Fachbereich Erziehungswissenschaft und Psychologie

der Freien Universität Berlin

Young children's understanding of and responses to moral transgressions

Dissertation

zur Erlangung des akademischen Grades

Doktorin der Philosophie (Dr. phil.)

vorgelegt von

M.A. Psychology

Vaish, Amrisha

Berlin, 2010

Erstgutachter: Prof. Dr. Michael Tomasello, Max Planck Institut für Evolutionäre Anthropologie, Leipzig

Zweitgutachter: Prof. Dr. Katja Liebal, Freie Universität Berlin

Datum der Disputation: 14. Juni 2010

TABLE OF CONTENTS

INTRODUCTORY NOTE

The nature of morality has been at the heart of philosophical and psychological discussion and debate for millennia. Does morality emerge from reason or emotions? Is morality innate or socially determined? Are humans basically good, basically evil, or blank slates? Can other animals be moral? These are not just academic questions; the answers are pivotal to any analysis of human behavior, to ascertaining what actions are wrong or right, to determining culpability and to legal systems more generally, and to knowing how best to raise moral children and thus to create a moral society.

One of the most promising ways to understand morality is to study its emergence and development in ontogeny. With this method, we can begin to get at the roots and the nature of human morality, the mechanisms that drive it, and the influences upon it. This, at a very broad and ambitious level, is the aim of this dissertation. More specifically, I investigated children's understanding of and their responses to third-party moral transgressions and the victims and transgressors therein.

1. INTRODUCTION

> It seems a happiness in the present theory, that it enters not into that vulgar dispute concerning the *degrees* of benevolence or self-love, which prevail in human nature… It is sufficient for our present purpose, if it be allowed..., that there is some benevolence, however small, infused in our bosom; some spark of friendship for human kind; some particle of the dove kneaded into our frame, along with the elements of the wolf and serpent.
>
> -- David Hume (1776/1965)

Humans are, for the most part, moral. We behave in moral ways: We help and share with others, we comfort those in distress, and we cooperate, in big groups and small, to achieve far more than any one of us could achieve alone. We also have a sense of morality, which consists of thoughts and feelings about rights and duties, good and bad character traits, and right and wrong motives and behaviors (Krebs, 2008). Moreover, our moral apparatus contains feelings of entitlement and obligation, and moral emotions such as gratitude, guilt, shame, disgust, and indignation (Tangney, Stuewig, & Mashek, 2007). Together, these thoughts and feelings lead us to monitor, judge, and react to our own motives and behaviors but also to others' motives and behaviors (Fehr & Fischbacher, 2003, 2004). The obvious question thus arises: Why?

Before tackling this question, it is important to take up a more basic question: What is morality? This question has no simple answer, but it is helpful to delineate how I will use the term. The term "morality" can be used either descriptively to refer to a code of conduct put forward by a society or some other group (e.g., a religion) or accepted by an individual for her own behavior, or the term can be used normatively to refer to a code of conduct that, given specified conditions, would be put forward by all rational persons. What both uses of the term have in common is a 'code of conduct' that helps distinguish right from wrong, good from bad, virtue from vice. This will be the very broad *psychological* sense in which I use the term: the sense of good versus bad, virtue versus vice, that is apparent in people's behaviors, judgments, and emotions. I thus agree with Blasi's (1990) general point that psychology

cannot and indeed, need not, evaluate whether a person's actions, judgments, or emotions are morally right or wrong, or what the source of morality should be (e.g., reasons or emotions); that is the realm of philosophy. Instead, psychology should aim to describe and explain the psychological underpinnings of morality. This, within a developmental context, will be my aim in this dissertation. With this in mind, let us return to the question of why we are moral.

There are at least four ways to ask this question (laid out in a seminal paper: Tinbergen, 1963). One can ask two *ultimate* level questions: How did morality come about in the course of evolution (phylogeny) and how does it contribute to our survival and reproduction (ultimate mechanisms)? One can also ask two *proximate* level questions: What are the mechanisms that bring about or motivate moral behavior in an individual (proximate mechanisms) and how does morality come about in the lifetime of an individual (ontogeny)? In this dissertation, I will explore the ontogeny of morality. However, it is useful to briefly consider the other three levels of questions, as I will also discuss these from time to time where relevant.

1.1 Ultimate level questions

From an evolutionary perspective, moral behavior can seem rather problematic[1]. For instance, helping another individual does not always benefit the helper (and indeed, can be costly to the helper, as in altruism) and would thus seem to be evolutionarily and motivationally unfeasible (Sober & Wilson, 1998). Yet it exists. What's more, at least some aspects of human morality likely have deep evolutionary roots (see de Waal, 2009; Warneken & Tomasello, 2009b). Recent experiments have shown, for instance, that our nearest primate

[1] By 'moral behavior,' I am referring to *prosocial* or *cooperative* behaviors (voluntary behaviors that benefit other individuals; Eisenberg & Fabes, 1990; Zahn-Waxler & Radke-Yarrow, 1982). Prosocial or cooperative behaviors (I will use the two terms interchangeably in this dissertation) include helping another individual achieve her individual goal, sharing a valuable resource (such as food) with another individual, comforting a distressed individual, and informing others of things that they need or want to know (Warneken & Tomasello, 2009b). More sophisticated moral behaviors also include following the group's moral norms oneself and enforcing those norms on other group members (Boyd & Richerson, 1992; Krebs, 2008). Notice that this use of the term does not commit to the motivations behind the behaviors, which, as will become clear, can range from entirely selfish to entirely selfless.

relatives, chimpanzees, retrieve out-of-reach objects for humans and assist unrelated conspecifics in obtaining food, even in the absence of rewards (Warneken, Hare, Melis, Hanus, & Tomasello, 2007; Warneken & Tomasello, 2006; but see K. Jensen, Call, & Tomasello, 2006; Silk et al., 2005). There is also some evidence that chimpanzees are sensitive to reciprocity and are more likely to help those who had previously helped them (e.g., de Waal, 1997; Gomes, Boesch, & Mundry, 2009; Melis, Hare, & Tomasello, 2008; but see Yamamoto & Tanaka, 2009). Collaborative activities (working together to achieve a shared goal) seem to be quite limited among apes, although some evidence has been found in the wild (e.g., cooperative hunting; Boesch & Boesch, 1989) and more recently in the laboratory as well (Melis, Hare, & Tomasello, 2006). Similarly, active sharing occurs rarely if ever, although chimpanzees may tolerate other individuals taking food from them and mother chimpanzees occasionally hand over low-quality food to their infants (e.g., de Waal, 1989; Ueno & Matsuzawa, 2004). There is also some evidence that chimpanzees and even monkeys are sensitive to some aspects of fairness (Brosnan & de Waal, 2003; Brosnan, Schiff, & de Waal, 2005; but see Bräuer, Call, & Tomasello, 2006). All in all, despite the contradictory findings, it seems that at least some elements of human morality are evolutionarily old. What, then, are the evolutionary functions of morality?

The basic argument for the evolutionary advantages conferred by morality is captured nicely by Krebs (2008), who writes, "The domain of morality pertains to the formal and informal rules and sanctions that uphold the systems of cooperation that enable members of groups to survive, to reproduce, and to propagate their genes" (p. 168). In line with this, the critical first step to the evolution of morality was most likely the formation of groups. Darwin (1874) argued that animals form groups for good reason: They benefit from living in groups by exchanging goods and services and by coordinating their efforts to obtain food, defend themselves against predators, build shelters, and so on. However, group living also leads to

problems, because although individuals benefit by being part of a group, it may be in each individual's interest to make others give more and take less than their share, while they themselves give less and take more. Such dilemmas are intensified by natural selection. If those who are naturally disposed to behave selfishly contribute more than their share of offspring to future generations, the proportion of selfish members in a group will increase and the proportion of cooperators will decrease. With ever fewer cooperators for the selfish individuals to exploit, there will be a decrease in the benefits of selfishness, eventually leading to the collapse of the cooperative system. The key to understanding how morality evolved, then, is to understand how animals resolve such fundamental social dilemmas (Krebs, 2008).

Several ultimate mechanisms have been proposed in this regard. One such mechanism is *kin selection*. The central concept of kin selection is *inclusive fitness* (Hamilton, 1964), or the idea that an individual's genetic fitness is measured not only by the survival and reproduction of the individual and his or her offspring, but also by the enhancement of the fitness of other relatives who share the same genes. This allows for the selection for acts that may be beneficial to others but non-beneficial or even detrimental to the individual's own fitness. Kin selection thus implies that natural selection favors genetic tendencies in which one confers benefits towards one's kin: the nearer the kin, the more helpful the behavior and the greater readiness to sacrifice one's own fitness.

Although kin selection is certainly a powerful mechanism, one of the striking aspects of human cooperation is that it happens in large groups composed of various genetically unrelated members. Kin selection, at least in its most basic form, cannot explain this level of cooperation, but other mechanisms can. One such mechanism is *mutualism*, or behaving prosocially towards another in order to gain a benefit oneself. For instance, an individual lion may be unable to hunt an elephant alone but a group of lions vastly improve their chances of

successfully hunting the elephant and each individual gaining a benefit. Importantly, mutualism does not require ongoing relationships among the participants: The group of lions could hunt the elephant and each individual lion immediately gain a benefit, and the lions would have engaged in mutualism. However, another ultimate mechanism, *reciprocity*, does require such relationships. In reciprocity, one individual provides a benefit to another now and receives a benefit in return later. When individual A helps individual B now and B helps A later, A and B have engaged in *direct reciprocity* (related to *reciprocal altruism*; see Trivers, 1971). A related but more complex ultimate mechanism is *indirect reciprocity*, in which individual A establishes a good reputation for being helpful or otherwise following or enforcing the group's moral norms, as a result of which individual A later receives benefits, but from individuals other than those he directly benefited (e.g., Alexander, 1987; Nowak & Sigmund, 1998). The flip side of being rewarded for a good reputation is being punished (or shunned) for a bad reputation, which is also thought to contribute to the evolution and maintenance of helpfulness and cooperativeness (Boyd & Richerson, 1989).

A fourth major ultimate mechanism is *group selection* (Sober & Wilson, 1998). Under the premise that cooperative social life has obvious survival value, group selection proposes that natural selection can favor characteristics that benefit the group or species as a whole rather than the individual. That is, individuals will often act in ways that do not advance their own interest but contribute to the survival of the group. If groups of many highly cooperative individuals are more successful in competitions against groups consisting of very few cooperative individuals, and if between-group benefits outweigh the within-group costs of cooperative individuals, then cooperation can evolve. Although this between- versus within-group balance is not easy to achieve at the genetic level due to factors such as migration and intermarriage, it is more effectively achieved at the cultural level. The basic argument goes like this: Having a tendency to conform one's behavior to that of the majority of one's group

is adaptive, since it allows reliable and efficient access to those behaviors that are likely to be successful in the immediate environment without the risk of individual learning or experimentation (Boyd & Richerson, 1985; Henrich & Boyd, 1998, 2001). Genetic evolution has thus produced (at least in humans) psychological mechanisms that support conformist transmission, or the tendency to copy the most frequent behavior in the group (Tomasello, 1999, 2009). As a result of conformist transmission, within-group cultural variation is suppressed while between-group variation is enhanced. Once a meta-population of culturally distinct groups is in place, there is selective pressure in favor of the persistence and proliferation of those cultural traits that are most successful for survival, including moral traits such as helpfulness, some degree of self-sacrifice for the group, enforcement of moral norms by punishment, and so on. The resulting *cultural group selection* is thus thought to have contributed significantly to the evolution of (the extensive degree of human) morality.

Together, these mechanisms (and others besides) help to explain how, in our evolutionary history, animals may have resolved the social dilemmas that arise from living in and depending on groups, leading to the evolution of prosocial behavioral dispositions (Krebs, 2008). All of these mechanisms ensure the adaptive benefits of supporting kin and others on whom one's fitness depends, coordinating effort, exchanging goods and services, upholding groups, and avoiding punishment. Inasmuch as these benefits may be quite delayed and may require quite sophisticated cognitive apparatuses, however, it is unclear what motivational role, if any, they can play (de Waal, 2008). It thus becomes crucial to consider the proximate mechanisms that explain morality. That is, what motivates an individual in a given situation to help another individual?

1.2 Proximate mechanisms

1.2.1 Background and theories: Emotion versus reason

Any discussion of the motivators of morality must begin with at least a brief introduction to two major lines of thinking that have persisted and been hotly debated over centuries. Very broadly, the debate concerns whether the basis of morality lies in (or should lie in) emotions or reason. The most prominent philosopher to argue that emotions underlie morality was David Hume, who believed that sympathy, benevolence, and a "feeling for humanity" are the true moral motivations (Hume, 1776/1965). Contrary to "selfish school" philosophers such as Thomas Hobbes (1651/1994), Hume argued that benevolent acts and sentiments cannot simply be reduced to self-interest: We may derive joy from making others happy and we may even act from the combined motives of benevolence and enjoyment, but that does not make our benevolence identical to our self-enjoyment. In other words, enjoying being good need not mean it is the reason one is good. Hume argued that we recognize good and evil by our feelings of moral approval and disapproval, which arise from our sympathy with others (or the ability to share others' feelings). Reason, he argued, does not motivate moral action but only guides the impulse received from our feelings. Thus, reason can let us infer that a particular action will lead to the death of many innocent people, but unless we *care* about those people, unless we have some *sentiment* that values human life, reason alone cannot advise against taking the action. Hume argued that a person in full possession of reason yet lacking moral sentiment would have difficulty choosing any ends or goals to pursue, and would look like what we now call a psychopath (Hume, 1776/1965).

Importantly, Hume also sought to explain how moral judgments of approval and disapproval are possible given that people are relying on their individual, distinctive sentiments. He conjectured that in making moral judgments, individuals abstract in imagination from their own particular interests and adopt an impartial point of view,

becoming what he called "judicious spectators." From this point of view, individuals assess the effects of others' actions on the interests of everyone affected. The fact that we can all adopt this perspective in imagination, he argued, accounts for a general (albeit not universal) consensus in moral judgments.

David Hume's ideas have since been echoed, modified, and expanded upon. Some of the prominent philosophers to argue similarly for the emotional basis of morality have been Adam Smith (1759/2006), Arthur Schopenhauer (1841/1995), and more recently, Martha Nussbaum (2001). Versions of Humean ideas have also spilt into moral psychology. Humean moral psychologists argue that moral judgments are intuitive, rapid, automatic, and unconscious, and that they carry good-bad or like-dislike evaluations (Haidt, 2007; Hauser, 2006). Importantly, this line of thinking is heavily influenced by evolutionary psychology; it thus features as a central claim that our moral intuitions are evolutionarily (and socially) functional rather than truth seeking. The most prominent model (the Social Intuitionist Model; Haidt, 2001, 2007) stays close to Hume's ideas by focusing on the primacy of affective responses, arguing that upon perceiving a morally relevant event, we experience emotions that lead us to moral judgments about the event, with moral reasoning only occurring post-hoc to justify the judgments. An alternative model is Universal Moral Grammar (Cushman, Young, & Hauser, 2006; Hauser, 2006; Mikhail, 2007), which argues that all humans are endowed with a moral faculty, or the capacity to intuitively judge certain types of events as good or bad based on a set of abstract moral principles, while emotional responses and moral reasoning occur post-hoc.

Hume's emotivist approach to morality was countered by moral rationalists. Most prominently, Immanuel Kant's rationalist ethical theory (1785/1959) was created as an attempt to refute Hume. Kant argued for the critical roles of reason and rationality in morality, remarking, "When moral worth is in question, it is not a matter of actions which

one sees but of their inner principles which one does not see" (1785/1959, p. 23). He argued that being motivated by duty, or what one *ought* to do, uniquely expresses an agent's commitment to morality and thus conveys a special moral worth to the agent's actions. The supreme moral principle, which is arrived at through reasoning, is an authoritative, normative principle, not a descriptive one, and as such, serves to guide us in achieving a moral life. Kant argued for the need to ground morality in a priori principles; that is, he argued that the supreme moral principle must be discovered through pure moral philosophy, grounded exclusively on principles that are inherent in and revealed through the operations of reason. He thus rejected the empirical moral philosophy and the sentimentalism espoused by Hume.

Critical to Kant's proposal was the concept of a moral point of view that is achieved once individuals abstract from their particular interests and assess situations from an impartial point of view, akin to the impartial point of view proposed by Hume. However, rather than being mainly explanatory as Hume's was, Kant's impartial perspective served as a basis from which to assess and justify moral rules or principles, and was envisioned as a perspective that any reasonable person can and should adopt in deliberating about what he ought morally to do.

Kant's moral philosophy has gathered an enormous following, evident most prominently in the works of Thomas Nagel (1979, 1986) and Christine Korsgaard (1996). Akin to Kant's impartial perspective, for instance, Nagel (1979) stressed the importance of seeing things from "nowhere in particular" when making moral decisions. Along similar lines, John Rawls (1971) proposed that principles of justice should be conceived of under a "veil of ignorance" wherein the parties involved have no access to knowledge (such as their own place in society) that could distort their judgments and lead to principles that favor their personal circumstances, thus leading them to choose principles of justice that are fair for all.

The Kantian line of thinking has also been prevalent in moral psychology, originating with developmental psychologists Jean Piaget (1932/1997) and Lawrence Kohlberg (1969, 1976), who both believed that the development of higher-level moral reasoning depends on age-related advances in cognitive skills, especially perspective taking and the ability to assume an impartial viewpoint from which one evaluates right and wrong objectively. As I will discuss theories of moral development in more detail later, suffice it here to say that several moral psychologists, especially in the field of moral development, have championed a reasoning- and cognition-based moral psychology.

More recently, however, there has been a move in philosophy, psychology, and neuroscience towards more integrative approaches to morality wherein emotional intuition and conscious reasoning are both thought to play important roles. One prominent model is the Dual-process model of moral judgments (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001), according to which deontological moral judgments (judgments associated with concerns for "rights" and "duties") are driven by automatic emotional responses, while utilitarian or consequentialist moral judgments (judgments aimed at promoting the "greater good") are driven by more controlled cognitive processes. In this model, while emotional intuition is thought to be critical to moral judgments, reasoning is acknowledged to play a restricted but significant role as well (Greene & Haidt, 2002). Another hybrid approach is that of Shaun Nichols (2002, 2004), who proposes an Affect-Backed Normative Theory of moral judgment. That is, moral judgments rely on a Normative Theory about which actions are prohibited and an affective mechanism that produces strong emotional responses to those actions.

A third integrative approach is taken by neuroscientists such as Jean Decety and Tania Singer, who argue that we do have automatic, emotional (bottom-up) responses to morally relevant events but these responses are flexible and can be controlled by cognitive (top-down)

processes (Decety & Jackson, 2006; Decety & Lamm, 2006; Singer & Lamm, 2009). Some modern developmental psychologists such as Martin Hoffman (e.g., Hoffman, 2000) and Nancy Eisenberg (e.g., Eisenberg, Shea, Carlo, & Knight, 1991) take a similar integrative approach, as I will discuss later. In this dissertation, I will also adopt an integrated view of morality in which both emotions and cognition play important roles and influence each other. As such, I will not attempt to tease emotion and cognition apart but will discuss both aspects as and when they are relevant. Let us now consider more closely some of the behaviors and emotional responses that constitute the proximate mechanisms underlying human morality.

