three (Tomasello, 2018).

### **5.3. The Relationship between Prosociality and Morality**

The main result of the second study set was that 3-year-old children understand fairness as a moral normative notion. Moreover, a 3-year-old's sense of fairness is related to prosocial tendencies and other-regard, especially sympathy and altruism – that is, being interested in and concerned about others' well-being and motivated to act accordingly.

In three experiments, results revealed that children enforce third-party fairness norms as unaffected bystanders, even without receiving instructions of acting fairly, that they protest and intervene against injustice even when an authority dictates inequality, and finally that they are even willing to sacrifice their own resources in order to punish unfair behavior. Moreover, children's third-party normative expectations of fairness are closely linked to their own prosocial behavior, particularly to their concern for the welfare of others (empathic concern, sympathy) and to an altruistic interest in punishing norm violators.

Taken together, Study set 2 specifically investigated the mechanisms of the developmental understanding of moral norms, in particular the relationship between children's first-party prosociality and their third-party normative expectations about fairness.

There are several reasons these findings contribute to understanding the relationship between prosociality and morality. First, only a few studies have investigated the interrelation between (pre)moral capacities and early prosocial motivations (Schmidt & Sommerville, 2011; Sommerville et al., 2013).

Second, the developmental findings of this thesis contribute to the philosophical theory of moral sentimentalism (David Hume and Adam Smith) as well as to Batson's psychological model of altruistic motivation (Batson, 2010; Batson & Shaw, 1991; Nichols, 2004; Waldmann et al., 2012). Thus, theoretically, the current findings complement Tomasello's model, in which self-other equivalence is assumed to be the main driving force of morality and, in particular, of fairness considerations, rather than altruism and empathic concern for the welfare of others. Moreover, Tomasello regarded children's prosocial altruistic motivation (empathic concern, sympathy) as a pre-moral first step to the later agent-neutral normative understanding of morality (Tomasello, 2016). Thus, in his point of view, both constructs are seen as more separated than interrelated. Study set 2 therefore provides further information for Tomasello's model, since prosociality and in particular empathic concern and altruism has proven to be an important mechanism even in the case of the development of an agent-neutral norm understanding of morality.

Third, the current results of Study Set 2 support the social domain theoretical approach (Smetana & Braeges, 1990; Turiel, 1983, 2008) by showing that not only adults but also children consider fairness a moral norm by 3 years of age, at least along the criterion of "authority- contingency ". Moreover, since these previous results stem

mainly from hypothetical interview scenarios, the current results give further evidence by using a third-party protest paradigm. Thus, 3-year-old children showed not only, that they are able to adequately adhere to moral norms, but also show third-party norm enforcements, which is considered a key normative understanding of morality, because it showed children understanding the normative force (the “oughtness”) in an agent-neutral way (Chudek & Henrich, 2011; Fehr & Fischbacher, 2004). Furthermore, the current work extends previous behavioral findings, which suggested that preschool children have descriptive, but not normative, expectations about equal resource allocation (Geraci & Surian, 2011; Meristo et al., 2016). Therefore, the findings underpin important findings of previous behavioral (protest) studies that show that 3-year-old children are generally capable of intervening and protesting against norm violations, for instance conventional game rule violations (Rakoczy et al., 2008) and moral transgressions (Rakoczy et al., 2016; Rossano et al., 2011; Vaish et al., 2011). In addition, the results are consistent with previous findings showing third-party costly punishment behavior in preschoolers following moral violations (Kenward & Öst, 2012; Krasnow et al., 2016; McAuliffe et al., 2015; Robbins & Rochat, 2011).

Overall, the current findings of Study set 2 may help better understand to what extent the ontogeny of fairness norms can be characterized as moral, in that it is associated with children’s developing concern for the welfare of others in different contexts.

#### **5.4. Scope of Morality**

The main finding of the third study was that 5-to 6-year-old children treat in-group and out-group members equally in a moral normative sense of fairness. Moreover, the present study demonstrated the effectiveness of decategorization to reduce intergroup bias in preschool children. First, I showed that in a typical intergroup context, children

favor their own group over the out-group by sharing more stickers with in-group versus out-group members. Second, when presenting group members in a decategorized manner, that is when the recipients were presented as individuals with certain preferences for toys and sweets, intergroup bias was diminished. In this case, children shared the same number of stickers with in-group and out-group members.

There are several reasons these findings contribute to the scope of morality. First, the results showed that preschool children at the ages of 5 and 6 are already susceptible to intergroup distinction. This complements earlier results regarding intergroup bias in early childhood (Benozio & Diesendruck, 2015; Chalik & Dunham, 2018; Dunham et al., 2011; Misch et al., 2016; Over, 2018; Rutland et al., 2010, 2015; Sierksma et al., 2018).

Second, the current findings provide further evidence for approaches, which stated that the motivational and normative forces of mutual altruism and cooperation are limited in scope to their in-group (Brewer & Caporael, 2006; Fiske, 2000; Tomasello, 2016). That is, people need obligatory interdependence within a certain small group to live together peacefully and in order to survive in competition with physiologically dominant animals or other rival groups. Thus, a relatively small, distinct in-group most effectively provide loyalty, trust and cooperation, where the chance of being exploited is relatively low (Tomasello, 2016). In line with that, the identification with the relevant in-group is at the same time coupled with a social comparison with other opposing out-groups, to which one automatically does not belong. This typically leads to social categorisation and intergroup distinction, accompanied by positive emotional attachment and consensus among members of the same group (Tajfel, 1982; Tajfel & Turner, 1979). My findings complements, that social categorization processes even under trivial criteria (e.g., group color yellow vs. purple) are sufficient enough to trigger in-group/out-group distinction (Tajfel, Billig, Bundy, & Flament, 1971) in preschoolers.

Third, the results of Study 3 are consistent with those from social psychology as well as with research of intervention methods in adults and school-children, indicating that decategorization is a suitable method to restore equal treatment of both in-group and out-group members (Bettencourt et al., 1992; Brewer, 1996, 1997; Cameron et al., 2006; Cameron & Rutland, 2006). This extends the scope of morality, in that not only in-group members as opposed to out-group members are more preferred (in-group bias), but decategorization leads to overcome this bias by showing equal treatment of both group members. Thus, the development of equality in intergroup contexts is not only an important capacity for a full-blown understanding of fairness (e.g., Rawls, 1999) but also important to overcome discrimination and prejudice, even at preschool years (Beelmann & Heinemann, 2014).