1.2.2 Proximate mechanisms

The mechanisms that I cover here do not form an exhaustive list of the proximate mechanisms underlying morality. Indeed, an exhaustive list likely does not exist yet because researchers are still working on understanding the numerous factors that motivate morality and the interactions between those factors (but see, e.g., Fessler & Haley, 2003, for a more extensive list). Here, I focus primarily on those mechanisms that are thought to be especially important and that have received the most theoretical and empirical attention in the literature. These are also the mechanisms most relevant to the studies in this dissertation.

Note that the proximate motivators of morality can operate in dyadic interactions (when one is directly involved in and affected by the interaction) or in third-party interactions (when one is aware of but not directly involved in or affected by the interaction; see Fessler & Haley, 2003), a distinction that will emerge below. For reasons that will become clear, third-party interactions are thought to be the litmus test of moral understanding.

(i) Empathy: A critical mechanism believed to underlie morality is empathy (Batson, 1991; Hoffman, 2000). Empathy is an affective response that stems from another's (rather than one's own) emotional state or condition and is congruent with the other's emotional state or condition. The associated response, sympathy, is a feeling of concern for another in

reaction to the other's emotional state or situation (Eisenberg, 1986; Eisenberg et al., 1991; Hoffman, 1981). The origins of empathy lie in emotional contagion, that is, the tendency to automatically mimic others' emotional expressions facially, vocally, and behaviorally, and to oneself experience traces of the same emotions (Hatfield, Cacioppo, & Rapson, 1994). However, true empathy is thought to involve at least some cognitive elements, such as self-other differentiation and an awareness that another's inner states may differ from one's own (Hoffman, 1982, 2000). The cognitive ability to perspective-take, that is, imagine how the other feels, can also contribute to empathic processes by providing information about the other's situations and feelings and leading to empathic processes even in the absence of any perceptible cues to the other's distress (Eisenberg et al., 1991). These emotional and cognitive processes together create a powerful and flexible mechanism for apprehending another's emotional state and, when necessary, acting prosocially to relieve the other's negative state (Hoffman, 2000).

Empathy allows us to quickly and automatically relate to the emotional states of others, which is essential for the regulation of social interactions, coordinated activity, and cooperation towards shared goals (de Waal, 2008). Indeed, there is ample evidence that adults do empathize and sympathize with those in distress, both when the distress is visible and when it requires perspective-taking, and that these empathic processes motivate prosocial behavior (Batson et al., 1991; Batson, Duncan, Ackerman, Buckley, & Birch, 1981; Decety & Jackson, 2006; Ruby & Decety, 2004; Singer & Lamm, 2009).

(ii) Guilt: Empathy for another's distress can occur both when one is an uninvolved bystander who happens to witness the other's distress (i.e., in a third-party interaction) and when one has caused the other's distress (i.e., in a dyadic context). In the latter case, the empathy for the distress combined with the awareness of being the cause of that distress can

lead to the aversive emotion of guilt[2] (Hoffman, 1976). Guilt focuses attention on the action and the harm done (or help not given) to the other, inflicts subjective discomfort on the actor due to its unpleasant valence, and motivates the actor to make amends by aiding or otherwise compensating the victim (Baumeister, Stillwell, & Heatherton, 1994). Guilt is thus tuned to identifying and reversing the damage done to a cooperative relationship. Furthermore, anticipation of one's own guilt often leads one to refrain from intentionally harming others. Guilt can thus enhance cooperation in multiple ways, and work with adults provides convincing evidence for these functions of guilt (Cunningham, Steinberg, & Grev, 1980; Ketelaar & Au, 2003; Regan, Williams, & Sparling, 1972).

Interestingly, because experiencing (or anticipating) guilt serves important moral functions, it has been suggested that displaying guilt after transgressing serves such functions as well: It conveys to others that one is also suffering, that one intends to make amends, and that one is committed to the norms of the group and promises to do better in the future. A remorseful transgressor is thus more likely than an unremorseful transgressor to elicit forgiveness, affiliation, and cooperation from the victim and from other group members (e.g., Darby & Schlenker, 1989; Goffman, 1967; Keltner & Anderson, 2000). Indeed, among adults, the guilt displays of transgressors in both dyadic and third-party contexts do seem to serve these appeasement functions and to restore cooperation (e.g., O'Malley & Greenberg, 1983; Ohbuchi, Kameda, & Agarie, 1989).

[2] Guilt is a member of the family of "self-conscious emotions" that are evoked by self-reflection and self-evaluation (Tangney et al., 2007). The other emotions in this group are shame, embarrassment, and pride. Shame follows the failure to live up to expectations (one's own or those of significant others) that define one's ideal or core self (e.g., Tangney, 1992). Like guilt, shame can also be elicited by moral transgressions, but guilt is the more advantageous of the two emotions as it leads to reparations and thus benefits individuals and their relationships; shame is less adaptive and more costly as it causes individuals to withdraw from social contact (Barrett, 1995; Tangney, 1991) and does not lead to an increase in prosocial or cooperative behavior (de Hooge, Zeelenberg, & Breugelmans, 2007). Embarrassment is the aversive state of abashment that follows transgressions of social conventions rather than moral transgressions (Keltner & Buswell, 1997). Similarly, pride is generated by appraisals that one is socially valued or is responsible for a socially valued outcome, but has been examined most often in contexts of achievement (e.g., academic or athletic achievement) rather than moral events (Tracy & Robins, 2004). Although all four self-conscious emotions are no doubt important, I will focus my discussion on guilt and its effects, as guilt is the quintessential moral emotion (see Eisenberg, 2000).

(iii) Punishment: Social and moral norms in human societies are enforced partly due to the expectation that violations will be punished (Fehr & Fischbacher, 2004). When an individual violates a societal norm, other individuals may punish him by withholding cooperation from him, making him the target of gossip, shunning him, and so on (Boyd & Richerson, 2005). The sanctioning individual may be a second party, i.e., in a dyadic interaction with the transgressor and thus punishing to protect his own interests. However, if only second parties imposed sanctions, a very limited number of norms could be enforced because norm violations often do not directly hurt other people. Thus, second-party punishment alone is insufficient to support the emergence of the kind of large-scale cooperation seen in human societies. This requires third-party punishment, wherein the sanctioning individual is an uninvolved party who is not directly impacted by the transgression and thus acts out of an understanding of and desire to enforce the norms, perhaps in order to ensure successful future cooperation in the group and/or to conform to the majority of the group who also punish in such contexts (Boyd & Richerson, 1992; Fehr & Fischbacher, 2003, 2004; Henrich & Boyd, 2001). Third-party punishment is thus considered an especially powerful tool for maintaining morality in large groups.

Empirical work has shown that adults do indeed punish unfair or uncooperative individuals in third-party interactions (e.g., Fehr & Fischbacher, 2003, 2004; Fehr & Gächter, 2002; Henrich et al., 2006). Moreover, the empathy network in adult brains is not activated, and indeed, the reward areas of adult (male) brains are activated, when a previously unfair individual is in pain (Singer et al., 2006). Third-party punishment and related processes are thus prevalent among adults.

While punishing transgressors is essential and effective, it also seems sensible to distinguish the occasional, accidental transgressor from the frequent, intentional one. The former may otherwise be a good, norm-following group member whereas the latter may not,

a distinction that is relevant to ascertaining the degree of punishment to be meted out and the degree of cooperation to be maintained. The punishment mechanism thus needs to be flexible, that is, able to take into account various factors such as whether the transgression was intentional or accidental, whether the harm was foreseeable, whether the transgressor was remorseful, and so on. Indeed, legal systems are flexible in these ways, often holding an intentional transgressor more responsible than an accidental one (e.g., murder vs. manslaughter). It turns out that adults are also flexible in their punishment decisions: They take into account not only the consequences of a transgressor's actions but also other aspects such as the transgressor's mental states (e.g., intentions and knowledge prior to committing the transgression) and whether the transgressor apologized or showed remorse (e.g., Cushman, 2008; Fehr & Schmidt, 1999; Leary, Landel, & Patton, 1996; Ohbuchi et al., 1989; Piaget, 1997). Thus, while transgressions generally evoke punishment from adults, this response is sophisticated and flexible.

(iv) Reward: Just as it is important to punish violators of social (especially moral) norms, it would seem important to reward those who follow the norms and are thus good group members. In a dyadic interaction, when an individual benefits as the result of an actor's costly, intentional, and voluntary action, she is likely to reciprocate by benefiting the actor in the future and by defending the actor's interests (Trivers, 1971). This type of reciprocity is likely mediated by gratitude (see McCullough, Kimeldorf, & Cohen, 2008). Alternatively, one could also reward a third-party benefactor. Rewarding moral, cooperative individuals can increase cooperation by increasing the attractiveness of adhering to moral norms, although overall, punishment or the threat thereof seems to be more salient than reward as an incentive for cooperation or generosity, at least in experimental economics games (Andreoni, Harbaugh, & Vesterlund, 2003; but see Rand, Dreber, Ellingsen, Fudenberg, & Nowak, 2009).

All in all, multiple proximate mechanisms allow human adults to regulate and enforce moral behavior. These mechanisms include affective mechanisms such as empathy, cognitive mechanisms such as understanding the moral norms, and actions such as prosocial behavior and (flexible) punishment. Moreover, proximate mechanisms function in both dyadic and third-party interactions, although third-party intervention seems to be critical for the emergence and maintenance of large-scale cooperation.

In the next section, I will consider the ontogenetic roots of human morality. I will argue that the widespread morality seen in adults has deep ontogenetic roots, as seen in young children's moral judgments, behaviors, and emotions.

1.3 Ontogeny

1.3.1 Piaget and Kohlberg

The systematic study of moral development began with Piaget (1932/1997), who described how children's moral reasoning changes from a rigid acceptance of the rules of authorities to an appreciation that moral rules are a product of social interaction and are thus modifiable. He believed that interactions with peers, more than adult influence, account for advances in children's moral reasoning. By observing children playing rule-based games and by interviewing children to examine their thinking about issues such as transgressions of rules, the role of intentionality in morality, and fairness of punishment, Piaget concluded that there are two stages of development in children's moral reasoning, and a transitional period between the stages. The first stage (*morality of constraint*) is characteristic of children younger than 7 or 8 years of age who view rules and duties as unchangeable "givens." Children at this stage think of justice as whatever authorities (adults, rules, laws) say is right, and authorities' punishments are always justified. Acts that are not consistent with rules and authorities' dictates are "bad;" acts that are consistent with them are "good." At this stage, children believe that what determines whether an action is good or bad is the consequences of

the action, not the motives or intentions behind it. Piaget suggested that young children's belief that rules are unchangeable is due to two factors, one social and one cognitive. First, Piaget argued that parental control of children is coercive and unilateral, leading to children's unquestioning respect for adults and their rules. Second, children's cognitive immaturity causes them to believe that rules are "real" things that exist outside people and are not the product of the human mind.

According to Piaget, the period from age 7 or 8 to age 10 years represents a transitional period between the morality of constraint and the next stage. During this phase, children typically have more interactions with peers than previously, and these interactions are more egalitarian and involve more give-and-take than their interactions with adults. In games with peers, children learn that rules can be constructed and changed by the group. They also increasingly learn to take one another's perspective and to cooperate. As a consequence, children start to value fairness and equality and begin to become more autonomous in their thinking about moral issues. Piaget viewed children as taking an active role in this transition, using information from their social interactions to figure out how moral decisions are made and how rules are constructed.

By age 11 or 12, Piaget's second stage of moral reasoning, the stage of *autonomous morality*, emerges. Children no longer accept blind obedience to authority as the basis of moral decisions; rather, they understand that rules are the product of social agreement and can be changed if the majority of a group agrees to do so. In addition, they consider fairness and equality among people as important factors to consider when constructing rules. Children at this stage also believe that punishments should fit the crime and that punishments delivered by adults are not necessarily fair. They also now consider individuals' motives when evaluating their behavior rather than only the outcomes.

Heavily influenced by Piaget, Kohlberg (1969, 1976; Colby & Kohlberg, 1987) assessed moral reasoning by presenting children and adults with hypothetical moral dilemmas and examining how people resolved these dilemmas. In his best-known dilemma, a man named Heinz must decide whether he should break into a drug store to steal a drug that may save the life of his dying wife. On the basis of the reasoning underlying participants' responses, Kohlberg proposed that moral development proceeds through three discontinuous and hierarchical stages. He claimed that children start at the *preconventional* level of moral reasoning, during which they are egoists, focused on getting rewards and avoiding punishment. The second, c*onventional* level of moral reasoning centers around social relationships, with a focus on compliance with social duties and laws. The third level, *postconventional* moral reasoning, centers on ideals and moral principles. Kohlberg argued that people in all parts of the world move through his stages in the same order, although they differ in how many stages they attain.

Like Piaget, Kohlberg believed that age-related advances in cognitive skills, especially perspective-taking, underlie the development of higher-level moral judgment. His Cognitive Developmental theory proposed that the idea that moral obligations are rooted in convention (conventional moral reasoning) precedes the idea that moral obligations are rooted in natural law (postconventional moral reasoning). Moreover, the development of postconventional moral reasoning is related to the development of general cognitive skills, including the ability to distance oneself from what is ego- or consensus-based and to transcend to an impartial vantage point from which one evaluates right and wrong objectively, akin to Nagel's (1979) idea of seeing things from nowhere in particular and Rawls' (1971) veil of ignorance (see above).

Piaget's and Kohlberg's views have been extremely influential to the study of moral development, and their theories have generally been supported by empirical research. For

instance, consistent with their proposal that cognitive development is critical to the development of moral judgment, children's perspective-taking skills and their performance on other intelligence tests have been found to be associated with their level of moral judgment (Lickona, 1976). However, some aspects of their theories have found little support in recent research. Here, I focus on two important aspects.

(i) Moral norms versus conventional norms: One Piagetian-Kohlbergian idea that has been opposed is that young children view morality as dictated by authority figures. The opposition has come most prominently from Elliot Turiel and colleagues (e.g., Killen, 1991; Smetana, 2006; Turiel, 2006; Turiel, Killen, & Helwig, 1987), who it should be noted are otherwise in general agreement with the Piagetian-Kohlbergian tradition. Turiel and colleagues propose a social-cognitive domain theory according to which people think about the consequences of an action before determining whether the action is a moral violation. Actions that lead to injustice, harm, or violation of rights are recognized as falling within the moral domain, while actions that involve no injustice, harm, or violations of rights are treated either as violations of social conventions (locally agreed-upon customs or regulations intended to ensure social coordination and social organization, such as choices about table manners) or as personal issues (involving areas of individual prerogative, such as choice of friends or recreational activities).

Contrary to Piaget and Kohlberg, Turiel and colleagues argue that the ideas of morality and convention are not one in development but rather undergo separate courses of elaboration during development. Also, moral understandings do not emerge out of conventional understandings but rather coexist with them during early childhood. The differentiation of the moral from the conventional is explained not by reference to the development of rational reasoning but by reference to the distinguishing qualities of social interactions. A paradigmatic moral experience is the child's personal observation of the

consequences of hitting a helpless person, whereas a paradigmatic conventional experience is the extrinsic social demand for persons of a specific gender to dress in specific ways. Through social interaction and direct experience, children come to distinguish those events that possess an inherently moral quality (connecting them to issues of harm, justice, and welfare) from those whose rightness or wrongness is a matter of social consensus or personal choice.

In support of their theory, Turiel and colleagues have shown that children in many cultures differentiate between moral and social conventional issues (e.g., Nucci, Camino, & Sapiro, 1996). Their research has also shown that 3- to 4-year-old children already know a lot about morality, clearly differentiating between moral norms and social conventions, viewing moral (but not social) norms as obligatory and universally applicable, and judging moral violations as more wrong than violations of social conventions. By age 4, children generally recognize that moral violations, but not social violations, are wrong even if an adult does not know about them and even if adult authorities have not said they are wrong (for reviews, see Smetana, 2006; Turiel, 2006).

(ii) The role of intentions (and other factors) in moral judgments: Piaget's proposal that only by 11 to 12 years of age do children begin to consider a transgressor's motives when evaluating the transgression has also been criticized. In part, this is because much younger children and even infants have some understanding of others' mental states, including intentions (see Tomasello, Carpenter, Call, Behne, & Moll, 2005). For instance, infants as young as 6 months encode others' actions towards objects as being goal-directed, 9- to 12-month-old infants grasp the basics of goal-directed action (such as that actors try persistently if they fail to achieve a goal) and can distinguish between accidental and intentional actions, and infants just a few months older understand the rationality of intentional actions (Behne, Carpenter, Call, & Tomasello, 2005; Gergely, Bekkering, &

Kiraly, 2002; Gergely, Nadasdy, Csibra, & Biro, 1995; Meltzoff, 1995; Woodward, 1998, 1999). One-year-old infants can also use others' emotions to guide their own behaviors, as in social referencing (Mumme, Fernald, & Herrera, 1996; Vaish & Striano, 2004). By 18 months of age, infants grasp that others have desires and that these desires might be different from their own desires (Repacholi & Gopnik, 1997). Moreover, by 2.5 years of age, children can use others' desires to guide their interpretations of and predictions about others' actions and emotions (Stein & Levine, 1989; Wellman & Woolley, 1990). The one critical aspect of mental state understanding that was, for decades, thought to emerge quite late was belief (especially false belief) understanding, but even this has now been shown to be present during the second year (e.g., Buttelmann, Carpenter, & Tomasello, 2009; Onishi & Baillargeon, 2005; for a review, see Baillargeon, Scott, & He, 2010).

These data on children's mental state understanding (or *theory of mind*) are relevant to our understanding of moral development because mature moral reasoning depends on a consideration of mental states (see Astington, 2004). The fact that very young children turn out to have quite a sophisticated grasp on mental states of all kinds would thus suggest that they might also be far more sophisticated in their moral reasoning than Piaget believed. Indeed, a flurry of research since Piaget's time has convincingly shown that Piaget underestimated young children's ability to appreciate the role of intentionality in morality. When Piagetian moral vignettes are presented in ways that make the individuals' intentions clearer, children as young as 5 years of age (but not younger) judge individuals with bad intentions as being naughtier, more deserving of punishment, and less likable, than individuals with benign intentions (Grueneich, 1982; Imamoğlu, 1975; Yuill & Perner, 1988; Zelazo, Helwig, & Lau, 1996).

Following from this work, a host of other studies have assessed the role of factors other than intentions and outcomes on children's moral judgments. This work has revealed

impressive results. Children around 4 to 5 years of age have been found to take into account factors such as whether a transgressor apologized and whether she had already been punished (D. T. Miller & McCann, 1979; C. E. Smith, Chen, & Harris, in press; Wellman, Larkey, & Somerville, 1979). Children also take into account factors in the psychological domain such as the victim's and the children's own relationship with the perpetrator (Slomkowski & Killen, 1992; Wellman et al., 1979). Impressively, children of 4 to 5 years of age also know that simply showing distress does not automatically make the distressed person a victim and the actor a transgressor. Thus, when presented with a hypothetical scenario in which a child was crying but for no good reason (he wanted another child's cookie but the other child ate his own cookie), children did not think that a moral transgression had been committed (Leslie, Mallon, & Dicorcia, 2006). Together, all of this work points to a sophisticated moral understanding quite early in development.

Importantly, research thus far suggests that moral judgments of children below 4 or 5 years of age are not as sophisticated; thus, according to prior work, children around 3 years of age can make appropriate moral judgments about and assign punishment to actors who cause harm, but they do not seem to take factors other than the consequences of the actions into account (e.g., Wellman et al., 1979; Zelazo et al., 1996). These findings are surprising given that, as discussed earlier, children of this age are already sophisticated in their social and emotional understanding, including their understanding of others' mental states. This is thus an intriguing discrepancy that deserves to be further addressed.

In sum, Piaget and Kohlberg set the stage and provided a great impetus for the study of moral development, but recent work has demonstrated that children's moral judgments are sophisticated at younger ages than Piaget and Kohlberg estimated. Importantly, children's moral behaviors and emotions provide evidence for even deeper ontogenetic roots of morality. It is to some of this work that I now turn.

1.3.2 Prosocial behavior and motives

Research on children's prosocial behavior presents an impressive picture. At the most basic level, children's moral behavior is evident in the proclivity of 14-18-month-old infants to instrumentally help others (even strangers) in simple, instrumental ways without being praised or rewarded (Warneken & Tomasello, 2006, 2007). Indeed, rewarding children's helping behavior *reduces* its frequency, suggesting that the behavior is intrinsically rather than extrinsically motivated (Warneken & Tomasello, 2008). Young children also occasionally display other types of prosocial behavior such as spontaneously sharing with a peer or sibling (Dunn, 1988; Eisenberg & Fabes, 1998; Warneken & Tomasello, 2009b), although spontaneous sharing was not found among 18- or 25-month-old children in an experimental setting (Brownell, Svetlova, & Nichols, 2009).

It has been suggested that these early prosocial tendencies are efforts to engage positively with other people and are likely indiscriminate in their targets (Hay & Cook, 2007; Tomasello, 2009; Warneken & Tomasello, 2009a). As toddlers develop preferences, become more sensitive to context, and learn the moral and social norms surrounding prosocial behavior, their early prosocial impulses become transformed into more deliberate, selective, rule-governed, and morally informed choices (Hay & Cook, 2007). From an evolutionary point of view, it is argued that children begin by being naturally cooperative, especially while they are under the constant care of close kin and thus do not need to monitor others or be selective in their prosocial behavior; the various mechanisms thought to maintain cooperation become relevant only as children become older and must deal with others independently – at which point it becomes important to establish a reputation for being cooperative but equally important to not be exploited (Warneken & Tomasello, 2009a).

Empirically, there is some evidence that even 12-month-old infants make some distinctions between recipients of their prosocial actions, being more likely to share objects

with their peers and with their own mothers than with the peers' mothers (G. Young & Lewis, 1979). Certainly, children's prosociality seems to be discriminating by the third year. For instance, 3-year-olds' sharing of toys with a peer increased if that peer had previously shared toys with them, suggesting a sensitivity to reciprocity by this age (Levitt, Weber, Clark, & McDonnell, 1985). Similarly, when acting on a doll's behalf, 3.5-year-old children allocated more resources to the doll's kin and friends than to strangers, as well as to dolls who had been generous to their doll or to other dolls than to non-generous dolls (Olson & Spelke, 2008). By about 4 years of age, children share (even at a cost to themselves) with their friends more than with non-friends or strangers (Birch & Billman, 1986; Moore, 2009). Moreover, Fehr et al. (2008) showed that from 8 years of age, but not earlier, children allocate more resources to in-group than out-group members. Thus, young children's indiscriminate prosocial behavior is thought to blossom into older children's selective prosocial behavior. However, beyond these few factors, the selectivity of children's prosocial behavior, in particular in response to moral factors, still remains largely unexplored. For example, do young children help victims more than non-victims? Do they help moral transgressors less than non-transgressors? Given that selective prosocial behavior is an important aspect of morality from both ontogenetic and evolutionary perspectives, the nature of young children's prosocial selectivity with regard to moral factors deserves far more attention.

Like infants' prosocial behavior, infants' empathic and sympathetic responses are also evident by 14-18 months of age. Thus, when presented with a person (parent or stranger) showing pain or distress, infants as young as 14 months of age show negative affect that is resonant with the victim's affect as well as concern for the victim, and attempt to alleviate the victim's distress by comforting, helping, or sharing with her (Ungerer et al., 1990; Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992; Zahn-Waxler, Robinson, & Emde,

1992). Indeed, young children's empathy and sympathy relate positively with their prosocial behavior (Eisenberg & Miller, 1987; Hoffman, 1982) and negatively with their antisocial and aggressive behavior (P. A. Miller & Eisenberg, 1988), suggesting that empathic and sympathetic responses serve as prosocial motives. (The development of empathy-related processes will be discussed in more detail in Chapter 2).

Interestingly, Hoffman (1981, 2000) hypothesizes that in order for empathy-related responses to be reliable prosocial motives, they need to be multi-determined and flexible. That is, we should be able to empathize with others in response to whatever distress cues are available, ranging from direct emotional cues to purely situational knowledge. Indeed, adults do empathize and sympathize with victims in multiple and flexible ways, and these responses do motivate their prosocial behavior (Batson et al., 1991; Decety & Jackson, 2006; Ruby & Decety, 2004; Singer & Lamm, 2009). However, thus far, developmental work has not assessed how multi-determined or flexible young children's sympathetic responses are. In most research on the ontogeny of sympathy, children are presented with an overtly distressed person and their sympathetic responses are observed. This leaves open the important issue of whether, when the victim's distress is not perceptible, young children can nevertheless sympathize with the victim (perhaps via some form of perspective-taking) and whether this sympathy also serves as a prosocial motive.

The empathic abilities of infants develop into the conscience of toddlers. By the end of the second year, children begin to attribute causality to themselves for actions that cause distress in others, to show an appreciation for moral standards and rules, and to exhibit signs of guilt, such as remorse, apologies, and attempts to repair, when they transgress (Cole, Barrett, & Zahn-Waxler, 1992; Kopp, 2001; Zahn-Waxler & Robinson, 1995). By about 3.5 or 4 years of age, with the ability to hold multiple representations in mind, children's guilt experiences and attempts at reparations become more complex, less egocentric, and more

organized with reference to the internal state of the victim (Mascolo & Fischer, 2007). There is also evidence for a link between the experience of guilt and moral development (Kochanska, Gross, Lin, & Nichols, 2002; Zahn-Waxler & Kochanska, 1990). The experience of guilt in dyadic interactions (i.e., when the child has caused harm) thus emerges and plays an important role in morality from early in ontogeny.

Importantly, though, almost nothing is known about children's understanding of guilt displayed in third-party interactions. As discussed above, guilt displays are argued to serve important moral functions. For example, a transgressor's display of guilt shows that he did not intend harm, that he is aware of and committed to the norms of the group, and that he promises better conduct in the future; it thus reduces the likelihood that he will be punished and, compared to an unremorseful transgressor, increases the likelihood that others will affiliate and cooperate with him (e.g., Keltner & Anderson, 2000; Leary et al., 1996). Whether young children are sensitive to these appeasement functions of guilt displays is a fascinating question that has yet to be studied.

1.3.3 Responses to transgressions

Researchers have also assessed young children's responses to others' transgressions, in both dyadic and third-party interactions. In dyadic interactions, when young children perceive themselves to be the victims of moral violations, they actively defend themselves. Toddlers protest, argue, and use justifications in disputes with family members, especially disputes concerning violations of their rights (Dunn & Munn, 1987). By 4 years of age, children tattle on their peers, particularly about moral violations such as physical harm (Ingram & Bering, in press). Importantly, although these behaviors may provide evidence for some basic moral understanding and may well mark the beginnings of more extensive moral development (Dunn, 1988), they are not the best test cases for moral understanding as such. This is because responding to moral violations against oneself may be driven by emotional

reactions (e.g., anger) or self-interest alone. Indeed, even chimpanzees retaliate against conspecifics that steal food from them, and they do so as a function of how angry they are at the theft (K. Jensen, Call, & Tomasello, 2007). The real test cases for moral understanding are thus third-party interactions, as moral behaviors in such interactions show agent-neutral applications of moral norms[3]. Indeed, in evolutionary analyses of the origins of morality, the key situations are instances of third-party intervention, which is also thought to be human-specific (Gintis, Bowles, Boyd, & Fehr, 2003; Krebs, 2008; Tomasello, 2009).

Surprisingly little is known about children's reactions to third-party transgressions, but recent work suggests that infants and young children do at least seem to monitor and evaluate third-party moral interactions. For instance, already during the first year of life, infants watching a protagonist character being helped or hindered by other characters expect the protagonist to approach the helper rather than the hinderer (Kuhlmeier, Wynn, & Bloom, 2003); infants also prefer to touch the helpful more than a neutral character, and the neutral more than the hindering character (Hamlin, Wynn, & Bloom, 2007). Whether these very basic social evaluations impact children's subsequent interactions with the helpers and hinderers is an exciting but still unanswered question.

More generally, prior work leaves unclear whether children engage in the kind of third-party intervention that is considered so critical for the emergence and maintenance of human morality. Do young children punish third-party transgressors and/or reward third-party helpers? If so, how flexible is this intervention? For instance, do children (like adults) consider a transgressor's intentions, or whether a transgressor displayed guilt, in their

[3] This is not to say that dyadic interactions are irrelevant. On the contrary, they constitute our most common interactions, wherein we are emotionally or motivationally involved rather than detached bystanders. For this reason, responses to dyadic versus third-party interactions can also differ in notable ways. For instance, Rochat et al. (2009) found that children are more selfish when distributing resources in a dyadic context (when they are potential recipients and are thus emotionally invested in the outcome) versus in a third-party context (when they are not potential recipients). Thus, although I stress that third-party interactions are the test cases for moral understanding as such, it is important to keep in mind that dyadic interactions form an absolutely essential piece of the morality puzzle as well.

decision to punish? We also know almost nothing about whether children actively intervene in third-party transgressions (e.g., by protesting and tattling, as they do in response to dyadic transgressions; see above). These gaps in our knowledge are surprising given that third-party intervention (whether punishment, reward, protest, or tattling) would be evidence that children are not only aware of basic moral norms but also attempt to enforce these norms in an agent-neutral manner rather than simply out of self-interest, and given the import accorded to third-party intervention for the emergence and maintenance of large-scale human morality.

1.4 Focus of the dissertation

Broadly speaking, the purpose of this dissertation was to explore the ontogenetic origins of human morality. Specifically, in four studies, I explored young children's understanding of and their responses to moral transgressions and the victims and transgressors therein. The moral transgressions featured in all of the studies were third-party interactions. Primarily, this was for the reasons mentioned above, namely, that third-party interactions are considered the litmus test of moral understanding and intervention in third-party interactions is considered essential to human morality. There was also a third, ethical reason: Having children witness and react to third-party interactions keeps children from themselves having to assume the role of victim or transgressor.

As previously mentioned, I consider morality to be comprised of moral behaviors, judgments, and emotions. As such, each study assessed one or more of these aspects of morality; which one(s) depended on the question(s) being addressed in the particular study. Also, the ages I studied ranged from 18 months to 5 years; again, this was determined by the question(s) being addressed in each study and the ages suggested by relevant prior work. The inclusion of multiple aspects of morality and of a relatively wide age range are, I think, valuable, as each study contributes depth to our understanding of moral development but the dissertation as a whole also contributes breadth.

In Study 1, I explored 18- and 25-month-old children's sympathetic and prosocial responses to a victim. As discussed above, in most research on the ontogeny of sympathy, young children are presented with an overtly distressed person and their sympathetic responses are observed. This work leaves unclear whether, when the victim's distress is not perceptible, young children can nevertheless sympathize via a more cognitive route, perhaps by perspective-taking. This was the question addressed in Study 1.

Studies 2 and 3 assessed various forms of third-party intervention. The two parts of Study 2 assessed third-party intervention in the form of selective prosocial behavior towards actors in third-party interactions. In Study 2a, I asked whether 3-year-old children reduce their prosocial behavior towards a moral transgressor and/or increase their prosocial behavior towards a helpful actor. In Study 2b, following upon but extending the long tradition of work on children's moral judgments (see above), I explored whether 3-year-old children's prosocial behavior towards a transgressor varies depending on the intentions of the transgressor.

In Study 3, I examined whether children actively enforce moral norms by intervening and attempting to prevent third-party moral transgressions from taking place. In particular, I studied 3-year-old children's intervention (verbal protest) as they observed a moral transgression, as well as their subsequent tattling and prosocial behavior towards the victim of that transgression.

Finally, Study 4 examined the flexibility of children's moral understanding. I asked whether 4- and 5-year-old children, who make quite sophisticated moral judgments that take into account multiple factors about the transgressor and the context of the transgression (see above), also have the flexibility to modify those judgments based upon the transgressor's subsequent remorse. Specifically, I explored whether children judge a transgressor who displays guilt differently from one who displays no guilt.

2. STUDY 1: SYMPATHY THROUGH AFFECTIVE PERSPECTIVE-TAKING

2.1 Introduction

Empathy and sympathy are critical mechanisms underlying morality (Batson, 1991; Hoffman, 2000). Empathy and especially sympathy are thought to lead to prosocial behaviors such as helping and away from antisocial behaviors such as aggression (Batson, 1998; Batson et al., 1991; Eisenberg & Miller, 1987; Hoffman, 1982, 2000; P. A. Miller & Eisenberg, 1988). Thus, to better understand how humans experience and share others' emotions as well as the motivations behind human pro- and antisocial behavior, it is important to examine the development of these emotional responses.

Empathy and sympathy have been defined in a number of ways (Batson, 1998; Decety & Jackson, 2006; Eisenberg & Strayer, 1987; Feshbach, 1978; Hoffman, 1982; Preston & de Waal, 2002; Wispé, 1986, 1987). Here, I will adopt a conceptualization offered by Nancy Eisenberg and colleagues (Eisenberg et al., 1994; Eisenberg, Spinrad, & Sadovsky, 2006; see also Hoffman, 1982; Hoffman, 2000) in which empathy is an affective response that stems from the apprehension or comprehension of another's emotional state and is similar to what the other person is feeling or would be expected to feel, and sympathy is an affective response that stems from the apprehension or comprehension of another's emotional state but is not the same as the other's state, consisting instead of feelings of sorrow or concern for the other. The related response, personal distress, is an aversive, self-oriented affective reaction to the apprehension and comprehension of another's emotion (Batson, 1987; Eisenberg et al., 1991).

Under this conceptualization, empathy-related responses (empathy, sympathy, and personal distress) can be aroused in multiple ways (Hoffman, 2000). At the most basic level, they can be aroused by preverbal, automatic, and essentially involuntary processes, including motor mimicry of the other's emotional cues, emotional contagion, classical conditioning,

and direct association of cues from the victim with one's own painful past experience. These three processes arouse an involuntary affective response and require only a shallow level of cognitive processing. Still, these simple modes of empathic arousal are important because they show that humans are built to involuntarily and forcefully experience another's emotion (Hoffman, 2000). These three modes are crucial for arousing empathy in early childhood but they continue to provide an involuntary route to empathy throughout life, especially in face-to-face situations in which the victim's distress is directly perceptible.

Importantly, Hoffman (2000) argues that the arousal of empathy does not require the victim's distress to be observable, since, even in the absence of observable cues, the capacity to represent events and imagine oneself in another's place can still lead to empathic responses. One such higher-level cognitive process that contributes to empathy involves activating and accessing elaborated cognitive networks (made up of social scripts in which extensive information about situations and people in general are embedded). These networks develop out of direct or vicarious experiences, and once created, can be informative about similar experiences in the future. A second higher-level cognitive process is affective perspective-taking, which involves viewing a target's situation from his perspective and thereby making an inference about his affective state (Eisenberg et al., 1991; Feshbach, 1978; Hoffman, 1982, 1984). As these modes involve more sophisticated levels of cognition than the simpler processes, they are expected to emerge later in development (around 2-3 years of age, according to Hoffman, 1982, 1984).

Eisenberg et al. (1991) argue that sympathy and personal distress are generally outcomes of sophisticated cognitive processes such as retrieval of cognitive networks or affective perspective-taking because these processes involve an analysis of the source of the vicarious feeling and therefore lead to a focus on the other rather than (or in addition to) the self. These cognitive processes then either lead directly to sympathy and/or personal distress,

or lead to empathy, which, with some additional cognitive processing of the sophisticated kind, then leads to sympathy and/or personal distress. Thus, when someone sympathizes with another person, the sympathizer has likely engaged in some relatively sophisticated cognitive processes (see Eisenberg et al., 1991, for an extensive discussion of the affective and cognitive aspects of empathy-related responding).

Much of the research on the development of empathy-related responding has assessed how infants and young children respond to observable (visible or audible) distress cues. The typical paradigm involves presenting children with a person (infant, mother, or stranger) displaying distress, and measuring children's emotional responses to the victim (as evident in their facial, vocal, gestural, or physiological reactions). This research suggests that soon after birth, infants automatically cry in response to other infants' cries (e.g., Sagi & Hoffman, 1976; Simner, 1971). This *global empathy* is thought to be a precursor of empathy (Hoffman, 1982, 2000). Around 12-14 months of age, infants show *egocentric empathic distress*, in which they respond to another's distress as if they themselves were in distress. This happens during the developmental period in which infants can feel empathic distress (from early preverbal arousal modes) but they still lack a clear differentiation between self and other. Slightly older infants show *quasi-egocentric empathic distress*, in which they realize that the distress is the other's and not their own, but they confuse the other's inner states with their own and try to help by doing for the other what they would themselves find comforting, such as giving the person a teddy bear that they themselves find comforting (Hoffman, 1982; Ungerer et al., 1990; van der Mark, van IJzendoorn, & Bakermans-Kranenburg, 2002). True or v*eridical empathy* emerges in the second half of the second year, when children more fully differentiate between self and other and thereby understand that others are separate beings (as seen in their ability to recognize themselves in the mirror; Lewis, Sullivan, Stanger, & Weiss, 1989). With this important cognitive development, toddlers realize that others have

independent inner states, and they now begin to show increasing empathic and sympathetic responses to others' visible distress as well as appropriate, other-directed comforting and prosocial behavior (Bischof-Köhler, 1991; Eisenberg & Fabes, 1998; Zahn-Waxler & Radke-Yarrow, 1982; Zahn-Waxler, Radke-Yarrow et al., 1992; Zahn-Waxler, Robinson et al., 1992).