Overall, the findings suggest that even preschool children at the age of 5- to 6 are not only vulnerable to intergroup distinction, but I could also show results to overcome this bias, which leads to equal treatment of in-group and out-group members.

## **5.5. Limitations and Implications for Future Research**

The present thesis offered new insights into the prosocial roots of children's developing morality. However, further research is needed to unravel the phenomenon of young children's prosocial motivation to understand, adhere to, and enforce moral norms. Therefore, in the following I will discuss the limitations of the results of each study, raise some future research questions and point out implications.

Study 1 examined whether 18-month-old children distinguish between conventional and moral-norm violations. However, further research is needed to disentangle the phenomenon of infants' capacity to distinguish between moral and conventional situations. I provided first evidence using a novel eye-tracking paradigm, but anticipatory gaze behavior and pupil dilation are indirect measures of children's

capacity to distinguish between conventional and moral violations. Future research should use a combination of eye-tracking measures and behavioral tasks to investigate the relation between children's first-party behavior (empathic concern, altruism) and third-party descriptive expectations of morally relevant situations. In addition, the present paradigm was limited to certain norm violations, such as property rights. Examining different moral violations (e.g., fairness, other harm-related situations) would complement the current findings. Furthermore, there is much evidence that witnessing social interactions, in particular between the victim and the perpetrator, is a crucial factor in investigating children's understanding of morally relevant situations; in particular being concerned for the victim and/or being more emotionally aroused because of the desire to punish the perpetrator (e.g., Hepach, 2017; Hepach et al., 2017; Vaish et al., 2009). However, the first study focusses on the tool use. An interesting question for future research would be whether children exhibit greater pupil dilation when looking at the victim (empathic concern) or at the perpetrator (desire to punish). A further interesting aspect to explore would be the influence of individual differences on the relationship between children's prosocial arousal (empathic concern for the victim) and their descriptive expectations, when witnessing different moral situations. Finally, a gaze-contingency paradigm could shed further light on whether infants are both, able to possess third-party descriptive expectations, and to punish the perpetrator through gaze control (norm enforcement).

Study 2 demonstrated that 3-year-old children understand fairness as a moral normative notion. Moreover, clear evidence about the relation between children's prosocial motivation (empathic concern, altruism) and their third-party moral norm enforcement was found. Thus, these results are specific for fairness norms, but it would be interesting to investigate other types of moral norms, for instance harm-related issues, as well. Moreover, results revealed that 3-year-old children show a normative

understanding of fairness at least at a specific criterion (authority independency) of Turiel's (1983, 2006) theoretical account, but other criteria such as generalizability are also necessary to explore. As a future direction, normative expectations about fairness as well as their relationship to children's prosocial motivation could be examined in an intergroup context. Moreover, it would be interesting to explore the relation between children's prosocial motivation and other fairness considerations such as merit or need. Finally, a closer look at the relation between other capacities, such as intelligence, executive functions, inhibition or theory of mind, and children's moral norm understanding would be interesting to investigate.

Results from Experiment 3 of Study set 2 are ambiguous and give rise to further possible investigation. On the one hand, the explicit question (i.g., "Is the distribution good or bad?") was useful to measure whether the children understand fairness correctly as sharing resources equally. However, we could not investigate what the children who judged unfair distribution as good would do in the costly sharing task, because we did not give them the opportunity to perform this task. Moreover, it is not yet clear what the children's motivation was, who answered the explicit question not fully wrong but mixed (i.g., an unequal distribution was judged as good and in other trials was judged as bad). It might be that they played strategically based on egotistical motivation, but possess normative fairness considerations, and simply were not willing to sacrifice own resources to foster equal treatment. Furthermore, we conducted a paradigm which could not fully separate retributive (i.e., taking resources away to punish the perpetrator) from restorative justice considerations (i.e. take it from the perpetrator and give it to the victim back to restore justice as a moral norm). It might be interesting to examine children's behavior in a pure restorative justice situation. A future study could examine how children understand the sense of justice at no personal cost by allowing them to take resources away from the perpetrator to give to the victim. It would also be



interesting to investigate whether younger children also understand restorative justice. For example, an eye-tracking paradigm could be used here as a language-free paradigm, and specifically a gaze-contingency paradigm. Children could then interactively redistribute resources.

Study 3 revealed that decategorization is a suitable intervention method to foster equal treatment of in-group and out-group members. In this context, the question arises how children would react to a third-person distribution task, in which they distribute stickers directly between in-group and out-group members at no personal cost. Moreover, interesting descriptive findings were found regarding the intergroup bias and the direction in which it was removed. Following the distribution pattern in all four conditions, it seems that children distributed more stickers to the recipient of their in-group in the impersonal categorization condition than in the other three remaining conditions. Hence, it could be that children at this age show a decrease in prosociality towards in-group members but not an increase in prosocial behavior towards the out-group members. Future research could investigate whether children show an in-group bias but not yet active forms of out-group derogations. Study 3 revealed the scope of morality in an intergroup context, but what about the relationship between children's prosocial motivation and their moral norm understanding in an intergroup context? It might be that children not only show an intergroup bias in a minimal group paradigm, but also process differences in their prosocial motivation (less empathic concern for an out-group victim/more concern for in-group victim). And furthermore, the question arises whether children are equally concerned when using an intervention method such as decategorization. Results revealed significant effects on specific in-group as well as out-group members, even when drawing them anonymously. However, it might be that the other group members are still homogenized (i.e., I like Nisa from the purple group, but I still hate the purple group). However, it is not yet clear whether the findings that

are related to specific individual group members can be extended to the whole group or to new in-group/ out-group members. In the future, a generalization effect should therefore be investigated. A further limitation of the study is a transfer of our findings to real life situations. First, decategorization is a suitable measure as a laboratory experiment, but further research is needed in the field. The experiment was based on arbitrary and artificial categories, but we do not yet know how it is with deeply rooted prejudices, e.g. towards minorities or whether effects are short or long lasting.