Surprisingly, prior work has not assessed the emergence of the more sophisticated cognitive routes to sympathy and empathy despite the fact that these sophisticated processes are crucial ways to understand others' minds and experiences, and despite the extensive theorizing about the emergence and development of these processes (see especially Eisenberg et al., 1991; Hoffman, 1984, 2000). Behavioral and neuroscientific work has shown that adults do frequently engage in these sophisticated processes (Batson et al., 1997; Ruby & Decety, 2004; see Blair, 2005; Decety & Jackson, 2006, for reviews). The question addressed in the present study, therefore, was whether as early as the second year, children can rely on sophisticated cognitive processes (in the absence of emotional cues) to infer a person's state and thereby sympathize with that person.

There are two lines of developmental research relevant to this topic, but both are problematic. The first involves the use of picture and story assessments (Eisenberg-Berg & Lennon, 1980; Feshbach & Roe, 1968; Iannotti, 1985), in which the child is typically told brief stories while being shown pictures of hypothetical protagonists in emotion-eliciting situations (e.g., a child has lost her dog or a child is at a birthday party) but no information about the protagonists' feelings. After each story, the child is asked to indicate verbally or by pointing to pictures of facial expressions how she herself feels, under the assumption that the child has responded empathically if her reported emotion is similar to that of the protagonist. Eisenberg and Miller (1987) argue, however, that among other problems, these assessments create strong demand characteristics and thus tap children's inclination to provide the socially

appropriate response rather than their empathic responses (see also Eisenberg & Lennon, 1983; Eisenberg et al., 2006).

The second relevant line of research has examined children's affective perspective-taking skills. Similar to the picture and story assessments, affective perspective-taking tasks typically involve the child being told stories about a character's situations and then being asked how the character feels. By 2.5 to 3 years, children are quite successful at these tasks, suggesting that young children have some understanding of others' affective states (Dunn & Hughes, 1998; Harwood & Farrar, 2006; Wellman, Phillips, & Rodriguez, 2000; Wellman & Woolley, 1990). However, these tasks do not test children's sympathy, i.e., it remains unclear whether, having comprehended the other's affective state, the child feels concern for the other. Finally, a problem common to both lines of research is that the tasks used require relatively sophisticated cognitive abilities (e.g., the ability to understand hypothetical stories) and linguistic skills, which limits the age testable using these tasks.

To my knowledge, only one recent study has assessed sympathy in the absence of observable distress. Hobson, Harris, García-Pérez, and Hobson (2009) tested 11-year-olds with autism, 11-year-olds with learning disabilities, and typically-developing 6-year-olds (all groups had verbal mental ages of around 6 years). In their task, participants and two experimenters each drew a picture. Then, in the experimental condition, one experimenter (the perpetrator) unexpectedly tore up the other experimenter's (the victim's) drawing, or, in the control condition, the first experimenter tore up a blank sheet of paper. In both cases, the victim observed the first experimenter neutrally (i.e., did not display any emotion). Children's looks to and concern for the victim were analyzed. In the experimental condition, a significantly higher percentage of children without autism than children with autism looked immediately and spontaneously to the victim and showed concern for the victim. These differences did not emerge in the control condition.

I adapted Hobson et al.'s (2009) task, for two reasons: (i) it does away with the affective cues typically provided in work on sympathy, thus allowing for a test of sympathy in the absence of emotional cues, and (ii) since the task is non-hypothetical and non-linguistic, it does away with the difficult task demands placed on children in picture/story and existing affective perspective-taking tasks. I extended their task by introducing several scenarios in addition to the drawing scenario (hereafter called *sympathy situations*).

Like Hobson et al. (2009), along with measuring patterns of children's looks to the victim, I also examined children's concern for the victim. One potential problem with measuring concern is that perhaps children look concerned about the generally negative situation (e.g., someone tearing someone else's picture) without really being concerned for the victim. To address this issue, I took two steps. First, like Hobson et al., I only coded those concerned looks that were directed towards the victim. Second, extending Hobson et al.'s work, I assessed children's prosocial behavior towards the victim in a subsequent task (hereafter called *prosocial situation*). This step was taken because sympathy is thought to play an important role in motivating prosocial behavior. As mentioned earlier, during the second year of life, children display prosocial behavior such as comforting or making helpful suggestions (e.g., S. K. Young, Fox, & Zahn-Waxler, 1999; Zahn-Waxler, Robinson et al., 1992), and their empathy and sympathy for victims who show overt emotional cues correlate with their prosocial behavior (e.g., Eisenberg et al., 1989; Eisenberg & Miller, 1987). Thus, to test whether children's expressions of concern represented sympathy for the victim rather than more general concern, I examined whether children's prosocial behavior towards the victim was greater after they had witnessed situations that aroused sympathy for the victim than after situations that were neutral in nature. Finally, to better compare the present study with prior work, I also assessed associations between children's concern and their subsequent prosocial behavior.

I predicted that toddlers would show more concern for an adult when she was harmed than when she was not. Note that although the victim showed no emotional response, I nevertheless assessed children's emotional response (concern for the victim). My aim was thus not to assess children's cognitive skills per se, but rather to assess whether children could arrive at an affective response without any overt affective input (and thus without use of the simple modes of empathic arousal such as emotional contagion). Based upon past work, I also predicted that toddlers would subsequently help the victim more and that there would be an association between concern for and prosocial behavior towards the victim.

Since prior work has mostly been conducted with children in the second year, I too tested 1.5- and 2-year-olds to assess whether at these ages, children can sympathize with a victim not only when she shows overt emotions but even when she does not. I also piloted the procedure with a few 14-month-old infants but found that they did not fully grasp the sympathy situations. I therefore did not further test this age group. Finally, given that sympathy and prosocial behavior have sometimes been found to vary by gender, I assessed the effects of gender, but I did not have any specific predictions regarding this variable due to the mixed results from prior work (e.g., Holmgren, Eisenberg, & Fabes, 1998; Zahn-Waxler, Radke-Yarrow et al., 1992; see Eisenberg & Lennon, 1983).

2.2 Method

Participants

Participants were 18-month-olds (n = 32, 16 girls) between 17;01 and 18;28 (M = 18;2; SD = 11.7 days) and 2-year-olds (n = 32, 16 girls) between 23;15 and 26;28 (M = 25;16; SD = 31.8 days) from a medium-sized German city. Children were recruited from a database of parents who volunteered to participate in child development studies. All participants were native German speakers and most came from middle class backgrounds. No information concerning parents' education or occupation was collected. Although precise

information about the sample's ethnicity was not available, about 98% of the population from which the sample was drawn is native German. All children were seen in a child laboratory for a 45-min "play" session. Five additional children were tested but excluded due to fussiness or inattentiveness during the sympathy situations (two 2-year-olds and two 18-month-olds) and equipment failure (one 18-month-old). All participants were tested by the same two female experimenters playing the same roles each time.

Materials

Each child saw four sympathy situations in which the following materials were used: two similar-looking necklaces with large, colorful beads; two similar-looking black belts with large, colorful beads; blank sheets of white paper and a color pencil; and a blue and a red ball of clay. Before each sympathy situation, children and one experimenter (E1, who would later play the victim) played together with one of two filler toys: an age-appropriate puzzle or a "climber" toy (consisting of a ladder and a wooden man). During the prosocial situation, three similar-looking colorful balloons were used, one filled with helium and the others filled with air. The helium balloon was E1's balloon and was tied to a piece of string, whereas the two air balloons were the children's and were tied to plastic yellow sticks that were easy for children to hold. Between the last sympathy situation and the prosocial situation, a ball and a stuffed toy served as filler toys.

Setting

During the sympathy situations, children sat on their parent's lap at a 120 X 70 X 75 cm table, facing E1, while a second experimenter (E2, who played the perpetrator) sat beside the children on their right. For the prosocial situation, the child and E1 moved to a red carpet (200 X 140 cm) in the same room while the parent sat on a chair close-by.

Procedure

E1 and E2 first played with children in a waiting room for about 10 minutes and E1 obtained parents' informed consent. Throughout, E1 wore one of the necklaces and belts described above (in order to make it seem as though these really belonged to her and she enjoyed wearing them). When children were comfortable, parents and children were taken to the testing room, where everyone took their designated seat. Parents were asked not to provide the children with cues and to look away if children looked at them during the study. The overall experimental procedure was as follows: All children saw four sympathy situations, but half of the children saw them in the *harm* condition and half saw them in the *neutral* condition. After the sympathy situations, all children took part in the same prosocial situation.

To get started, E1 and the children played with the climber toy for 2 or 3 minutes, after which E1 put the toy away and the first of four sympathy situations began. Each situation (in both conditions) began with a phase (45 s) in which E1 acted on one of the four target objects (which, in the harm condition, would later be taken or destroyed by E2). The four situations were as follows:

Necklace. E1 admired and showed off her necklace. This involved looking admiringly at it, taking it off to examine it, commenting on the different beads, stating how much she liked it, and so on. The second, similar-looking necklace lay on a tray to the right of E2, visible but inaccessible to the children.

Belt. E1 admired and showed off her belt in a similar way as with her necklace. The second, similar-looking belt lay on the tray to the right of E2.

Picture. E1 happily and proudly drew a picture of a house and an apple tree, commenting the entire time about what a pretty picture it was, how much she liked it, and how happy it made her. To begin drawing, E1 picked up a stack of blank paper that had been lying out of the

children's view, took one sheet for herself, and left the remainder of the stack on the table, visible but inaccessible to the children.

Clay. E1 happily and proudly made a clay bird using either the blue or the red clay ball and commenting as in the picture situation. To begin this task, E1 picked up a small tray that had thus far been lying out of the children's view and that held both balls of clay; she took one of the balls of clay for herself and left the other ball on the tray, visible but inaccessible to the children.

The necklace and belt comprised *possession* situations, and the picture and clay comprised *effort* situations. Since this was the first study of its kind with young children, I was unsure about what kinds of situations might elicit sympathy when the victim provided no emotional cues. I thus used two different kinds of situations in order to increase the chances that children would show sympathy in at least one kind. I did not have any predictions about which (if any) type of situation might elicit more sympathy.

In each case, E1 acted on the object for 45 seconds, during which time she occasionally looked to the child to share her excitement or to re-engage the child but was mostly focused on the objects and her actions. During these 45 seconds, E2 watched E1's actions with mild interest but did not speak and did not look at the child. For each situation, when the 45 seconds were over, E1 placed the target object (her necklace, belt, picture, or bird) in front of her on the table while still looking admiringly at it. At this point, the experimental manipulation began.

Half of the children in each age group were randomly assigned to the harm condition, and the other half to the neutral condition. In the harm condition, E2 grabbed the target object as soon as E1 had put it down, said in a mildly aggressive tone, "I'm going to take/tear/break this now," and proceeded to do so mildly aggressively for 15 seconds. Specifically, E2 put on the necklace or belt and looked at it admiringly, or tore up the picture or broke apart the bird

into small bits and threw the bits into a bin lying to her right on the ground. In the neutral condition, E2 said the same words in a neutral tone of voice and produced the same actions in a more neutral way upon the second (similar) object; i.e., in the necklace and belt situations, E2 put on the necklace or belt lying on the tray; in the picture situation, E2 tore up a blank sheet of paper; and in the clay situation, E2 broke apart the second ball of clay. Critically, regardless of E2's action or the condition, E1 silently watched E2's actions with a neutral face; she neither spoke to nor looked at the child or anywhere else during this time. E2 also only watched her own actions; she did not look at E1 or the child during this time. Children's looks to E1 (the victim) were coded during these 15-second periods (see Coding and Reliability section below). After 15 seconds, E2 stopped acting upon the target object, which indicated the end of the trial to E1; E1 then neutrally picked up a filler toy and engaged children with it for approximately 1 minute while E2 neutrally looked away (e.g., at the bin lying near her) and did not engage in the play.

A manipulation check was conducted on a random 25% of participants (n = 16; eight in each age group and in each condition) to ensure that E1 maintained a neutral expression during the 15-second intervals in which E2 acted. A coder who was blind to condition coded E1's facial expression on a 5-point scale, consisting of -2 (very negative), -1 (somewhat negative), 0 (neutral), 1 (somewhat positive), and 2 (very positive). The scores were 0 in 62 of 64 instances, and 1 in the remaining two instances (M = .03, SD = .18), indicating that E1 did indeed maintain a neutral expression.

Each child saw all four situations in counterbalanced order, alternating between possession and effort situations, after which E1 and the children moved to the floor to play with filler toys while parents moved to a chair near the carpet and were given something to read. Parents were told that if children offered them a balloon, they could take it but then should put it on the floor and go back to reading.

After 2 or 3 minutes of play, E2 took out the three balloons and said excitedly, "Look (name of child), look what I found! Balloons!" She handed the two air balloons to the child and the helium balloon to E1. E1 played happily with her balloon and did not engage with the children, and children generally played with their own balloons. About 1 minute later, E1 "accidentally" let go of her balloon (which floated to the ceiling), gasped, pointed to her balloon, and said in a shocked voice, "Oh no, my balloon!" She then "attempted" to bring it down, failed, and sat back down. She was then vocally and facially obviously sad. During the next two minutes (from the moment E1's balloon hit the ceiling), children's behavior was coded (see Coding and Reliability section below).

During these 2 minutes, E1 never looked at children's hands or at their balloons and only very rarely looked at them at all so as to prevent giving them hints or pressuring them to help. After the 2 minutes, E1 stood on a chair, brought her balloon down, and was obviously happy. The prosocial situation did not last the full 2 minutes (1) if children became very upset, in which case the study was cut short and E1 brought down her balloon, or (2) if children handed one or both of their balloons to E1, in which case E1 gratefully took and then handed back the balloon(s) before bringing down her own balloon. For children who had seen the harm sympathy situations, after the entire procedure was completed, E2 apologized to E1 while the children were paying attention, and E1 accepted the apology. This was done so as not to end the session on a negative note, and in order to show children that E2's behavior had been wrong.

Coding and Reliability

In the sympathy situations, the four 15-second intervals during which E2 acted upon the objects were coded. The primary coder (the first author, who was not blind to condition) used Interact (Mangold, 2007) to code looks to E1's face, E2's face, E2's actions, and away. (However, looks to E2's face, E2's actions, and away were not analyzed and will not be

discussed further.) Reliability was assessed on a randomly selected 25% of children (eight in each age group) by two secondary coders who were blind to condition, one of whom coded six children in each age group (three in each condition), and the other of whom coded four children in each age group (two in each condition). Agreement with the primary coder was excellent: $\kappa = .81$ for 2-year-olds and $\kappa = .80$ for 18-month-olds.

The primary coder also coded the quality of all looks to E1 using three categories (based party on Hobson et al., 2009): concerned, checking, and other looks. Concerned looks were those expressing concern for E1. For a look to be coded as 'concerned,' children's facial expression while looking to E1 had to involve either a furrowing or raising of the brow and sadness or concern in the eyes. In addition, their expression had to be different from that just before they turned to look to E1 as well as different from the overall facial expression that they had shown during E1's presentation. Counting only those looks of concern that were directed at E1 made the measure of concerned looks rather conservative since, of course, a child might experience concern for the victim even when she is looking away from the victim. However, since a concerned expression not directed at the victim might be the result of a general worry or confusion about the situation, I thought it safer to count only those concerned looks specifically directed at the victim.

Checking looks were looks meant to evaluate the situation, E1's response, and what might happen next (somewhat similar to the hypothesis-testing category used by Zahn-Waxler, Radke-Yarrow et al., 1992). These looks were accompanied by neutral facial expressions or facial expressions that were no different from those just before the children turned to look to E1 and from the children's overall facial expression during E1's presentation. Checking looks were coded to gauge children's expectation of a reaction from E1. I predicted that even if children did not show concern, they would show more checking looks in the harm than in the neutral condition because they perceived the harm condition as

affecting E1 more. Finally, other looks were any looks that were not coded as concerned or checking (e.g., looks during which children smiled at E1). However, as almost no significant differences emerged with regard to other looks, and as these looks were not theoretically interesting, they will not be discussed further.

Due to the subjective nature of the coding of quality of looks, two secondary coders who were blind to condition assessed reliability on 100% of children: One secondary coder coded 24 children in each age group (12 in each condition) and the other coded eight in each age group (four in each condition). Agreement with the primary coder was excellent: $\kappa = .82$ for 2-year-olds and $\kappa = .80$ for 18-month-olds. Despite the high reliability, I used the blind coders' coding of quality of looks in analyses to avoid any bias in the primary coder's coding.

In the prosocial situation, the primary coder coded the 2 minutes or, if the trial was shorter, the full trial length, using the following categories, with their associated scores in parentheses (ordered from the highest to lowest level of prosocial or emotional response): Helps/shares (3), Shows distress; Describes situation for self or E1 (2), Attends to situation (1), or Ignores situation (0; see Table 1 for details). 'Shows distress' and 'Describes situation for self or E1' were assigned scores of 2 because I took these to be greater emotional responses to or involvement in the other's situation than 'Attends to situation' or 'Ignores situation.' These categories were based partially on prior work (Zahn-Waxler, Radke-Yarrow et al., 1992).

Although children could show any or all of these prosocial behaviors, for analyses, children's *prosocial score* consisted of each child's highest score. Since no child's prosocial score was 0 (Ignores situation), this category was not included in analyses. Two coders who were blind to condition assessed reliability on 25% of children: One coded six children in each age group (three in each condition), and the other coded four children in each age group

Table 1. Coding scheme for prosocial situation

Prosocial score	Category	Behaviors
3	Helps/shares	• Gives own balloon to E1: fully approaches E1 and clearly offers her one or both balloons • Puts balloon near E1 or throws it towards E1; may then move away: tosses balloon(s) in E1's direction or places it/them near her and then retreats, usually still watching her. If it was clear during testing that the child intended to give the balloon(s) to E1, E1 picked up the balloon(s) and the 2 minutes were cut short, but if E1 was unsure about what the child intended, she continued displaying sadness • Comforts E1: hugs or pats E1 • Describes the situation to parent: verbal or gestural descriptions about the situation (e.g., "The balloon is gone") directed to parent in an effort to draw the parent's attention to the situation; akin to Zahn-Waxler, Radke-Yarrow et al.'s (1992, p. 129) "indirect helping" • Makes suggestions to E1: suggests ways to retrieve balloon (e.g., "Ladder") or to cheer E1 up (e.g., "Ball," referring to the ball that E1 had previously enjoyed playing with)
2	Shows distress	• Shows distress: including whimpering or crying
	Describes situation	• Describes situation for self or E1: verbal descriptions (e.g., "Balloon is up") or gestures (e.g., pointing to balloon at ceiling), while looking not to parent but to situation or E1; akin to Zahn-Waxler, Radke-Yarrow et al.'s (1992, p. 129), "hypothesis testing" • Points out to self or E1 that s/he has balloon(s): verbal (e.g., "I have a balloon") or gestural communication (e.g., pointing to own balloon(s)) while looking not to parent but to situation or E1
1	Attends to situation	• Watches E1 and situation in a serious way; stops play • Goes to parent or moves away but continues watching E1
0	Ignores situation	• Shows no involvement or interest in the situation • Goes to parent and tries to engage him/her

(two in each condition). Agreement on the prosocial scores was excellent: $\kappa = .80$ for 2-year-olds and $\kappa = .81$ for 18-month-olds.

2.3 Results

I first report results from the sympathy situations, followed by results from the prosocial situation, and finally the correlations between the two. All tests of significance are two-tailed. Effect sizes were calculated using partial eta-squared (η^2_p).

2.3.1 Sympathy situations: Patterns of looks

To assess patterns of children's looks to E1, four dependent measures were used: number of the four trials in which children looked to E1, average latency of first look to E1, average total duration of all looks to E1, and average number of looks to E1. Average latency and duration were obtained by averaging across only those trials in which children looked to E1.

As a preliminary analysis, I compared patterns of looks in possession versus effort situations, and found two significant effects: Children looked to E1 in a significantly higher number of possession than effort trials, $F(1, 60) = 13.74$, $p < .0005$, $\eta^2_p = .186$, and children also directed a greater number of looks to E1 in possession than effort trials, $F(1, 60) = 8.65$, $p = .005$, $\eta^2_p = .126$. However, these variables did not interact with condition or age group. Furthermore, average latency and duration of looks did not differ across possession versus effort situations, nor did they interact with condition or age group, all ps $> .095$. Thus, for analysis of these four dependent measures, data were collapsed across possession and effort situations.

The main analysis consisted of a MANOVA using the same four dependent measures. The fixed factors were condition (harm, neutral), age group (18 months, 2 years), and gender. There was a significant multivariate effect of condition, Wilks' Lambda = .520, $F(4, 50) = 11.53$, $p < .0005$, $\eta^2_p = .480$. Univariate tests revealed striking condition differences in all four

variables: Compared to the neutral condition, children in the harm condition looked to E1 in a significantly higher number of trials, more quickly, for longer, and more often (see Table 2).

Table 2. Means and standard deviations for various measures of children's looks in the sympathy situations

Dependent Measure	Condition		
	Neutral *M (SD)*	Harm *M (SD)*	Effect size (η^2_p)
Number of trials child looked to E1	2.13 (1.14)	3.16 (1.0)**	$\eta^2_p = .20$
Average latency to look to E1	9.67 s (2.71)	6.19 s (2.94)***	$\eta^2_p = .28$
Average duration of looks to E1	1.40 s (.54)	1.71 s (.66)***	$\eta^2_p = .29$
Average number of looks to E1	.79 (.64)	1.44 (.83)**	$\eta^2_p = .17$

Note. * $p < .005$. *** $p < .0005$.

The MANOVA also revealed a nearly significant condition x gender interaction, Wilks' Lambda = .833, $F(4, 50) = 2.50$, $p = .054$, $\eta^2_p = .167$. Univariate tests revealed that this interaction was only significant for average duration of looks, $F(1, 53) = 8.774$, $p = .005$, $\eta^2_p = .142$. Simple main effects (Bonferroni corrected) showed that girls in the harm condition looked to E1 for a significantly longer duration ($M = 2.32$ s, $SD = .75$) than did girls in the neutral condition ($M = 1.05$ s, $SD = .47$, $F(1,53) = 31.0$, $p < .001$), whereas duration of boys' looks did not differ across conditions (harm: $M = 1.93$ s, $SD = .76$; neutral: $M = 1.59$ s, $SD = .56$, $p = .48$). The MANOVA did not reveal any other significant main effects or interactions, all *p*s > .288.

2.3.2 Sympathy situations: Quality of looks

To assess the quality of looks to E1, I analyzed (1) the number of children who showed concerned and checking looks, (2) the number of the four sympathy situations in which children showed concerned and checking looks, and (3) the proportion of individuals' looks that were concerned and checking looks. The means for the second measure (the number of situations) are presented in Figure 1.

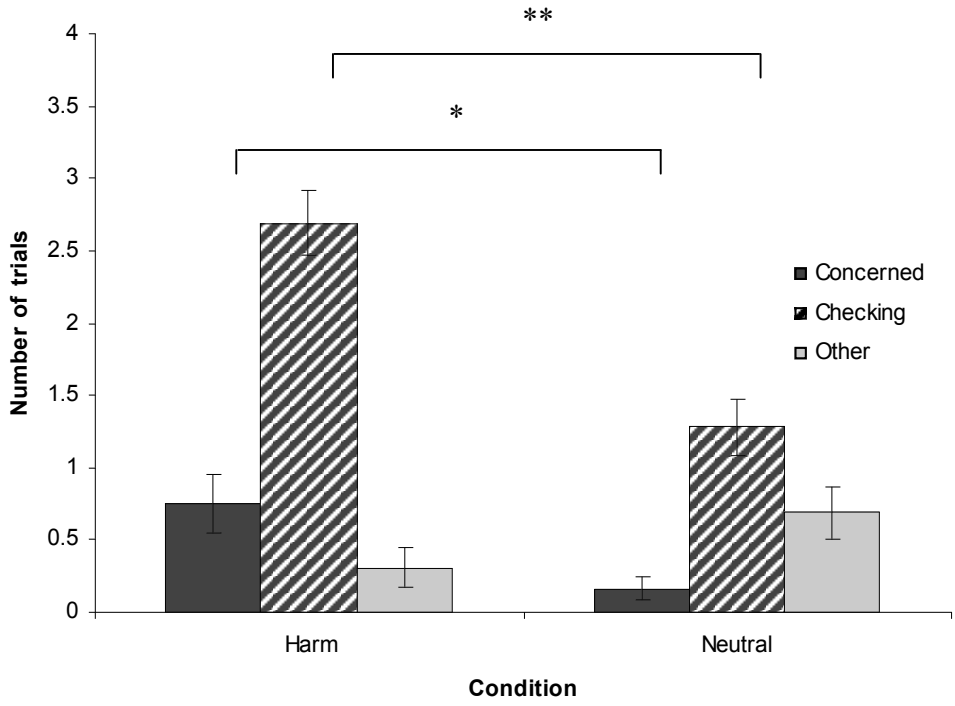

Figure 1. Mean number of trials (± *SE*) in which children showed each type of look. *$p < .05$. **$p < .0005$.

Concerned looks. As predicted, children showed more concern for E1 in the harm than in the neutral condition. Specifically, more children showed concerned looks in the harm (13 of 32, or 40.6%) than in the neutral condition (4 of 32, or 12.5%), $\chi^2(1, N = 64) = 6.49$, $p = .011$, with no difference between type of situation (possession versus effort), age group, or gender, all *p*s > .395. Children also showed concerned looks in a significantly greater number

of the four harm situations (M = .75, SD = 1.11) than the four neutral situations (M = .16, SD = .45; independent-samples $t(62) = 2.81$, p = .008).

However, this difference might be explained by the fact that children simply looked to E1 much more in the harm than in the neutral conditions. To control for this difference, for each child I calculated the number of situations in which the child showed concern as a proportion of the number of situations in which the child looked to E1. Thus, if a child looked to E1 in two of the four situations, and showed concern in one of those situations, the child received a proportion of 50. Using these proportions revealed the same result: Children showed concerned looks in a greater proportion of harm (M = 20.97, SD = 29.57) than neutral situations (M = 7.22, SD = 21.3; $t(59) = 2.09$, p = .041). Finally, analyses of the proportion of individuals' looks that were concerned looks revealed a similar albeit non-significant pattern (harm: M = 15.70, SD = 23.23; neutral: M = 6.44, SD = 20.23; $t(59) = 1.66$, p = .102). Note that three children (one in the harm condition and two in the neutral condition) were excluded from the last two analyses because they did not look to E1 in any of the four trials.

Given that children were presented with four sympathy situations in succession, it is conceivable that children's concern was primarily evident in the first few harm situations and faded with repeated presentation. However, a repeated-measures analysis indicated no significant difference across the four harm situations, p = .985: In all four situations, the proportion of looks that were concerned looks ranged between 15.56% and 18.89%. Thus, children's concerned looks did not fade across the four harm situations despite E1's lack of response.

Checking looks. As with concerned looks, and as would be expected, more children showed checking looks towards E1 in the harm (97%) than in the neutral condition (75%), p = .026 (Fisher's exact test, used due to low expected count in some cells). However, this effect was mediated by situation type and age. That is, significantly more children showed

checking looks to E1 in the possession (77%) than in the effort (55%) situations, McNemar test, $\chi^2 = 6.50$, $p = .009$. Still, in both types of situations, more children showed checking looks in the harm than in the neutral condition (possession: 28 of 32 in harm versus 21 of 32 in neutral, χ^2 $(1, N = 64) = 4.27$, $p = .039$; effort: 22 of 32 in harm versus 13 of 32 in neutral, $\chi^2(1, N = 64) = 5.11$, $p = .024$).

In addition, more 18-month-olds (97%) than 2-year-olds (75%) showed checking looks, $p = .026$ (Fisher's exact test). Analyzing the age groups separately revealed that among the 2-year-olds, whereas 15 of 16 (94%) showed checking looks in the harm condition, only nine of 16 (56%) did so in the neutral condition, $p = .037$ (Fisher's exact test). There was no difference between conditions for the 18-month-olds (16 of 16 in the harm condition and 15 of 16 in the neutral condition showed checking looks, $p = 1.000$, Fisher's exact test). The number of children who showed checking looks did not differ by gender, $p = .148$.

Children in the harm condition also showed checking looks in a significantly greater number of the four sympathy situations ($M = 2.69$, $SD = 1.23$) than did children in the neutral condition ($M = 1.28$, $SD = 1.05$, $t(62) = 4.91$, $p < .0005$). To control for the baseline difference in amount of looking to E1 across conditions, I again calculated proportion scores (i.e., the number of situations in which a child showed checking looks as a proportion of the number of situations in which the child looked to E1). This more conservative measure revealed the same result (harm: $M = 88.17$, $SD = 21.17$; neutral: $M = 64.72$, $SD = 40.69$; $t(59) = 2.84$, $p = .007$). Finally, a similar pattern emerged in the proportion of individuals' looks that were checking looks (harm: $M = 79.22$, $SD = 25.53$; neutral: $M = 66.54$, $SD = 39.43$), but this difference was not significant, $t(59) = 1.50$, $p = .14$. Again, the three children who did not look to E1 in any of the four trials were excluded from the last two analyses.

2.3.3 The effect of condition on subsequent prosocial behavior

The distribution of children's prosocial scores across condition, age group, and gender are presented in Table 3. As expected, significantly more children helped or shared with E1 (i.e., received a prosocial score of 3) if they had previously experienced the harm rather than the neutral condition (harm: 21 of 32, or 65.6%; neutral: 12 of 32, or 37.5%; χ^2 (1, N = 64) = 5.07, p = .024). The number of children who helped or shared did not differ by age group or gender, both ps = .802.

Table 3. Percent of children who received each score as their highest prosocial score

		Condition previously experienced		Age group		Gender	
Category	Score	Harm	Neutral	18 months	2 years	Females	Males
Helps/shares	3	65.6	37.5	53.1	50.0	53.1	50.0
Shows distress or Describes situation	2	15.6	25.0	15.6	25.0	25.0	15.6
Attends to situation	1	18.8	37.5	31.3	25.0	21.9	34.4
Ignores situation	0	0	0	0	0	0	0

An additional analysis of the effect of condition on prosocial behavior consisted of a univariate ANOVA using prosocial scores as the dependent measure and condition, age group, and gender as fixed factors. This revealed a main effect of condition, $F(1, 56) = 5.16$, p = .027, η^2_p = .084: Children who had previously seen E1 in the harm condition had higher prosocial scores towards her (M = 2.47, SD = .80) than did children who had seen her in the

neutral condition (M = 2.00, SD = .88). The ANOVA also revealed a nearly significant age group x gender interaction, $F(1,56) = 3.88$, $p = .054$, but simple main effects (Bonferroni corrected) revealed no significant gender differences in prosocial scores in either age group, both ps > .117. There were no further main effects or interactions, all ps > .100.

2.3.4 Relations between children's concerned looks and prosocial behavior

To assess the association between degree of concern in the sympathy situations and subsequent prosocial behavior, I conducted non-parametric correlations using the number of the four sympathy situations in which children showed concern and children's prosocial scores. As predicted, the two factors were positively correlated, Kendall's tau = .24, $p = .036$. This correlation was specific to concerned looks; a similar analysis conducted using children's checking looks was not significant, $p = .514$.

Since prosocial scores varied by condition (see above), the correlations between concerned looks and prosocial scores were also conducted separately for each condition. As predicted, in the harm condition, the correlation between number of situations with concerned looks and prosocial scores was positive, Kendall's tau = .26, although this was a non-significant trend, $p = .097$. In the neutral condition, the number of situations with concerned looks was not associated with prosocial scores, Kendall's tau = .015, $p = .928$. Note that correlational analyses conducted using the more conservative measure of proportion of situations (i.e., number of situations in which a child showed concerned looks divided by the number of situations in which the child looked to E1) revealed very similar results, as did correlational analyses using proportions of individuals' looks that involved concern.

One possible alternative interpretation of this correlation is that what I coded as concern actually indexed emotional arousal caused by the perpetrator's aggressive behavior in the harm condition (since, in order for the conditions to be believable, the perpetrator did behave mildly aggressively in the harm but not in the neutral condition). Furthermore,

perhaps those children who experienced this emotional arousal were then more susceptible to the victim's distress cues in the prosocial situation. This is potentially problematic given that children received a higher prosocial score for showing distress (a score of 2) than for only attending to the situation (1) or showing no response (0). Thus, perhaps the increased emotional arousal during sympathy situations and the resulting increased distress in the prosocial situation created a spurious correlation that does not index a sympathy-prosocial behavior link at all. However, this alternative interpretation of the correlation does not hold since, even when showing distress was excluded from the coding scheme and those three children who showed distress were assigned their next highest score (all three received a 1 for attending to the situation), the correlation between the number of trials in which children showed concerned looks and children's prosocial behavior persisted, Kendall's tau = .23, p = .045.

2.4 Discussion

I examined whether young children can, even in the absence of emotional cues, sympathize with a victim. Extending Hobson et al.'s (2009) study, I tested significantly younger children in multiple situations, and, in addition, examined the relation between children's sympathy and prosocial behavior. I found that by 18 months of age, children show concern for an adult stranger who is in a hurtful situation but shows no emotion. What is striking about these results is not that such young children showed sympathy, which was to be expected given past work (e.g., Bischof-Köhler, 1991; S. K. Young et al., 1999; Zahn-Waxler, Radke-Yarrow et al., 1992); what is striking is that this is, to my knowledge, the first demonstration that such young children can sympathize with a sufferer even in the absence of overt emotional cues. This study thus also extends past work on sympathy in toddlers, which had, up to this point, mostly focused on children's empathy and sympathy in response to a sufferer's overt emotional signals.

The claim that children were concerned for the sufferer (rather than, say, about the generally negative situation or the victim's potentially angry response) gains support from two additional findings. First, children in the harm condition later helped E1 significantly more than did children in the neutral condition. My interpretation of this finding is that observing someone experiencing negative situations increases the likelihood of children helping that person, presumably by inducing sympathy, which has been both theoretically and empirically linked to prosocial behavior (see, e.g., Batson, 1991; Eisenberg & Miller, 1987). This proposal is strengthened by a second finding: the correlation between children's concerned looks and their subsequent prosocial behavior towards the sufferer, which indicates that individual children who expressed concern for E1 were also more likely to help E1. Together, these findings substantiate the claim that I have measured sympathy and support my conclusion that the early ability to sympathize does not require overt emotional cues: In the absence of such cues, children can use situational cues to sympathize with another person.

One open question concerns the mechanism(s) children employed to arrive at sympathy. Obviously, sympathy in the present study did not result directly from exposure to the victim's emotional cues (e.g., via mechanisms such as mimicking the emotional cues, emotional contagion, etc.), as such cues were not provided. I thus propose that sympathy in this study resulted at least partially from cognitive processes[4]. Several cognitive processes can contribute to empathy-related responses (see Eisenberg et al., 2006; Feshbach, 1978; Hoffman, 1982, 1984). Simpler processes include direct association (e.g., seeing another's blood elicits distress in the observer due to blood being linked to the observer's own past distress) and classical conditioning. However, sympathy results from more sophisticated

[4] My claim is not that cognitive processes are entirely distinct and separable from affective processes. On the contrary, my claim would be that cognitive and affective processes are interdependent such that a cognitive construal of someone else's situation or state can arouse an affective response (such as the sympathy aroused in this study), affective appraisal can give rise to cognitive construal, and that the two work closely together to jointly give rise to behavior (see Pessoa, 2008).

processes that involve an analysis of the source of the vicarious feeling and therefore a focus on the other (Eisenberg et al., 1991).

One such sophisticated cognitive process is affective perspective-taking, that is, making inferences about the other's affective state by putting oneself in the other's place and basing one's responses on those inferences (Eisenberg et al., 1991; Hoffman, 1984). In the absence of emotional cues, one way to make this inference is via simulation, which involves imagining oneself in another's situation (e.g., Decety & Sommerville, 2003; Harris, 1995). An alternative but related possibility (and the one preferred by Hobson et al., 2009) is that the observer can feel her way into the experience of and feel for the other person because she identifies with that person's attitudes. According to Hobson et al., in their study, children with autism did not identify with the victim's attitudes and could therefore not experience concern for the way the victim would be expected to feel, whereas children with learning disabilities and typically-developing children did not have difficulties with identification and could thus experience concern for the victim. Importantly, whether via simulation, identification, or some other mechanism(s), one eventually takes the other's perspective and apprehends the other's affective state, which can activate affective responses such as sympathy and can thereby motivate prosocial behavior (Batson, Fultz, & Schoenrade, 1987; Feshbach, 1978; Krebs & Russell, 1981). Plausibly, then, in the present study, children apprehended the victim's state by taking her affective perspective, which motivated their sympathy and prosocial behavior.

This finding might be surprising given that thus far, affective perspective-taking has only been demonstrated in the third year and beyond (Denham, 1986; Wellman et al., 2000). However, tasks used in prior work required children to display quite sophisticated cognitive and linguistic skills, such as comprehending hypothetical stories and answering questions about their own feelings. These skills might not amply develop until the third year. It is thus

possible that children younger than 3 years of age possess some affective perspective-taking abilities but the methods used in prior work have not been sensitive enough to tap these abilities. Relevant here is recent work on children's theory of mind, in which the use of sensitive, implicit measures shows that, rather than emerging around 4 years of age, as previously believed, a basic theory of mind is already present during the second year (e.g., Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, 2007). Thus, when appropriate measures are used, children in their second year could well demonstrate some affective perspective-taking skills as well.

Depending on how familiar children were with situations like the current sympathy situations, they might additionally have relied on their past experiences to infer the victim's affect. That is, if children had previously directly or vicariously experienced such situations on multiple occasions, perhaps they had formed scripts about people's responses to such situations and, in this study, were partially relying on these scripts to infer the victim's affect. On the other hand, if the situations were novel for children, then children likely engaged in perspective-taking (see Blair, 2005; Eisenberg et al., 1991; Karniol, 1982). It is possible that the sympathy situations were somewhat familiar to children, especially to those with siblings and those in day-care. Thus, perhaps some children in the harm condition (those familiar with such situations) relied less on affective perspective-taking and more on scripts than did children who were unfamiliar with such situations. However, even if the situations were to some degree familiar to children, it is highly unlikely that children had ever witnessed precisely the situations that they witnessed in the present study (e.g., an adult tearing up another adult's drawing), and so although they might have had some scripts to rely on, they also had to engage in some affective perspective-taking. In any case, children did sympathize, indicating that they can arrive at sympathy without relying on perceptible emotional signals.