Finally, what implications can be derived from the results of the present work? In Study 1 and Study set 2, the importance of the relationship between children's altruistic motivation (empathic concern, sympathy) and their understanding of morally relevant situations was found, even at an early preschool age. The question therefore arises as to what decisive role empathic concern might play for the parenting style at home, the educational style in kindergarten or at school. Since children also seem to have an intrinsic altruistic motivation, it might be useful not to undermine their motivation. In the third study, the effects of inequality in an intergroup context were examined. I demonstrated how equality can be restored through appropriate intervention methods such as decategorization. Thus, suitable intervention methods in school as well as in kindergarten are of crucial importance to help children develop into moral autonomous beings without showing discrimination and resentments against others who not belong to "us".

## 5.6. Conclusion

The current dissertation provides a novel overview of the prosocial roots of children's developing morality. In particular, the present findings suggest that the ability to feel sympathy helps 18-month-old infants distinguish moral from conventional norm violations, at least having third-party descriptive expectations. Moreover, 3-year-old children's concern for the welfare of others in different contexts of fairness revealed that the ontogeny of fairness norms can be characterized as moral. Finally, the presentation of out-group members as individuals may be a powerful tool to reduce in-group bias and foster equal treatment in a moral sense of in-group and out-group members in 5- to 6-year-old preschool children. However, further research is certainly needed to unravel the phenomenon of young children's prosocial motivation to understand, adhere to, and enforce moral norms.

Taken together, the current dissertation adds to the literature on the ontogeny of moral norm-psychology. The present thesis has revealed that prosociality and morality are not only separate developmental steps in the ontogeny of morality, but that both constructs are more closely related than previously thought. Moreover, prosocial (altruistic) motivation, especially empathic concern and sympathy, plays an important role in children's developing into full-blown moral agents.

To return to my story from the beginning of the thesis. What can I conclude from my experience with the woman who helped me and from the insights I gained from my work? The prosocial motivation behind moral behavior is an important aspect in determining whether someone acts out of moral (good and right) reasons, not only in adulthood but even in infancy.

## References

- Abrams, D. (2010). *Processes of prejudice: Theory, evidence and intervention*. Equality and Human Rights Commission.
- Aknin, L. B., Broesch, T., Hamlin, J. K., & Van de Vondervoort, J. W. (2015). Prosocial behavior leads to happiness in a small-scale rural society. *Journal of Experimental Psychology: General*, *144*(4), 788–795.  
<https://doi.org/10.1037/xge0000082>
- Alexander, L., & Moore, M. (2016). Deontological ethics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.  
<https://plato.stanford.edu/archives/win2016/entries/ethics-deontological/>
- Amsterdam, B. (1972). Mirror self-image reactions before age two. *Developmental Psychobiology*, *5*(4), 297–305.
- Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, *100*(401), 464–477.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, *55*(4), 1278–1289.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *Fitting Linear Mixed-Effects Models using lme4*.
- Batson, C. D. (2010). Empathy-induced altruistic motivation. In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial motives, emotions, and behavior: The better angels of*

*our nature*. (pp. 15–34). American Psychological Association.

<https://doi.org/10.1037/12061-001>

- Batson, C. D., & Shaw, L. L. (1991). Evidence for altruism: Toward a pluralism of prosocial motives. *Psychological Inquiry*, 2(2), 107–122.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3), 497–529.
- Beelmann, A., & Heinemann, K. S. (2014). Preventing prejudice and improving intergroup attitudes: A meta-analysis of child and adolescent training programs. *Journal of Applied Developmental Psychology*, 35(1), 10–24.  
<https://doi.org/10.1016/j.appdev.2013.11.002>
- Benenson, J. F., Pascoe, J., & Radmore, N. (2007). Children's altruistic behavior in the dictator game. *Evolution and Human Behavior*, 28(3), 168–175.  
<https://doi.org/10.1016/j.evolhumbehav.2006.10.003>
- Benozio, A., & Diesendruck, G. (2015). Parochialism in preschool boys' resource allocation. *Evolution and Human Behavior*, 36(4), 256–264.  
<https://doi.org/10.1016/j.evolhumbehav.2014.12.002>
- Bettencourt, A. B., Brewer, M. B., Croak, M. R., & Miller, N. (1992). Cooperation and the reduction of intergroup bias: The role of reward structure and social orientation. *Journal of Experimental Social Psychology*, 28, 301–319.
- Bischof-Köhler, D. (1993). *Spiegelbild und Empathie: Die Anfänge der sozialen Kognition* (Nachdr). Huber.
- Blake, P. R., & Rand, D. G. (2010). Currency value moderates equity preference among young children. *Evolution and Human Behavior*, 31(3), 210–218.  
<https://doi.org/10.1016/j.evolhumbehav.2009.06.012>

- Brandom, R. (1994). *Making it explicit: Reasoning, representing, and discursive commitment*. Harvard University Press.
- Brewer, M. B. (1996). When contact is not enough: Social identity and intergroup cooperation. *International Journal of Intercultural Relations*, 20(3–4), 291–303.
- Brewer, M. B. (1997). The social psychology of intergroup relations: Can research inform practice? *Journal of Social Issues*, 53(1), 197–211.
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues*, 55(3), 429–444.
- Brewer, M. B., & Caporael, L. R. (2006). An evolutionary perspective on social identity: Revisiting groups. In M. Schaller, J. A. Simpson, & D. Kendrick (Eds.), *Evolution and Social Psychology* (pp. 143–162). Psychological Press.
- Brewer, M. B., & Miller, N. (1984). Beyond the contact hypothesis: Theoretical perspectives on desegregation. In N. Miller & M. B. Brewer (Eds.), *Groups in contact. The psychology of desegregation*. Academic Press.
- Buss, D. M. (2016). The handbook of evolutionary psychology. In V. Weekes-Shackelford, T. K. Shackelford, & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science* (pp. 1–3). Springer International Publishing. [https://doi.org/10.1007/978-3-319-16999-6\\_1860-1](https://doi.org/10.1007/978-3-319-16999-6_1860-1)
- Cameron, L., & Rutland, A. (2006). Extended contact through story reading in school: Reducing children's prejudice toward the disabled. *Journal of Social Issues*, 62(3), 469–488.
- Cameron, L., Rutland, A., Brown, R., & Douch, R. (2006). Changing children's intergroup attitudes toward refugees: Testing different models of extended contact. *Child Development*, 77(5), 1208–1219.
- Cameron, L., Rutland, A., Hossain, R., & Petley, R. (2011). When and why does extended contact work? The role of high quality direct contact and group norms