It is noteworthy that in the harm condition, only some children (40%) showed concerned looks (although this proportion is similar to proportions reported in studies in which the victim provided emotional signals; e.g., Zahn-Waxler, Radke-Yarrow et al., 1992). One possible explanation for this might be that the degree of sympathy aroused is related to the level of observer-sufferer attachment (Batson, 1987). As E1 was a relative stranger, fewer children may have experienced sympathy than would have if the sufferer had been their parent (van der Mark et al., 2002; S. K. Young et al., 1999). There are also likely differences in individuals' tendency to outwardly express sympathy. Thus, more children might have experienced sympathy than expressed it facially. Indeed, given that all but one child in the harm condition showed checking looks, perhaps some checking looks were in fact sympathetic looks but without the accompanying overt expressions. However, concerned but not checking looks correlated with prosocial behavior, indicating that the two kinds of looks tapped into distinct responses, and that checking looks were not simply sympathetic looks without the overt expressions.

A related possibility for why more children did not show concern might have to do with the fact that concerned looks were only coded as such when they were directed at the victim. The measure of concern was thus quite conservative, and perhaps some children experienced concern for the victim but were not coded as doing so because they did not meet the conservative criterion. One way to get around this problem in the future might be to use physiological measures, such as heart rate and skin conductance, which are less vulnerable to such coding decisions (see, e.g., Eisenberg & Fabes, 1990, 1998; Hastings, Zahn-Waxler, & McShane, 2006). For instance, heart rate deceleration co-occurring with the processing of important information about another's emotional state is thought to tap into the experience of sympathy, and to distinguish sympathy from personal distress, which is associated with heart rate acceleration (Eisenberg et al., 2006; Hastings et al., 2006). Using such measures may

well reveal that a greater proportion of children experience sympathy or empathy than the proportion revealed in the present study. Thus, it is possible that by identifying only the outwardly expressive children as experiencing sympathy or empathy, I under-estimated the proportion of children in the second year who can sympathize or empathize in the absence of distress cues.

Even so, I found a correlation between sympathy and prosocial behavior, the strength of which is comparable to some prior work in which the victim presented emotional cues (e.g., .20 in Zahn-Waxler, Robinson et al., 1992), and is consistent with the general finding that the relation between sympathy and prosocial behavior exists but is not very strong (see Eisenberg & Miller, 1987; Eisenberg et al., 2006). This correlation could represent a causal link such that sympathy for a person leads to prosocial behavior towards that person, but it could also be due to a third variable such as temperament or emotion regulation (see Batson, 1998; Eisenberg & Miller, 1987; Hoffman, 1976, 1982). For instance, S. K. Young et al. (1999) found that inhibited children show less prosocial behavior and less empathy towards an unfamiliar experimenter (see also Eisenberg, 2005; Radke-Yarrow & Zahn-Waxler, 1984; van der Mark et al., 2002). A similar factor might also partially explain the correlation in the present study. Along similar lines, it could be argued that the correlation between children's concern and their subsequent prosocial behavior was actually a spurious correlation between the increased emotional arousal during sympathy situations and the resulting increased distress in the prosocial situation. However, this alternative interpretation of the correlation does not hold since the correlation emerged even when I excluded 'shows distress' from the coding scheme. Still, all in all, there are reasons to believe that there are dispositional as well as situational relations between sympathy and prosocial behavior (see Eisenberg, 2005); it is thus plausible that both factors were responsible for the present correlation.

An interesting aspect of these results was that four of the 32 children in the neutral condition showed concern for E1. At first glance, this seems strange considering that E1 was in no way affected by E2's actions in this condition, but my sense during testing was that some children nevertheless worried that E2's behavior might be threatening to E1. For instance, after E1 had just finished drawing a picture, E2 tore up a blank piece of paper for no reason, and perhaps some children perceived this as a threat to E1's drawing, which was lying within easy reach of E2. My aim in designing the neutral condition was to make it as similar as possible to the harm condition. This might, however, have led some children in the neutral condition to interpret E2's actions as negative for E1.

It is worth mentioning that I found almost no gender differences in the dependent measures and only one age difference in checking looks. Importantly, there were no age differences in children's show of concern, which suggests that the ability to sympathize without overt emotional cues from the sufferer is present by 18 months. On the one hand, this is striking considering the kinds of cognitive and affective experiences and abilities that are likely needed to sympathize in the absence of emotional signals. On the other hand, it is unsurprising given that even 14-month-olds have been shown to sympathize when a sufferer displays emotions (e.g., Zahn-Waxler, Radke-Yarrow et al., 1992). Future work could further simplify the current sympathy situations to test whether 14-month-olds also show concern without the aid of the victim's emotional signals.

A caveat about the present findings is that children might have been influenced by their parents. Parents were instructed not to provide cues and they were generally good at following these instructions. Still, future work might have parents sit to the children's side to prevent them from potentially subtly influencing their children's responses. A more fundamental caveat concerns the generalizability of the findings. What, for example, is the range of situations within which young children can sympathize without emotion reading? I

found that children sympathized in two kinds of situations: possession and effort. These categories cover many of the situations that children experience regularly. However, it is implausible that young children would sympathize in entirely novel categories of situations that they had no way to understand and in which they had no affective cues to guide them (e.g., hearing that someone did not get the job that he wanted). Clearly, the ability to sympathize, especially in the absence of emotional cues from the sufferer, rests on one's knowledge and understanding of the world, both of which develop with age.

There might also be cultural variation in the kinds of situations that elicit sympathy. The possession situations in the present study, for instance, might not elicit sympathy in a culture in which belongings tend to be shared and not to be the sole property of one person. This is a fascinating question that should be examined in future work. Finally, sympathy might vary depending on children's attachment to the victim or the victim's gender and race (see, e.g., Eisenberg & Lennon, 1983; S. K. Young et al., 1999). Thus, the current findings certainly do not generalize to all situations and all cultures. How and why children's sympathy varies are fascinating questions that deserve much more attention.

In sum, even 18-month-old children can sympathize with someone who is in a negative situation but shows no affective cues. Moreover, the sympathy thus experienced follows the patterns that true sympathy is expected to follow: It increases the likelihood of prosocial behavior and, within individuals, it correlates with prosocial behavior. These findings show that humans feel for and help people who are in hurtful situations, and we do so robustly, flexibly, and in multi-determined ways from very early in development.

However, morality is concerned not only with victims but also with actors, both harmful and helpful. If victims elicit empathy, sympathy, and prosocial behavior, then harmful actors (transgressors) should, in some sense, elicit the opposite of such responses. This might entail actively punishing transgressors or more passively withholding cooperation

from them, shunning or socially ostracizing them, and so on. On the other hand, moral, cooperative actors may be rewarded with increased cooperation. Both the punishment of transgressors and the rewarding of prosocial actors in third-party interactions are ways of maintaining cooperation in large groups, and adults seem to engage both mechanisms (although as mentioned before, punishment or the threat thereof seems to be more salient than reward as an incentive for cooperation). In the next chapter, I consider whether young children also engage these mechanisms. Specifically, I assess whether young children's prosocial behavior varies towards others based on the others' moral behaviors. In Study 2a, I ask whether 3-year-old children reduce their prosocial behavior towards a harmful actor and/or increase their prosocial behavior towards a helpful actor. In Study 2b, I explore whether 3-year-old children's prosocial behavior towards a transgressor varies depending on the intentions of the transgressor.

3. STUDY 2: SELECTIVITY IN PROSOCIAL BEHAVIOR

3.1 Introduction

One of the most important ways to enforce moral norms and maintain morality, especially in large-scale human societies, is to monitor third-party moral interactions and to take action by punishing moral norm violators and rewarding good, moral individuals (Boyd & Richerson, 2005; Fehr & Fischbacher, 2003, 2004). It has been proposed that from an evolutionary perspective, humans may well be attuned to two different social strategies of prospective partners: a cheating strategy and a cooperative strategy, and that selection of future partners could then be based on this categorization (Harris, Núñez, & Brett, 2001). Indeed, as adults, we do punish harmful individuals and we may help and affiliate with helpful individuals (Fehr & Gächter, 2002; Krebs, 2008). Often, the currency of this punishment and reward is future cooperative behavior; thus, adults often punish norm violators by withholding cooperation from them in the future (what Boyd & Richerson, 1992, call "retribution," since it is punishment directed only at the transgressor rather than at the whole group). To date, almost no work has examined whether young children also display such selectivity in their prosocial behavior. This was the purpose of the present set of studies (Studies 2a and 2b).

Young children and even infants are highly prosocial: They help others in simple, instrumental ways, they comfort those in pain or distress, they provide information when others need it, and they even spontaneously share their own resources (see Warneken & Tomasello, 2009b). It has been suggested that this early prosociality is a way to engage positively with other people and is likely indiscriminate in its targets (Hay & Cook, 2007; Warneken & Tomasello, 2009a). However, as toddlers develop preferences, become more sensitive to context, and learn the moral and social norms surrounding prosocial behavior, their early prosocial impulses are thought to become transformed into more deliberate,

selective, and morally informed choices (Hay & Cook, 2007). Indeed, children's prosocial tendencies have been found to be discriminating by the third year. For instance, 3-year-olds were more likely to share toys with a peer if that peer had previously shared toys with them, suggesting a sensitivity to reciprocity by this age (Levitt et al., 1985). There is also some evidence that by about 4 years of age, children share with another child (even at a cost to themselves) if the other child is their friend rather than a non-friend or a stranger (Birch & Billman, 1986; Moore, 2009). Thus, very young children's indiscriminate prosocial behavior seems to become selective by around 3 years of age, at least in dyadic interactions. It remains an open question, however, whether this selectivity is also agent-neutral, that is, whether it is also apparent in the context of third-party interactions.

Infants and young children do seem to at least monitor and evaluate third-party moral interactions. For instance, as early as 6 months of age, infants are able to discriminate a helpful from a hindering character, and they seem to show a social preference for the helpful over a neutral character, and for the neutral over the hindering character (Hamlin et al., 2007). Moreover, children do show sensitivity to reciprocity in interactions that resemble third-party interactions. Specifically, Olson and Spelke (2008) asked 3.5-year-old children to help a protagonist doll decide how to allocate resources to other dolls. Children allocated more of the protagonist doll's resources to a doll who was generous to a different (not the protagonist) doll than to a non-generous doll, thus demonstrating an appreciation of indirect reciprocity, that is, reciprocity in interactions that did not directly impact the children (or the doll on whose behalf they were acting).

Olson and Spelke's (2008) results could also be interpreted as showing that young children consider others' moral behaviors in third-party interactions when deciding whom to share with. However, sophisticated moral judgments take into account several factors, including not only the consequences of but also the intentions behind a behavior (see

Cushman, 2008; Piaget, 1932/1997; Turiel, 2006; Turiel et al., 1987; Weiner & Peter, 1973). Thus, to show that children's prosocial behavior is mediated by others' moral behavior, one needs to show that the prosocial behavior is mediated not only by the outcomes of but also the intentions behind others' actions. In this chapter, I present work assessing whether young children show differential prosocial behavior towards agents who vary in the helpful or harmful intentions underlying their behavior.

A secondary issue, also raised by Olson and Spelke (2008), is that children's responses on behalf of others (the protagonist doll in their study) may not correspond to children's responses when they themselves must decide whom to share with or help. Study 2a thus tested whether children respond differentially to others' harmful or helpful behaviors even when they themselves are the ones sharing or helping. The primary question of Study 2b was whether intentionality mediates children's judgments of others as worthy recipients of prosocial acts.

Children in both studies (2a and 2b) were 3 years old because by that age, children's prosocial behavior has been shown to be selective (Levitt et al., 1985; Olson & Spelke, 2008). Moreover, in prior work on children's moral judgments, when transgressors' intentions and the consequences of their actions were pitted against each other, the understanding and use of the transgressors' intentions were only demonstrated in children 5 years of age and older (see Chapter 1). Although several studies have included children as young as 3 years of age, none have been able to demonstrate sensitivity to transgressors' intentions at this young age (e.g., Wellman et al., 1979). This thus seemed like a reasonable age to assess this sensitivity using the non-verbal, non-hypothetical measure of prosocial behavior.

The general procedure of both studies was as follows: In a between-subjects design, after warming up with three adults, the children viewed four familiarization trials (very

similar to the sympathy situations of Study 1). During these trials, one adult (the actor) behaved in a condition-specific way (e.g., harmfully or helpfully) towards another adult (the recipient). These four familiarization trials were followed by one helping test in which children were presented with a forced choice between helping the actor or the third (familiar but neutral) adult. The dependent measure in both studies was thus instrumental helping, which is relatively easy to elicit without explicit verbal instructions (Warneken & Tomasello, 2006, 2007). A forced choice was included because young children are highly motivated to help (see Warneken & Tomasello, 2009a) and might help a harmful actor if that were their only option; the forced-choice paradigm gave them the option of not helping the actor. Finally, once children had helped one person in the helping test, they were given the opportunity to help the other person if they so chose.

3.2 Study 2a

In Study 1, children saw the actor harming, helping, or, as a baseline measure, behaving neutrally towards the recipient. I predicted that children's subsequent helping would be mediated by the actor's behaviors towards the recipient.

3.2.1 Method

Participants

Participants were 3-year-olds (N = 54; 9 girls and 9 boys per condition) between 35 months, 21 days and 38 months, 17 days (M = 37 months, 11 days; SD = 22.46 days) from a medium-sized German city. Fifteen additional children were tested but excluded due to indecisiveness or unclear responses during the helping test (6; see Procedure section below for a description of a clear helping response), fussiness or inattentiveness (4), parents not following instructions (4), and experimenter error (1). For Studies 2a and 2b, the recruitment, language, ethnicity, socioeconomic status, education, and occupation information was

identical to Study 1. Three adult female experimenters played the same roles (of recipient, actor, and neutral person) for all children.

Materials and Setting

The set-up is shown in Figure 2. Each child saw four familiarization trials during which the child, recipient, and actor sat around a table; the neutral person sat to the side, visible to the child, reading a magazine. The following materials were used during familiarization trials: two necklaces with colorful beads, two belts with colorful beads, blank sheets of paper and a color pencil, and a blue and a red ball of clay in a container.

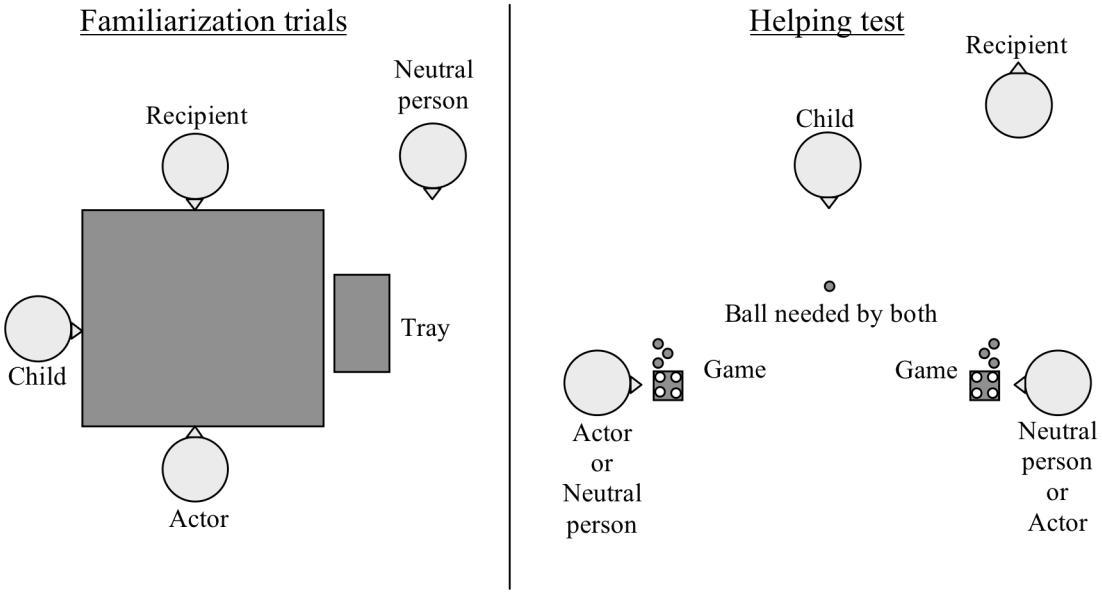

Figure 2. Experimental set-up for all conditions (Studies 2a and 2b).

In the subsequent helping test, two identical color-matching games were used, which consisted of a box with four holes marked by different colors into which balls of matching colors could be placed. The child stood at a pre-determined location and the color-matching games' boxes were placed 2 m from and on either side of the child (see Figure 1). Next to each box lay three of the four balls required for each game; a blue ball was missing from both sides. One blue ball lay in the middle, 1 m from the child. The parents sat directly behind the child, the actor and neutral person sat next to their game (side counterbalanced across

children), and the recipient stood to the side of and behind the child, facing away from the interaction, and kept time.

Procedure

After warming up with the children, the recipient introduced the actor and neutral person to children as her friends. The actor and neutral person then warmed up with children, being careful to interact equally and similarly with them. Throughout, the recipient wore one of the necklaces and belts described above. After about 10 min in the warm-up room, the recipient told children that she would like to show them the toys in a different room, and escorted the children and parents to the testing room, followed by the actor and the neutral person.

The study was between-subjects with three conditions (*Harm*, *Help*, and *Baseline*). All conditions had four familiarization trials followed by one helping test. Each familiarization trial (adapted from Study 1; see Chapter 2) began with a 45-s presentation in which the recipient presented one of the following objects (order counterbalanced):

Necklace and *Belt*. The recipient admired and showed off her necklace/belt. A second necklace/belt lay on a tray, visible but inaccessible to children.

Picture. On one sheet of paper from the stack, the recipient drew a picture, proudly commenting on how pretty it was and how happy it made her.

Clay. Using one ball of clay, the recipient made a bird, commenting as in the Picture trial. The second clay ball remained in the container.

These 45-s presentations ended differently depending on the condition:

In the Harm condition, the recipient placed her possession on the table and the actor said in a mildly aggressive tone, “I’m going to take/tear/break this now,” and put on the necklace or belt, or tore up the picture or bird and threw the bits into a bin. The actor did not display aggression in her facial expression (which was neutral while she spoke and during her

actions), nor in any other way before or during her actions. The recipient watched the actor sadly but silently. After 15 s, the recipient assumed a neutral expression and began the next demonstration.

In the Help condition, in contrast, the recipient was accident-prone: Instead of placing her possession on the table, she accidentally dropped her necklace, detached beads from her belt, tore her picture, or damaged her clay bird, and was sad about each mishap. The actor said sympathetically, "I'll get/fix it," then retrieved/repaired the object. The recipient watched the actor sadly but silently (as in the Harm condition). After 15 s, the actor placed the object on the table and smiled, and the recipient happily took it, put it aside, and neutrally began the next demonstration.