- in the development of positive ethnic intergroup attitudes amongst children. *Group Processes & Intergroup Relations*, 14(2), 193–206.  
<https://doi.org/10.1177/1368430210390535>
- Chalik, L., & Dunham, Y. (2018). Beliefs about moral obligation structure children's social category-based expectations. *Child Development*, 1–12.  
<https://doi.org/10.1111/cdev.13165>
- Christen, M., & Glock, H.-J. (2012). The (limited) space for justice in social animals. *Social Justice Research*, 25(3), 298–326. <https://doi.org/10.1007/s11211-012-0163-x>
- Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences*, 15(5), 218–226.  
<https://doi.org/10.1016/j.tics.2011.03.003>
- Cialdini, R. B., Schaller, M., Houlihan, D., Arps, K., Fultz, J., & Beaman, A. L. (1987). Empathy-based helping: Is it selflessly or selfishly motivated? *Journal of Personality and Social Psychology*, 52(4), 749–758.
- Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, 1(3), 98–101.
- Crisp, R. J., & Turner, R. N. (2009). Can imagined interactions produce positive perceptions? Reducing prejudice through simulated social contact. *American Psychologist*, 64(4), 1–28.
- Dahl, A. (2015). The developing social context of infant helping in two U.S. samples. *Child Development*, 86(4), 1080–1093.
- Darwin, C. (1871). *The descent of man*. BoD–Books on Demand.
- Davidov, M., Zahn-Waxler, C., Roth-Hanania, R., & Knafo, A. (2013). Concern for others in the first year of life: Theory, evidence, and avenues for research. *Child Development Perspectives*, 7(2), 126–131. <https://doi.org/10.1111/cdep.12028>

- Dawes, R. M., Van De Kragt, A. J., & Orbell, J. M. (1988). Not me or thee but we: The importance of group identity in eliciting cooperation in dilemma situations: Experimental manipulations. *Acta Psychologica*, *68*(1–3), 83–97.
- Dawkins, R. (1976). *The selfish gene*. Oxford University Press.
- de Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology*, *59*(1), 279–300.  
<https://doi.org/10.1146/annurev.psych.59.103006.093625>
- Decety, J., & Howard, L. H. (2013). The role of affect in the neurodevelopment of morality. *Child Development Perspectives*, *7*(1), 49–54.  
<https://doi.org/10.1111/cdep.12020>
- Decety, J., Michalska, K. J., & Kinzler, K. D. (2012). The contribution of emotion and cognition to moral sensitivity: A neurodevelopmental study. *Cerebral Cortex*, *22*(1), 209–220. <https://doi.org/10.1093/cercor/bhr111>
- Decety, J., & Sommerville, J. A. (2003). Shared representations between self and other: A social cognitive neuroscience view. *Trends in Cognitive Sciences*, *7*(12), 527–533. <https://doi.org/10.1016/j.tics.2003.10.004>
- Decety, J., & Yoder, K. J. (2016). Empathy and motivation for justice: Cognitive empathy and concern, but not emotional empathy, predict sensitivity to injustice for others. *Social Neuroscience*, *11*(1), 1–14.  
<https://doi.org/10.1080/17470919.2015.1029593>
- Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology/Psychologie Canadienne*, *49*(3), 182–185. <https://doi.org/10.1037/a0012801>
- Dobson, A. J. (2002). *An introduction to generalized linear models* (2nd ed.). Chapman & Hall/CRC.



- Dondi, M., Simion, F., & Caltran, G. (1999). Can newborns discriminate between their own cry and thy cry of another newborn infant? *Developmental Psychology*, 35(2), 418–426.
- Dunham, Y., Baron, A. S., & Banaji, M. R. (2008). The development of implicit intergroup cognition. *Trends in Cognitive Sciences*, 12(7), 248–253.  
<https://doi.org/10.1016/j.tics.2008.04.006>
- Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of “minimal” group affiliations in children. *Child Development*, 82(3), 793–811.  
<https://doi.org/10.1111/j.1467-8624.2011.01577.x>
- Eisenberg, N., Eggum Wilkens, N. D., & Spinrad, T. L. (2015). The development of prosocial behavior. In A. Schroeder & W. G. Graziano (Eds.), *The Oxford handbook of prosocial behavior* (pp. 114–136). Oxford University Press.
- Eisenberg, N., & Fabes, R. A. (1990). Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion*, 14(2), 131–149.
- Eisenberg, N., Fabes, R. A., Miller, P. A., Shell, R., Shea, C., & May-Plumlee, T. (1990). Preschoolers’ vicarious emotional responding and their situational and dispositional prosocial behavior. *Merrill-Palmer Quarterly (1982-)*, 507–529.
- Eisenberg, N., & Mussen, P. (1989). *The roots of prosocial behavior in children*. Cambridge University Press.
- Eisenberg, N., Spinrad, T. L., & Morris, A. (2014). Empathy-related responding in children. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (2nd ed., pp. 184–207). Psychology Press.
- Engel, P. (2011). Epistemic norms. In S. Bernecker & D. Pritchard (Eds.), *The Routledge Companion to Epistemology* (pp. 47–57). Routledge.

- Ensari, N., & Miller, N. (2001). Decategorization and the reduction of bias in the crossed categorization paradigm. *European Journal of Social Psychology, 31*, 193–216.
- Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature, 454*(7208), 1079–1083. <https://doi.org/10.1038/nature07155>
- Fehr, E., & Fischbacher, U. (2004). Social norms and human cooperation. *Trends in Cognitive Sciences, 8*(4), 185–190. <https://doi.org/10.1016/j.tics.2004.02.007>
- Feinberg, J. (2017). Noncomparative Justice 1. In *Justice* (pp. 41–82). Routledge.
- Fiske, S. T. (2000). Stereotyping, prejudice, and discrimination at the seam between the centuries: Evolution, culture, mind, and brain. *European Journal of Social Psychology, 30*, 299–322.
- Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology, 65*(1), 47–55.
- Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A., & Rust, M. C. (1993). The common ingroup identity model: Recategorization and the reduction of intergroup bias. *European Review of Social Psychology, 4*(1), 1–26. <https://doi.org/10.1080/14792779343000004>
- Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources: Evaluation of fairness of distributive actions. *Developmental Science, 14*(5), 1012–1020. <https://doi.org/10.1111/j.1467-7687.2011.01048.x>
- Göckeritz, S., Schmidt, M. F. H., & Tomasello, M. (2014). Young children's creation and transmission of social norms. *Cognitive Development, 30*, 81–95. <https://doi.org/10.1016/j.cogdev.2014.01.003>

- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*(5), 1029–1046. <https://doi.org/10.1037/a0015141>
- Greene, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*(5537), 2105–2108. <https://doi.org/10.1126/science.1062872>
- Grossmann, T., & Johnson, M. H. (2007). The development of the social brain in human infancy: The social brain in infancy. *European Journal of Neuroscience, 25*(4), 909–919. <https://doi.org/10.1111/j.1460-9568.2007.05379.x>
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814–834.
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus, 133*(4), 55–66. <https://doi.org/10.1162/0011526042365555>
- Haidt, J., Roller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology, 65*(4), 613–628.
- Hamlin, J. K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences, 108*(50), 19931–19936. <https://doi.org/10.1073/pnas.1110306108>
- Hardecker, S., & Tomasello, M. (2017). From imitation to implementation: How two- and three-year-old children learn to enforce social norms. *British Journal of Developmental Psychology, 35*(2), 237–248. <https://doi.org/10.1111/bjdp.12159>
- Haslam, A. S. (2004). The social identity approach. In A. S. Haslam (Ed.), *Psychology in organizations*. (2nd ed., pp. 17–39). Sage.

- Hastings, P. D., Miller, J. G., Kahle, S., & Zahn-Waxler, C. (2014). The neurobiological bases of empathic concern for others. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (2nd ed., pp. 411–434). Psychology Press.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion. Studies in emotion and social Interaction*. Cambridge University Press.
- Hechter, M., & Opp, K. D. (Eds.). (2001). *Social norms*. Russell Sage Foundation.
- Heffner, J., & FeldmanHall, O. (2019). Why we don't always punish: Preferences for non-punitive responses to moral violations. *Scientific Reports*, 9(1), 2045–2322. <https://doi.org/10.1038/s41598-019-49680-2>
- Hepach, R. (2017). Prosocial arousal in children. *Child Development Perspectives*, 11(1), 50–55. <https://doi.org/10.1111/cdep.12209>
- Hepach, R., Vaish, A., Müller, K., & Tomasello, M. (2017). The relation between young children's physiological arousal and their motivation to help others. *Neuropsychologia*. <https://doi.org/10.1016/j.neuropsychologia.2017.10.010>
- Hepach, R., Vaish, A., & Tomasello, M. (2012). Young children are intrinsically motivated to see others helped. *Psychological Science*, 23(9), 967–972. <https://doi.org/10.1177/0956797612440571>
- Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, 53, 575–604.
- Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. Cambridge University Press.
- Ickes, W. (1993). Empathic accuracy. *Journal of Personality*, 61(4), 587–610. <https://doi.org/10.1111/j.1467-6494.1993.tb00783.x>
- Jensen, K., Vaish, A., & Schmidt, M. F. H. (2014). The emergence of human prosociality: Aligning with others through feelings, concerns, and norms. *Frontiers in Psychology*, 5, 1–16. <https://doi.org/10.3389/fpsyg.2014.00822>

- Kahnemann, D., & Beatty, J. (1967). Pupillary responses in a pitch-discrimination task. *Perception & Psychophysics*, 2(3), 101–105.  
<https://doi.org/10.3758/BF03210302>
- Kane, J. (1996). Justice, impartiality, and equality: Why the concept of justice does not presume equality. *Political Theory*, 24(3), 375–393.
- Kant, I. (1959). *Foundations of the metaphysics of morals*. SophiaOmni Foundation.
- Kaufmann, A. S. (2015). *Kaufmann assessment battery for children (KABC-II). Deutsche Version* (2nd ed.). Melchers, U./Preuß, U.
- Kauppinen, A. (2018). Moral sentimentalism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.  
<https://plato.stanford.edu/archives/win2018/entries/moral-sentimentalism/>
- Kelly, H. H., & Thibaut, J. W. (1979). Interpersonal relations-a theory of interdependence. *Applied Ergonomics*, 10(4), 249.
- Kenward, B., & Östh, T. (2012). Enactment of third-party punishment by 4-year-olds. *Frontiers in Psychology*, 3, 1–9. <https://doi.org/10.3389/fpsyg.2012.00373>
- Killen, M., & de Waal, F. B. M. (2000). The evolution and development of morality. In F. Aureli & F. B. M. de Waal (Eds.), *Natural conflict resolution* (pp. 352–372). University of California Press.
- Killen, M., & Smetana, J. G. (2015). Origins and development of morality. In R. M. Lerner (Ed.), *Handbook of Child Psychology and Developmental Science* (pp. 1–49). John Wiley & Sons, Inc.  
<https://doi.org/10.1002/9781118963418.childpsy317>
- Knafo, A., Zahn-Waxler, C., Van Hulle, C., Robinson, J. L., & Rhee, S. H. (2008). The developmental origins of a disposition toward empathy: Genetic and environmental contributions. *Emotion*, 8(6), 737–752.  
<https://doi.org/10.1037/a0014179>