The Baseline condition established the appropriate chance level for analyses of the other conditions and thus involved the actor behaving neutrally. After Necklace and Belt demonstrations, the recipient told children that the necklace/belt on the tray was not hers. The actor then commented on the necklace/belt on the tray ("The necklace has so many beads" or "This belt can be adjusted here") and counted the beads or adjusted the belt. After Picture and Clay demonstrations, the actor said, "These [remaining] sheets are all blank" or "This [second ball of] clay is stuck to the container," and proceeded to examine the paper or un-stick the clay. The actor's comments and expression were neutral and the recipient watched the actor neutrally. After 15 s, the actor put the object away and the recipient began the next demonstration.

In each condition, the familiarization trials were followed by a helping test in which the actor and neutral person simultaneously but individually played their respective color-matching game. After placing the three available balls in their slots, both simultaneously reached for the ball in the middle and maintained this reach, looking only at the ball. If children did not act within 15 s, the recipient silently cued parents to ask children to "give the

ball." If this was ineffective, parents asked children to point to the person they wanted to give the ball to. As a last resort, the recipient sat between the actor and neutral person, held up the ball, and asked children to give it or point to the person they wanted to give it to. Whomever children chose gratefully accepted the ball. Only placing the ball in or near one person's hand or pointing to one person were considered clear responses. If children did not respond within 1 min, the test was ended.

After the helping test, children received a second ball to hand to whomever they chose. This helped resolve the situation by giving children the chance to help both individuals. Finally, in the Harm condition, the actor made amends (e.g., by returning the recipient's necklace and belt) and apologized, and the recipient accepted the apology.

Coding and Reliability

During testing, the experimenter playing the actor coded whom children chose to give the ball to. A second coder (blind to condition and hypotheses) coded this measure from videotapes for a random 50% of children (9 per condition). Agreement was perfect, $\kappa = 1.0$. For the same 50%, before coding giving, the second coder judged (using relevant cues such as one person reaching slightly further than the other) whether either person was likelier to receive the ball. She judged the actor as likelier in seven cases and the neutral person in five cases, but these numbers were unrelated to condition and to who received the ball (Fisher's exact tests, $p = .47$ and $p = 1.00$, respectively).

3.2.2 Results

Since preliminary analyses revealed no gender effects, gender was not included in further analyses. All reported *p*-values are two-tailed. In the Baseline condition, 12 of 18 children (67%) helped the actor; 67% was thus the appropriate test proportion for the other conditions. Following Hamlin et al. (2007), I used binomial tests to analyze the experimental conditions. In the Harm condition, a significantly lower proportion of children than 67%

helped the actor (4 of 18, or 22%; binomial probability, $p < .0005$). This difference did not emerge in the Help condition (11 of 18 children, or 61%, helped the actor; binomial probability, $p = .760$; see the first three bars of Figure 3).

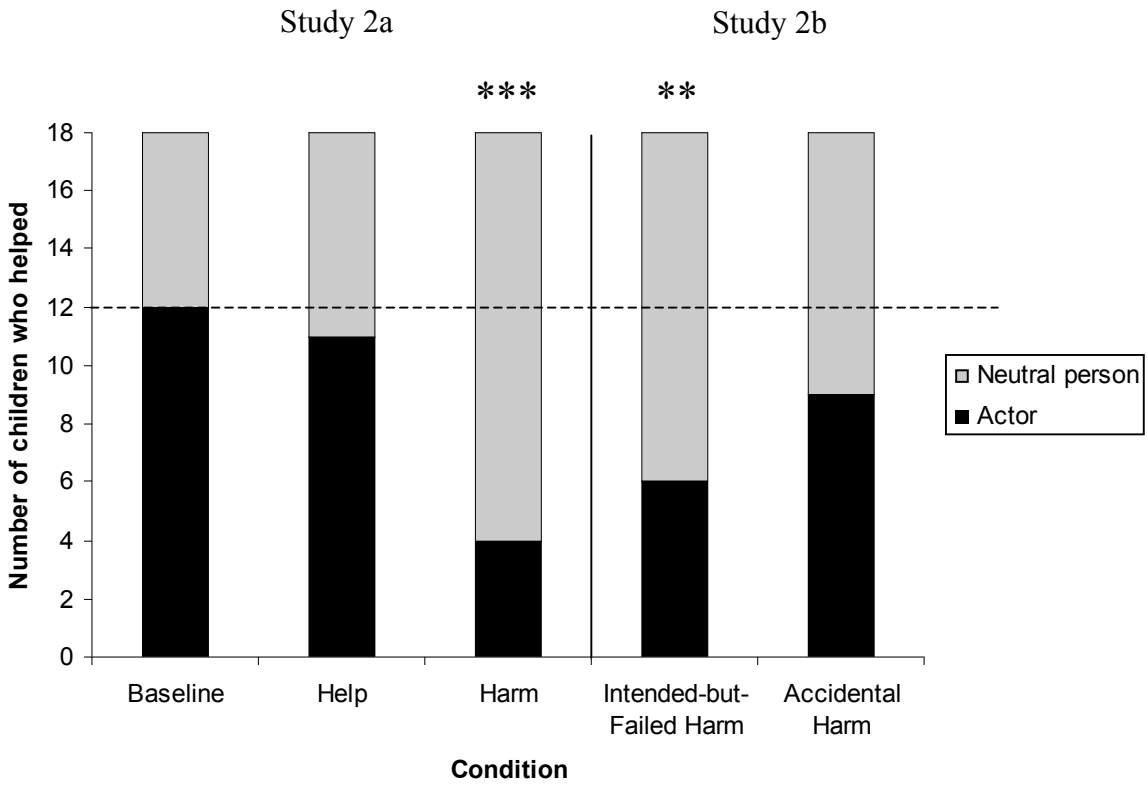

Figure 3. Number of children who helped the actor or the neutral person in Studies 2a and 2b. The dashed line represents chance (67%) as established by the Baseline condition. $^{**}p < .01$. $^{***}p < .0005$.

Secondary analyses using chi-square tests revealed that the Harm condition differed significantly from the Baseline condition ($\chi^2[1, N = 36] = 7.20, p = .007$) and from the Help condition ($\chi^2[1, N = 36] = 5.60, p = .018$). The Help and Baseline conditions did not differ, $p = .729$.

An alternative explanation for the results of the Harm condition is that children refrained from helping the actor because they were afraid of her (e.g., due to her mildly aggressive tone of voice). However, the phase in which children received a second ball suggests otherwise: All 14 children who had first helped the neutral person gave the second

ball to the actor, indicating they were not afraid of her but, when forced to choose during the helping test, chose not to help her.

3.2.3 Discussion: Study 2a

This study showed that 3-year-olds take into account others' harmful behaviors towards third parties when deciding whether or not to help them. This extends Olson and Spelke's (2008) findings to children's own prosocial behavior, and suggests that children select whom to help and selectively direct prosocial behavior away from harmful people. Thus, by 3 years of age, children show discriminate cooperative behavior even in the context of third-party interactions, i.e., even when the harmful behavior did not directly impact them. Young children thus monitor and take action against harmful actors.

Interestingly, the actor's harmful but not her helpful actions impacted children's behavior. This suggests a negativity bias, that is, a greater impact of negative than of positive information. This bias has been extensively documented in adults (e.g., Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001) and, in a recent review paper, some colleagues and I documented it in children's social-emotional development as well (Vaish, Grossmann, & Woodward, 2008). For instance, we found that most studies on infants' social referencing behavior suggest a far greater potency of negative information (e.g., fearful facial and vocal signals) to deter infants from exploring a novel object than of positive information (e.g., happy facial and vocal signals) to encourage infants to explore a novel object (e.g., Hornik, Risenhoover, & Gunnar, 1987; Mumme et al., 1996). The findings of Study 2a suggest a similar bias in children's moral development: Children's prosocial behavior is decreased towards a harmful individual but not increased towards a helpful individual. This result is also consistent with the research on children's moral judgments that indicates that children correctly identify "bad" acts substantially earlier than "good" acts (Hill & Hill, 1977; Rhine, Hill, & Wandruff, 1967; see Karniol, 1978, for a discussion). Future studies should further

explore the origins and implications of the negativity bias in moral development (see Aloise, 1993; Leslie, Knobe, & Cohen, 2006, for relevant work).

Notably, although most children in the Harm condition helped the neutral person first, they handed a second ball to the actor. This concords with Olson and Spelke's (2008) finding that although 3.5-year-old children allocated resources selectively when resources were limited, when resources were sufficient, children divided them equally among recipients, displaying a sense of fairness. When a second ball was available, children in the present study also displayed fairness, or at the very least, they were willing to help the actor when she was the only one who needed help.

In Study 1, young children helped a harmful individual less, but whether this was in response to the individual's harmful intentions or to the negative outcomes she caused remains unclear because the Harm condition featured both intentional harm and negative consequences. The critical question of why children helped a harmful actor less was addressed in Study 2b.

3.3 Study 2b

The intention-outcome distinction is essential to the study of moral development, given that sophisticated moral judgments are thought to rely on not only the consequences of but also the intentions behind others' behavior (Cushman, 2008; Karniol, 1978; Piaget, 1932/1997; Turiel et al., 1987). As discussed previously (Chapter 1), developmental work suggests that when intentions and outcomes are pitted against each other, children around 5 years of age and above reliably use the perpetrator's intentions when making moral judgments and assigning punishment; prior to this, they rely largely on outcome information (D. T. Miller & McCann, 1979; Wellman et al., 1979; Zelazo et al., 1996). Much of this past work has assessed children's verbal evaluations of hypothetical moral transgressions, which may not be optimal for very young children whose language skills are limited, and which do

not always correspond to children's actual behavior (Astington, 2004; Darley & Shultz, 1990; Wainryb, Brehl, & Matwin, 2005). I thus asked in this study whether young children's ability to respond to the intentions behind moral transgressions might be evident in their prosocial behavior.

In Study 2b, I teased apart intentions and outcomes using two new conditions: In the *Intended-but-Failed Harm* condition, the actor intended but was unable to harm the recipient (non-negative outcome, negative intention), and in the *Accidental Harm* condition, the actor accidentally harmed the recipient (negative outcome, non-negative intention). If the intentions behind others' harmful actions mediate children's prosocial behavior, children should subsequently help the actor less than the neutral person in the Intended-but-Failed Harm condition but not in the Accidental Harm condition.

3.3.1 Method

Participants

Participants were 3-year-olds (N = 36; 9 girls and 9 boys per condition) between 35 months, 18 days and 38 months, 15 days (M = 37 months, 7 days; SD = 24.93 days). Seven additional children were tested but excluded due to indecisiveness or unclear responses during the helping test (3), fussiness or inattentiveness (2), parent not following instructions (1), and experimenter error (1). The duration of sessions was the same as in Study 2a. The same experimenters played the same roles as in Study 2a.

Materials and Setting

The materials were the same as in Study 2a with minor changes. In the Intended-but-Failed Harm condition, the sheets of paper had a border of transparent tape such that they could not easily be torn. In the Accidental Harm condition, a set of beads was wrapped around the belt such that it could easily fall off, and the sheets of paper had an inconspicuous rip such that they could easily tear. The setting was identical to Study 2a.

Procedure

The warm-up session was identical to Study 2a. The two between-subjects conditions of Study 2b again involved four familiarization trials and a helping test. The recipient's presentations during familiarization trials were similar to Study 2a except that she decorated her clay bird with beads. The four 45-s presentations ended with the recipient placing her possession on the table. In the Intended-but-Failed Harm condition (like the Harm condition), the actor then said mildly aggressively, "I'm going to tear/break this now," and attempted (with obvious effort) to tear up the picture, dislodge the beads from the clay bird, or break the necklace or belt. She did not display aggression facially or in any other way before or during her actions. Importantly, she was unable to cause harm as she could not tear through the tape around the paper, the beads were lodged too deeply into the clay, and pulling on the beads of the necklace and belt did not break them off. The recipient watched the actor sadly but silently. After 15 s, the actor gave up and placed the object back on the table. The recipient examined the object, smiled briefly (to indicate satisfaction that it was intact), placed it to the side, and neutrally began the next presentation.

In the Accidental Harm condition, the actor accidentally destroyed the recipient's objects. Thus, the actor admired the picture but accidentally tore it while returning it to the recipient. Similarly, the beads fell off the belt and the knot of the necklace came undone while she was admiring these objects. In the clay situation, the actor was looking at her watch when the bird was placed on the table so that when she turned to admire the bird, her arm collided with it and broke it. After each mishap, the actor said in a sorry tone of voice, "I didn't want that to happen" or "That wasn't on purpose" (in alternating order, beginning with the former). The actor apprehensively examined the broken object but did not repair it or apologize, and the recipient watched the actor silently but sadly. After 15 s, the actor placed the object to the side and the recipient neutrally began the next presentation.

As in Study 2a, familiarization trials were followed by a helping test and a final phase in which children could hand over a second ball. Finally, in the Intended-but-Failed Harm condition, the actor apologized and the recipient accepted the apology.

Coding and Reliability

Coding and reliability were conducted as in Study 2a. Agreement was perfect between the primary and second coders' coding of whom children gave the ball to, $\kappa = 1.0$. The second coder also judged that the actor was more likely to receive the ball in one case and the neutral person in four cases, but these numbers were unrelated to condition and to who received the ball (Fisher's exact test, $p = .40$ and $p = 1.00$, respectively).

3.3.2 Results

Since preliminary analyses revealed no gender effects, gender was not included in further analyses. All reported p-values are two-tailed. I again used the test proportion (67%) established by Study 1's Baseline condition. Binomial tests revealed that in the Intended-but-Failed Harm condition, a significantly lower proportion of children than 67% helped the actor (6 of 18, or 33%; binomial probability, $p = .008$). This difference did not emerge in the Accidental Harm condition (9 of 18 helped the actor, $p = .204$; see last two bars of Figure 3).

Chi-square analyses revealed that the Intended-but-Failed Harm condition differed significantly from the Baseline condition of Study 1 ($\chi^2[1, N = 36] = 4.00, p = .046$) but not from the Harm condition, $p = .457$. In contrast, the Accidental Harm condition differed marginally from the Harm condition ($\chi^2[1, N = 36] = 3.01, p = .083$) but not from the Baseline condition, $p = .310$.

Results from the final (second ball) phase again suggest that children were not afraid of the actor in the Intended-but-Failed Harm condition: Of the 12 children who helped the neutral person during the helping test, 10 handed the second ball to the actor.

3.3.3 Discussion: Study 2b

In Study 2b, 3-year-olds decreased their prosocial behavior towards a person who had harmful intentions towards a third party even if she did not cause negative outcomes. Moreover, children did not significantly decrease their prosocial behavior towards a person who caused negative outcomes without harmful intentions. These findings are the first evidence that by age 3, children selectively reduce prosocial behavior towards intentionally harmful – and thus morally blameworthy – individuals regardless of the consequences of those individuals' actions.

Study 2b contributes to the moral development literature since it suggests that as early as 3 years of age, children respond differently to harm that is caused intentionally versus unintentionally. In prior work in which intentions and outcomes were pitted against each other in hypothetical moral transgressions, children reliably used a perpetrator's intentions in their moral evaluations starting around 5 years of age (e.g., Wellman et al., 1979; Zelazo et al., 1996). Interestingly, however, Imamoğlu (1975) found that even when older children (5-year-olds) fail to differentially evaluate intentional versus accidental acts, they respond differentially on other measures such as like-dislike judgments of the perpetrator. Thus, young children's understanding of a perpetrator's intentions may be more apparent in their liking of or willingness to help the perpetrator than in their verbal evaluations of the transgression. Affiliation and helping may thus be precursors to and important facets of children's moral evaluations. More generally, non-verbal, non-hypothetical behavioral measures complement verbal measures of children's judgments and reasoning about intentions. Future work should accordingly employ both verbal and behavioral measures focusing on both the act and the actor in order to obtain a fuller picture of children's moral judgment-making.

It should also be noted that the findings of Study 2b are consistent with some findings from outside the moral domain. First, the findings make sense in light of the extensive research showing an impressive understanding of goals, intentions, emotions, desires, and beliefs by the third year of life (see, e.g., Baillargeon et al., 2010; Tomasello et al., 2005). Moreover, a study on children's understanding of deontic concepts (i.e., concepts of duty and obligation) showed that 3-year-old children do distinguish between accidental versus intentional non-compliance to deontic rules: These children were much more likely to judge deliberate violations of deontic rules as being naughtier than the accidental violations of such rules (Núñez & Harris, 1998; see Harris, 2000). Thus, contrary to what is suggested by the extensive work on 3-year-old children's moral judgments, children at this age do consider others' intentions in their judgments. Quite early in development, then, children are already making sophisticated judgments in which they take into account not only the consequences of others' behaviors but also the intentions behind those behaviors.

3.4 General Discussion: Studies 2a and 2b

The present studies demonstrate that young children's prosocial behavior is mediated by others' moral behavior. In Study 2a, 3-year-olds helped a harmful adult less than a neutral adult, extending Olson and Spelke's (2008) findings to children's own prosocial behavior. The important new finding from Study 2b was that 3-year-olds grasped the intentions behind harmful behavior and selectively decreased their prosocial behavior towards the actor if and only if she could be held morally responsible for her actions (i.e., when she had harmful intentions), even if she was unsuccessful in causing harm. Thus, by age 3, children selectively withhold help from morally blameworthy individuals.

It could be argued that children in the Harm and Intended-but-Failed Harm conditions helped the actor less due to her mildly aggressive tone of voice rather than her harmful behavior. However, some aggression is likely a reliable cue to intentional as opposed to

accidental harm. Indeed, the rationale for including mild aggression was that having the actor speak entirely neutrally would create an unnatural harming situation in which the actor's intentions would be ambiguous. For instance, children might infer from a neutral voice that the actor had not registered or did not remember that the objects belonged to the recipient and she thus did not intend harm. Thus, although future studies could control for the actor's aggression, I believe that some mild aggression is naturally linked with intentionally harmful behavior and retaining it creates a more ecologically valid situation.

There are also several reasons to believe that the actor's mild aggression alone does not explain the results. First, the actor was only aggressive in her tone of voice and only before the transgressions, not in any other way or at other times during the procedure. Also, nearly all children in both the Harm and Intended-but-Failed Harm conditions gave a second ball to the actor, suggesting that her aggression had not made them afraid of her (though it is possible that children were afraid of the actor but their fear did not stop them from helping her in this context). Finally, as will be seen in the next chapter (Study 3), children do seem to recognize moral transgressions even in the absence of aggression from the transgressor. It is thus unlikely that the present findings are due solely to the slightly aggressive way in which the actor spoke before her harmful actions.