- Korsgaard, C. M. (1996). *The sources of normativity*. Cambridge University Press.
- Krasnow, M. M., Delton, A. W., Cosmides, L., & Tooby, J. (2016). Looking under the hood of third-party punishment reveals design for personal benefit. *Psychological Science, 27*(3), 405–418.  
<https://doi.org/10.1177/0956797615624469>
- Lamm, C., Rütgen, M., & Wagner, I. C. (2019). Imaging empathy and prosocial emotions. *Neuroscience Letters, 693*, 49–53.  
<https://doi.org/10.1016/j.neulet.2017.06.054>
- Leo, I., Angeli, V., Lunghi, M., Dalla Barba, B., & Simion, F. (2018). Newborns' face recognition: The role of facial movement. *Infancy, 23*(1), 45–60.
- Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science, 7*(3), 297–307.
- Maibom, H. L. (2017). *The routledge handbook of philosophy of empathy*.
- Malti, T., Gummerum, M., Keller, M., Chaparro, M. P., & Buchmann, M. (2012). Early sympathy and social acceptance predict the development of sharing in children. *PLoS ONE, 7*(12), e52017. <https://doi.org/10.1371/journal.pone.0052017>
- Marcus-Newhall, A., Miller, N., Holtz, R., & Brewer, M. B. (1993). Cross-cutting category membership with role assignment: A means of reducing intergroup bias. *British Journal of Social Psychology, 32*(2), 125–146.
- McAuliffe, K., Jordan, J. J., & Warneken, F. (2015). Costly third-party punishment in young children. *Cognition, 134*, 1–10.  
<https://doi.org/10.1016/j.cognition.2014.08.013>
- Meristo, M., Strid, K., & Surian, L. (2016). Preverbal infants' ability to encode the outcome of distributive actions. *Infancy, 21*(3), 353–372.  
<https://doi.org/10.1111/infa.12124>

- Miglietta, A., Gattino, S., & Esses, V. M. (2014). What causes prejudice? How may we solve it? Lay beliefs and their relations with classical and modern prejudice and social dominance orientation. *International Journal of Intercultural Relations*, *40*, 11–21. <https://doi.org/10.1016/j.ijintrel.2013.12.004>
- Milward, S. J., & Sebanz, N. (2016). Mechanisms and development of self–other distinction in dyads and groups. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 1–4. <https://doi.org/10.1098/rstb.2015.0076>
- Misch, A., Over, H., & Carpenter, M. (2016). I won't tell: Young children show loyalty to their group by keeping group secrets. *Journal of Experimental Child Psychology*, *142*, 96–106. <https://doi.org/10.1016/j.jecp.2015.09.016>
- Morton, J., & Johnson, M. H. (1991). Conspic and concern: A two-process theory of infant face recognition. *Psychological Review*, *98*(2), 164–181.
- Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology*, *103–122*.
- Nagel, T. (1986). *The view from nowhere*. Oxford University Press.
- Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment. *Cognition*, *84*(2), 221–236. [https://doi.org/10.1016/S0010-0277\(02\)00048-3](https://doi.org/10.1016/S0010-0277(02)00048-3)
- Nichols, S. (2004). *Sentimental rules: On the natural foundations of moral judgment*. Oxford University Press.
- Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, *314*(5805), 1560–1563. <https://doi.org/10.1126/science.1133755>
- Over, H. (2018). The influence of group membership on young children's prosocial behaviour. *Current Opinion in Psychology*, *20*, 17–20. <https://doi.org/10.1016/j.copsyc.2017.08.005>

- Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *Annual Review of Psychology*, *60*(1), 339–367. <https://doi.org/10.1146/annurev.psych.60.110707.163607>
- Park, B., & Judd, C. (2005). Rethinking the link between categorization and prejudice within the social cognition perspective. *Personality and Social Psychology Review*, *9*(2), 108–130.
- Paulus, M., & Moore, C. (2012). Chapter 8—Producing and understanding prosocial actions in early childhood. In Janette B. Benson (Ed.), *Advances in Child Development and Behavior* (Vol. 42, pp. 271–305). JAI.
- Paulus, M., & Moore, C. (2014). The development of recipient-dependent sharing behavior and sharing expectations in preschool children. *Developmental Psychology*, *50*(3), 914–921. <https://doi.org/10.1037/a0034169>
- Pletti, C., Scheel, A., & Paulus, M. (2017). Intrinsic altruism or social motivation—What does pupil dilation tell us about children’s helping behavior? *Frontiers in Psychology*, *8*. <https://doi.org/10.3389/fpsyg.2017.02089>
- R Core Team. (2019). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. <https://www.R-project.org/>
- Rakoczy, H., Kaufmann, M., & Lohse, K. (2016). Young children understand the normative force of standards of equal resource distribution. *Journal of Experimental Child Psychology*, *150*, 396–403. <https://doi.org/10.1016/j.jecp.2016.05.015>
- Rakoczy, H., & Schmidt, M. F. H. (2013). The early ontogeny of social norms. *Child Development Perspectives*, *7*(1), 17–21. <https://doi.org/10.1111/cdep.12010>



- Rakoczy, H., & Tomasello, M. (2009). Done wrong or said wrong? Young children understand the normative directions of fit of different speech acts. *Cognition*, *113*(2), 205–212. <https://doi.org/10.1016/j.cognition.2009.07.013>
- Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology*, *44*(3), 875–881. <https://doi.org/10.1037/0012-1649.44.3.875>
- Rawls, J. (1999). *A theory of justice* (Rev. ed). Belknap Press of Harvard University Press.
- Riedl, K., Jensen, K., Call, J., & Tomasello, M. (2015). Restorative justice in children. *Current Biology*, *25*(13), 1731–1735. <https://doi.org/10.1016/j.cub.2015.05.014>
- Robbins, E., & Rochat, P. (2011). Emerging signs of strong reciprocity in human ontogeny. *Frontiers in Psychology*, *2*. <https://doi.org/10.3389/fpsyg.2011.00353>
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–244). Russell Sage Foundation.
- Rossano, F., Rakoczy, H., & Tomasello, M. (2011). Young children's understanding of violations of property rights. *Cognition*, *121*(2), 219–227. <https://doi.org/10.1016/j.cognition.2011.06.007>
- Roth-Hanania, R., Davidov, M., & Zahn-Waxler, C. (2011). Empathy development from 8 to 16 months: Early signs of concern for others. *Infant Behavior and Development*, *34*(3), 447–458.
- Roughley, N. (2018). Socioecological pressures, proximal psychological mechanisms and moral normativity. Situating Tomasello's natural history of human morality. *Philosophical Psychology*, *31*(5), 639–660. <https://doi.org/10.1080/09515089.2018.1486602>

- Roughley, N., & Bayertz, K. (Eds.). (2019). *The normative animal?: On the anthropological significance of social, moral, and linguistic norms*. Oxford University Press.
- Rutland, A., Hitti, A., Mulvey, K. L., Abrams, D., & Killen, M. (2015). When does the in-group like the out-group? Bias among children as a function of group norms. *Psychological Science*, 26(6), 834–842.  
<https://doi.org/10.1177/0956797615572758>
- Rutland, A., Killen, M., & Abrams, D. (2010). A new social-cognitive developmental perspective on prejudice: The interplay between morality and group identity. *Perspectives on Psychological Science*, 5(3), 279–291.  
<https://doi.org/10.1177/1745691610369468>
- Salali, G. D., Juda, M., & Henrich, J. (2015). Transmission and development of costly punishment in children. *Evolution and Human Behavior*, 36(2), 86–94.  
<https://doi.org/10.1016/j.evolhumbehav.2014.09.004>
- Schmidt, M. F. H., Butler, L. P., Heinz, J., & Tomasello, M. (2016). Young children see a single action and infer a social norm: Promiscuous normativity in 3-year-olds. *Psychological Science*, 27(10), 1360–1370.  
<https://doi.org/10.1177/0956797616661182>
- Schmidt, M. F. H., & Rakoczy, H. (2018). *Developing an understanding of normativity* (A. Newen, L. De Bruin, & S. Gallagher, Eds.; Vol. 1). Oxford University Press.  
<https://doi.org/10.1093/oxfordhb/9780198735410.013.36>
- Schmidt, M. F. H., & Rakoczy, H. (2019). On the uniqueness of human normative attitudes. In N. Roughley & K. Bayertz, *The Normative Animal?* (pp. 121–136). Oxford University Press. <https://doi.org/10.1093/oso/9780190846466.003.0006>

- Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition*, *124*(3), 325–333. <https://doi.org/10.1016/j.cognition.2012.06.004>
- Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2013). Young children understand and defend the entitlements of others. *Journal of Experimental Child Psychology*, *116*(4), 930–944. <https://doi.org/10.1016/j.jecp.2013.06.013>
- Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2019). Eighteen-month-old infants correct non-conforming actions by others. *Infancy*, *24*(4), 613–635. <https://doi.org/10.1111/infa.12292>
- Schmidt, M. F. H., & Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *PloS One*, *6*(10), 1–7.
- Schmidt, M. F. H., Svetlova, M., Johe, J., & Tomasello, M. (2016). Children's developing understanding of legitimate reasons for allocating resources unequally. *Cognitive Development*, *37*, 42–52.
- Schmidt, M. F. H., & Tomasello, M. (2012). Young children enforce social norms. *Current Directions in Psychological Science*, *21*(4), 232–236. <https://doi.org/10.1177/0963721412448659>
- Searle, J. (2010). *Making the social world: The structure of human civilization*. Oxford University Press.
- Shutts, K. (2015). Young children's preferences: Gender, race, and social status. *Child Development Perspectives*, *9*(4), 262–266.
- Sierksma, J., Lansu, T. A. M., Karremans, J. C., & Bijlstra, G. (2018). Children's helping behavior in an ethnic intergroup context: Evidence for outgroup helping. *Developmental Psychology*, *54*(5), 916–928. <https://doi.org/10.1037/dev0000478>

- Simner, M. L. (1971). Newborn's response to the cry of another infant. *Developmental Psychology*, 5(1), 136–150.
- Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(6), 679–692. <https://doi.org/10.1002/wcs.1323>
- Smetana, J. G. (1981). Preschool children's conceptions of moral and social rules. *Child Development*, 52(4), 1333. <https://doi.org/10.2307/1129527>
- Smetana, J. G. (1984). Toddlers' social interactions regarding moral and conventional transgressions. *Child Development*, 55(5), 1767–1776. <https://doi.org/10.2307/1129924>
- Smetana, J. G., Ball, C. L., Jambon, M., & Yoo, H. N. (2018). Are young children's preferences and evaluations of moral and conventional transgressors associated with domain distinctions in judgments? *Journal of Experimental Child Psychology*, 173, 284–303. <https://doi.org/10.1016/j.jecp.2018.04.008>
- Smetana, J. G., & Braeges, J. L. (1990). The development of toddlers' moral and conventional judgments. *Merill-Palmer Quarterly*, 36(3), 329–346.
- Smetana, J. G., Jambon, M., & Ball, C. L. (2014). The social domain approach to children's moral and social judgements. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 23–45). Psychology Press.
- Smetana, J. G., Rote, W. M., Jambon, M., Tasopoulos-Chan, M., Villalobos, M., & Comer, J. (2012). Developmental changes and individual differences in young children's moral judgments: Young children's moral judgments. *Child Development*, 83(1), 683–696. <https://doi.org/10.1111/j.1467-8624.2011.01714.x>
- Smetana, J. G., Schlagman, N., & Adams, P. W. (1993). Preschool Children's Judgments about Hypothetical and Actual Transgressions. *Child Development*, 64(1), 202. <https://doi.org/10.2307/1131446>

- Sommerville, J. A., Schmidt, M. F. H., Yun, J., & Burns, M. (2013). The development of fairness expectations and prosocial behavior in the second year of life: Fairness expectations and prosocial behavior. *Infancy, 18*(1), 40–66.
- Steinbeis, N. (2016). The role of self–other distinction in understanding others’ mental and emotional states: Neurocognitive mechanisms in children and adults. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*, 1–7. <https://doi.org/10.1098/rstb.2015.0074>
- Sumner, W. G. (1906). *Folkways. A study of the sociological importance of usages, manners, customs, mores, and morals*. Ginn and Company.
- Svetlova, M., Nichols, S. R., & Brownell, C. A. (2010). Toddlers’ prosocial behavior: From instrumental to empathic to altruistic helping: toddlers’ prosocial behavior. *Child Development, 81*(6), 1814–1827. <https://doi.org/10.1111/j.1467-8624.2010.01512.x>
- Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues, 25*, 79–98.
- Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology, 33*(1), 1–39. <https://doi.org/10.1146/annurev.ps.33.020182.000245>
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology, 1*(2), 149–178.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations*. (pp. 33–47). Brooks/Cole.
- Tomasello, M. (2016). *A natural history of human morality*. Harvard University Press.
- Tomasello, M. (2018). The normative turn in early moral development. *Human Development, 61*(4–5), 248–263. <https://doi.org/10.1159/000492802>

- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5), 675–691.
- Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two key steps in the evolution of human cooperation: The interdependence hypothesis. *Current Anthropology*, 53(6), 673–692. <https://doi.org/10.1086/668207>
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46(1), 35–57.
- Turiel, E. (1983). *The development of social knowledge: Morality and conventions*. Cambridge University Press.
- Turiel, E. (2006). The development of morality. In W. Damon, R. M. Lerner, & N. Eisenberg (Eds.), *Handbook of child psychology, Vol. 3. Social, emotional, and personality development* (pp. 789–857). John Wiley & Sons, Inc.
- Turiel, E. (2008). The development of children's orientations toward moral, social, and personal orders: More than a sequence in development. *Human Development*, 51(1), 21–39. <https://doi.org/10.1159/000113154>
- Turner, J. C., Oakes, P. J., Haslam, A. S., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, 20(5), 545–463.
- Vaish, A., Carpenter, M., & Tomasello, M. (2009). Sympathy through affective perspective taking and its relation to prosocial behavior in toddlers. *Developmental Psychology*, 45(2), 534–543. <https://doi.org/10.1037/a0014322>
- Vaish, A., Missana, M., & Tomasello, M. (2011). Three-year-old children intervene in third-party moral transgressions: Children intervene in moral transgressions. *British Journal of Developmental Psychology*, 29(1), 124–130. <https://doi.org/10.1348/026151010X532888>

- Verschoor, S. A., Paulus, M., Spapé, M., Biro, S., & Hommel, B. (2015). The developing cognitive substrate of sequential action control in 9- to 12-month-olds: Evidence for concurrent activation models. *Cognition*, *138*, 64–78. <https://doi.org/10.1016/j.cognition.2015.01.005>
- Verschoor, S. A., Spapé, M., Biro, S., & Hommel, B. (2013). From outcome prediction to action selection: Developmental change in the role of action-effect bindings. *Developmental Science*, 801–814. <https://doi.org/10.1111/desc.12085>
- Vezzali, L., Hewstone, M., Capozza, D., Giovannini, D., & Wölfer, R. (2014). Improving intergroup relations with extended and vicarious forms of indirect contact. *European Review of Social Psychology*, *25*(1), 314–389. <https://doi.org/10.1080/10463283.2014.982948>
- Vezzali, L., Stathi, S., Crisp, R. J., Giovannini, D., Capozza, D., & Gaertner, S. L. (2015). Imagined intergroup contact and common ingroup identity: An integrative approach. *Social Psychology*, *46*(5), 265–276. <https://doi.org/10.1027/1864-9335/a000242>
- Vezzali, L., Stathi, S., Giovannini, D., Capozza, D., & Visintin, E. P. (2015). ‘And the best essay is...’: Extended contact and cross-group friendships at school. *British Journal of Social Psychology*, *54*(4), 601–615. <https://doi.org/10.1111/bjso.12110>
- Waldmann, M. R., Nagel, J., & Wiegmann, A. (2012). Moral judgment. *The Oxford Handbook of Thinking and Reasoning*, 274–299.
- Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301–1303.
- Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, *11*(3), 271–294.

- Warneken, F., & Tomasello, M. (2008). Extrinsic rewards undermine altruistic tendencies in 20-month-olds. *Developmental Psychology, 44*(6), 1785–1788.
- Wenzel, M., Okimoto, T. G., Feather, N. T., & Platow, M. J. (2008). Retributive and restorative justice. *Law and Human Behavior, 32*(5), 375–389.  
<https://doi.org/10.1007/s10979-007-9116-6>
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biom Bull, 1*, 80–83.



## List of Figures

<i>Figure 1.</i> Differences in anticipatory looking between tools A and B as a function of condition. Error bars indicate standard errors of the mean.....	45
<i>Figure 2.</i> Box-and-whisker plots of pupil diameter from the baseline measure (Movie 2) as a function of condition.....	46
<i>Figure 3.</i> Relative change in pupil diameter (Movie 4 minus Movie 2) as a function of condition. Error bars indicate standard errors of the mean.....	46
<i>Figure 4.</i> Experiment 1 revealed that emotional sharing, but not costly sharing, had a significant effect on children’s protest behavior against unequal distributions. Results are represented using estimates (beta coefficients) and their CIs from GLM (poisson), * $p < .05$ .....	61
<i>Figure 5.</i> Experiment 2 indicated that the children protested significantly more against the unequal versus equal distribution. Moreover, despite an authority’s command to act unequally, emotional but not costly sharing had a significant effect on children’s protest behavior against unequal distributions. Results are represented using estimates (beta coefficients) and their CIs from GLMM (binomial), ** $p < .01$ . .....	66
<i>Figure 6.</i> Mean sum score of protest for the four different emotional sharing groups (0 = ignore; 1 = attend; 2 = describe; 3 = help/share/comfort) and the distribution condition (equal versus unequal distribution).....	67
<i>Figure 7.</i> Experiment 3 showed that children’s third party costly punishment behavior of an unfair distributor is associated with children’s first party costly sharing behavior. Results are represented using estimates (beta coefficients) and its CIs from GLMM (poisson), * $p < .05$ .....	73
<i>Figure 8.</i> Mean number of stickers shared with in-group and out-group members as a function of condition. Error bars indicate standard errors of the mean.....	88