In the present studies, children directed less helping towards a person who harmed or intended to harm a third party. A similar result emerges in adults: In economic games, adults punish individuals who show (intentionally but not accidentally) unfair behavior even when they were not themselves affected by that behavior (Blount, 1995; Fehr & Fischbacher, 2003, 2004; Fehr & Gächter, 2002; Singer et al., 2006). Moreover, the reward areas of adult (male) brains are activated when a previously unfair individual is in pain (Singer et al., 2006). Third-party punishment and related processes are thus prevalent among adults. The current findings show that the ability to identify harmful individuals in third-party interactions and to punish

those individuals emerges early in childhood. As noted in Chapter 1, one may punish norm violators by, among other methods, withholding help from them, or shunning or socially ostracizing them (Boyd & Richerson, 2005). The findings from Studies 2a and 2b are consistent with both of these possibilities; that is, children may have specifically withheld help from the harmful actor, or they may have more generally shunned or avoided the harmful actor. Teasing apart these possibilities will require further research in which children's tendency to shun is assessed separately from their tendency to withhold help. For example, a future study could assess whether children affiliate with (e.g., play with) a harmful actor even though they withhold help from her. Regardless of the specific mechanism that children employed, the findings of Study 2 demonstrate that 3-year-old children do engage in some form of third-party sanctioning, probably as a result of an understanding of and desire to enforce moral norms (Boyd & Richerson, 1992; Fehr & Fischbacher, 2003, 2004).

The present results are limited to children's interactions with adults and may not generalize to children's interactions with their peers (see, e.g., Killen, 1991). Numerous authors have observed that child-child interactions provide a rich and unique context within which children develop a sense of fairness, equality, and justice (e.g., Arsenio & Lover, 1995; Damon & Killen, 1982; Dunn, Cutting, & Demetriou, 2000; Piaget, 1997; see Smetana, 2006). Thus, when in such interactions, children may display a more advanced command of these concepts and at still younger ages than I have shown here. Alternatively, my use of child-adult interactions in a laboratory may have made children comply with what they perceived as the adults' expectations. Note, though, that my helping measure did not entail providing responses about the moral transgressions directly to an adult; rather, it was a more implicit assessment that took place after, and in a different situation than, the transgressions. Children's responses are thus unlikely to have been greatly influenced by their

perceptions of the adults' expectations. Still, on the view that children were complying with such expectations, children may not show the same sensitivity in interactions with other children. Even so, the present findings at least demonstrate young children's ability to recognize harmful intentions in interactions with adults under controlled conditions. It remains for future research to assess this ability across multiple contexts and types of interactions.

Relatedly, although I assessed the impact of the intentions behind and outcomes of moral behavior, I recognize that these are only two of several criteria that fall under only one of several domains that impact children's moral judgments. Other criteria in the moral domain include, for instance, whether the perpetrator apologized and whether she was already punished (e.g., D. T. Miller & McCann, 1979). Criteria in the psychological domain include the victim's and the children's own relationship with the perpetrator (e.g., Slomkowski & Killen, 1992; Wellman et al., 1979). The context of the transgression and individual differences among children also play a role, and of course, all of the above domains and factors interact with each other (see Helwig, 2006; Smetana, 2006, for reviews). Intentions and outcomes thus form only one piece of the rich and multifaceted area of children's moral evaluations.

Furthermore, I have broadly used the term "moral behavior" here to mean acts that have consequences for others' rights or welfare (e.g., Smetana, Schlagman, & Adams, 1993). However, morality also includes other criteria, such as obligation and independence from authority sanctions, which distinguish it from the conventional, psychological, and personal domains of social knowledge (e.g., Smetana, 2006; Turiel et al., 1987). My focus here, however, was on whether children's prosocial behavior is mediated by an individual's harmful transgressions against others. Although these are precisely the types of transgressions that fall into the moral domain, my focus on prosocial behavior rather than on

moral judgments precluded an assessment of whether children perceived the transgressions as moral in the strict sense or not. Children may have withheld help for many reasons (such as how much they liked the perpetrator) that are related to but not the same as judging the transgressions to be moral. All the same, children's ability to recognize harmful behaviors and intentions as seen in their prosocial behavior might be an early step on the way to the full-blown, explicit moral judgments that children make just a few years later and as such, is important to explore.

All in all, punishment in third-party contexts is thought to be a crucial mechanism for maintaining morality in large-scale human societies (Boyd & Richerson, 2005; Fehr & Fischbacher, 2003, 2004). Studies 2a and 2b shed light on the quite sophisticated abilities of very young children to monitor third-party interactions and to subsequently punish (in the form of withdrawing help from) harmful actors. This suggests an early ontogenetic appearance of third-party punishment.

An equally important mechanism for enforcing moral norms and maintaining morality might be to actively attempt to prevent a moral transgression rather than (or in addition to) subsequently punishing the transgressor. Do young children also actively intervene in this way? This was the primary question that I addressed in Study 3 (next chapter). I also addressed two secondary questions that tied Study 3 back to Studies 1 and 2. First, in line with Study 2, I again assessed children's third-party punishment, but in Study 3, I asked whether children engage in a different sort of third-party punishment, namely, tattling on the transgressor. Second, to partially replicate and extend Study 1, I assessed whether children's prosocial behavior increases towards a victim.

4. STUDY 3: INTERVENTION IN MORAL TRANSGRESSIONS

4.1 Introduction

Preschool-age children know a lot about morality, clearly differentiating, for example, between moral norms (pertaining to issues of justice, rights, and welfare) and social conventions (the arbitrary, consensually agreed-upon behavioral norms that regulate social interactions within social systems). They recognize that moral norms, but not social norms, are obligatory and universally applicable (for reviews, see Smetana, 2006; Turiel, 2006). Preschoolers also make quite sophisticated moral judgments in which they take into account various aspects of the transgression, the transgressor, and the victim (e.g., Slomkowski & Killen, 1992; Wellman et al., 1979; Zelazo et al., 1996; see Chapter 1).

Most prior work assessing children's responses to moral transgressions has relied on children's judgments and verbal reasoning about the transgressions, and thus provides important insight into children's ability to distinguish various types of transgressions and to identify and evaluate their critical elements (see Chapter 1). Only limited research has investigated children's actual moral behavior during transgressions, which is surprising given that moral judgments may not always relate to moral behavior (e.g., Astington, 2004; Wainryb et al., 2005). Still, some researchers have employed naturalistic observations and have found that children's spontaneous responses to moral versus conventional transgressions differ significantly (see Smetana, 2006; Turiel, 2006). For instance, toddlers respond to moral transgressions (i) more frequently than to conventional transgressions, and (ii) by focusing transgressors on the intrinsic (hurtful and unjust) consequences of the acts on victims (e.g., Smetana, 1984, 1989). Older (school-aged) children respond to both types of transgressions but in distinct ways: They respond to moral transgressions like the younger children, by focusing on the intrinsic consequences of the acts, whereas they respond to conventional transgressions by focusing on features of social organization, rules, and normative

expectations (e.g., Nucci & Nucci, 1982a; Smetana, 1989). Surprisingly, however, children's behavioral responses to moral transgressions have never been explored in a controlled experiment.

More importantly, the prior observational work has primarily assessed situations in which children were themselves victims (e.g., when another child took their toy away) and may thus have assessed only children's emotional reactions to being harmed rather than their possession and understanding of moral norms per se. Indeed, chimpanzees retaliate against conspecifics that steal food from them, and they do so as a function of how angry they are at the theft (K. Jensen et al., 2007). As discussed earlier, the real test case for moral norms in moral behavior is the moral behavior when a third party has been harmed, as this behavior shows an agent-neutral application of the norm (see Chapter 1). Indeed, in evolutionary analyses of the origins of morality, the key situations are instances of third-party intervention or punishment (Gintis et al., 2003; Krebs, 2008).

In Study 2 (previous chapter), I assessed whether young children's prosocial behavior varies based on the beneficiary's moral behaviors in third-party interactions, and found that 3-year-olds do selectively direct prosocial behavior away from individuals who harm or even intend to harm others. These results provided evidence for an early emergence of third-party intervention (in the form of withdrawing cooperation or shunning). However, there are alternative forms of third-party intervention that might also serve to maintain morality (see Boyd & Richerson, 2005). One such method is to attempt to prevent a moral transgression from taking place by protesting or retaliating against the transgressor. This method would be especially valuable when the victim was unable to defend herself (say, because she was too weak or was absent while the transgression was being committed). Whether children actively intervene in this way remains unknown. In the present study, therefore, I asked whether young children actively enforce moral norms by intervening and attempting to prevent moral

transgressions from taking place. In particular, I studied 3-year-old children's intervention (verbal protest) as they observed a moral transgression.

The procedure of Study 3 was adapted from a recent study investigating children's protest during social conventional transgressions (Rakoczy, Warneken, & Tomasello, 2008), in which 2- and 3-year-olds learned the rules of a novel game and then saw a puppet either breaking or not breaking the rules. (A puppet was used because intervening against an adult is likely a daunting task for children). Rakoczy et al. (2008) found that children protested normatively and imperatively more often when the puppet broke the game rules than when it did not break the rules, and 3-year-olds protested more explicitly on the normative level than did 2-year-olds. I adapted this paradigm to assess 3-year-old children's responses to moral transgressions. Children thus saw an actor puppet either harming a recipient puppet by destroying her belongings (harm condition) or acting neutrally by destroying irrelevant objects (control condition). If children intervene because the actor's actions are harmful towards another individual rather than because the actions themselves are negative, then they should do so more in the harm than in the control condition. Importantly, the recipient was absent during the transgressions, since in both Study 1 and Study 2, I observed that children do not intervene very often when the victim is present, perhaps because they expect the victim to stand up for herself.

To gain a fuller picture of children's moral intervention, I addressed two further issues in Study 3. First, in line with Study 2, I again assessed third-party punishment, but I focused on a different sort of punishment, namely, tattling. Specifically, I assessed whether, upon the recipient's return, children tattled on the transgressor to the recipient. Ingram and Bering (in press) investigated children's tattling in a naturalistic setting, and found that tattling to an adult about norm (especially moral norm) violations was frequent among 3- to 4-year-olds. Interestingly, however, tattling in their study primarily concerned negative

behavior towards the tattler (i.e., in dyadic interactions); tattling on behalf of third parties was rare. However, as the authors suggested, children may have refrained from third-party tattling because they expected victims to speak up for themselves. Since the victim in the present study was absent during the transgressions, I expected to see more tattling upon her return if she had been harmed during her absence.

Second, to replicate and extend Study 1, I assessed children's prosocial behavior towards the recipient after her return. As previously discussed, infants and young children sympathize with distressed others and initiate prosocial responses towards them (e.g., Eisenberg et al., 2006; Zahn-Waxler, Radke-Yarrow et al., 1992). Moreover, as seen in Study 1, young children show similar responses even when the victim provides no distress cues. I thus predicted that, despite the victim's absence (and thus despite the lack of distress cues) during the transgressions, children would show greater subsequent prosocial behavior towards the victim if she had been harmed than if she had not been harmed.

4.2 Method

Participants

Participants were 3-year-old children (n = 32, 16 girls) between 36;00 and 39;28 (M = 38;6; SD = 38.3 days). One additional child was tested but excluded due to experimenter error. The children were recruited from urban daycare centers in a medium-sized German city. The language, ethnicity, socioeconomic status, education, and occupation information was identical to the previous studies. All participants were tested by two female experimenters (E1 and E2) who played the same roles (actor or recipient) each time. In a between-subjects design, 16 children (8 girls) received a harm condition and 16 children (8 girls) received a control condition.

Materials and Setting

Two hand puppets (a cow and an elephant) served as children's interaction partners throughout the procedure. A ball and a wooden puzzle were used as warm-up games. During the test trials, blank sheets of white paper, color pencils, and white balls of clay were used. The child, the actor puppet (always controlled by E1), and the recipient puppet (always controlled by E2) sat at a table, with the puppets sitting opposite one another. On the floor next to E1 lay an empty bin.

Procedure

Prior to testing, E1 and E2 played with the child in the child's playgroup. When the child was comfortable, he or she was taken to the testing room. E1 and E2 introduced the puppets (which puppet was the actor or recipient was counterbalanced across children) and began the warm-up games in which the child and the puppets first played with a ball and then with a puzzle. The puzzle game was designed to familiarize children with situations in which they could intervene. Thus, during this game, each puppet made a mistake (e.g., placed a puzzle piece in the wrong position) while the other puppet was turned away. If the child did not intervene within a few seconds, the other puppet turned back to the table, remarked that the puzzle piece was wrongly placed, and prompted the child to help place it correctly. If the child had still not done so, the puppet that had prompted the child would have placed the puzzle piece correctly. However, this last step was not required as all children corrected both puppets' mistakes either spontaneously or after being prompted. The warm-up phase lasted 6-7 min.

Children then received two test trials (Picture and Clay; order counterbalanced across children). These trials were based closely on the Picture and Clay situations used in Studies 1 and 2. Each trial consisted of an activity phase and a testing phase. The actor always initiated

the first test trial and the recipient initiated the second test trial. The activity phases were as follows:

Picture. The assigned puppet brought out four blank sheets of paper and three color pencils, and handed one of each to the child, the other puppet, and herself. Realizing that she had one extra sheet of paper, she placed it to the side on the table. Then the child and the puppets each drew a picture: The actor drew a house, the recipient drew a flower, and the child drew whatever she liked. During this activity, each puppet happily and proudly showed off her drawing twice to demonstrate how much she liked it, and both puppets took an active interest in the child's drawing. When they were all finished (after about 8 min), the recipient said she would clear up, and placed the child's drawing on the floor between herself and the child, and placed her own and the actor's drawings on the table next to the blank sheet of paper. The locations of the recipient's drawing and the blank sheet were counterbalanced: For half the children, the blank sheet lay between the two puppets' drawings, and for the other half, the recipient's drawing lay between the blank sheet and the actor's drawing. The actor's drawing always lay closest to the actor.

Clay. The assigned puppet brought out a bowl with four white balls of clay and handed one ball each to the child, the other puppet, and herself. Realizing that she had one extra ball, she placed it to the side on the table and put the bowl away. The child and the puppets each made a clay sculpture: The actor made a tortoise, the recipient made a snail, and the child sculpted whatever she liked. Again, each puppet happily and proudly showed off her sculpture twice during the activity, and both puppets took an active interest in the child's sculpture. When they were all finished (after about 8 min), the recipient said she would clear up, and placed the child's sculpture on the floor, and placed her own and the actor's sculptures on the table next to the extra clay ball (locations counterbalanced as in the Picture situation).

After each activity phase, the recipient puppet (and E2) left the room (saying she had "forgotten something outside") and the testing phase began according to condition (assigned randomly prior to testing). In the harm condition, the actor puppet looked at the recipient's drawing or sculpture and said in a neutral but firm manner, "Well, I don't like the cow's flower/snail. I'm going to tear/break it now." (Note that the harmful actor did not speak in an aggressive tone, thus addressing the issue raised in Study 2 about whether children only responded to the harmful actor's aggressive tone of voice rather than to her moral transgression.) In the control condition, the puppet looked at the blank sheet of paper or extra clay ball and said in the same manner as in the harm condition, "Well, I don't like the blank sheet of paper/ball of clay. I'm going to tear/break it now." In both conditions, the puppet then moved towards the target object (about 5 s), looked at it, and said, "Yes, I'm going to tear/break it now." The puppet picked up the object and returned to her original location (about 5 s), placed the object in front of her, said, "Yes, I'm going to tear/break it now," and started to destroy the object and throw the pieces into the bin (about 30 s). The actor's intention was repeated and her actions presented in this step-wise manner to provide children with enough time and occasions to protest. If children physically intervened by taking away or otherwise physically protecting the object at any stage, the testing phase was ended. During testing phases, children's protest and physical intervention were coded (see Coding and Reliability section below for details).

After destroying the object, the actor puppet was silent and looked straight ahead at the table while E1 signaled the recipient (by coughing) to return. Upon reentering, the recipient puppet looked into the bin and neutrally said "Hmm" to show that she had noticed something in the bin. She looked at the remaining object on the table (either her own object or the control object, depending on condition), again said "Hmm" neutrally, and looked back into the bin. In the harm condition, the recipient then said in a somewhat surprised and sad

tone, "Oh, that was my flower/snail," waited about 6 s, said "Oh well" mildly despondently, and returned to her seat. In the control condition, the puppet behaved the same way except that she noted in a neutral tone that the object in the bin had been the blank sheet of paper or extra ball of clay. From the time the recipient reentered the room until she returned to her seat, children's tattling and prosocial behavior were coded (see Coding and Reliability section below for details).

The second test trial began when the recipient was reseated. In both conditions, the recipient was somewhat quiet at the start of the second activity phase (for about 50 s), but was then happy again as in the first activity phase. In the harm condition, this helped ensure that children believed that the recipient puppet really cared about her objects and was sad when they were destroyed (which helped maintain the validity of the second test trial). However, it was also done in the control condition so as to keep the two conditions as similar as possible. After the second test trial in the harm condition, the actor apologized for destroying the recipient's objects and drew her a new picture.

Coding and Reliability

The primary coder coded all sessions from videotape. For the warm-up puzzle game, children's interventions when the puppets made mistakes were coded as spontaneous or prompted. For reliability, a second coder coded a randomly-selected 25% of the sample. Reliability was perfect, $\kappa = 1$.

Children's protest was coded using a modified version of Rakoczy et al.'s (2008) coding scheme. Children's responses during both testing phases were first transcribed and the following codes (and scores) were then assigned to the responses (see Table 4 for details): 'normative protest' (score of 3), 'imperative protest' (2), 'hints of protest' (1), or 'no protest' (0). For analyses, the highest score assigned across the two testing phases was used as the final protest score for that child. A second coder (blind to hypotheses) coded a randomly-

selected 25% of the sample. Reliability on whether or not a child showed protest was $\kappa = 1$, and on the highest code for each trial was $\kappa = .88$.

Table 4. Coding scheme for protest

Category	Protest score	Behaviors
Normative protest	3	Child intervenes in a normative way, using normative vocabulary, reference to the rule ("No, you're not supposed to do that" or "You may not do that"), or reference to the recipient's emotional state ("She will be sad then")
Imperative protest	2	Child expresses an imperative, such as a command to stop the action, without use of normative elements ("No! Don't tear it!"), or expresses simple disagreement with the actor's action ("No!")
Hints of protest	1	Child protests but clear attribution to the other two categories is not possible; includes using a protesting tone of voice in exclamations ("Hey!"), questions ("Why are you doing that?"), or statements ("But I like the cow's flower" or "That is not nice")
No protest	0	Child shows no protest

Three additional, secondary measures were coded. Children were coded as physically intervening during the transgressions if they physically prevented the target object from being destroyed. Upon the recipient's return, children's tattling and prosocial behavior were also coded (see Table 5 for details).

Physical intervention, tattling, and prosocial behavior were all coded as binary (yes/no) variables (these were not coded in more detail because they were secondary variables that were not used to address the primary question of the study, namely, whether children verbally protest against moral transgressions). A second coder (blind to hypotheses)

coded these measures for a randomly-selected 25% of children. Reliability was excellent: κ = 1, κ = .81, and κ = 1 for physical intervention, tattling, and prosocial behavior, respectively.

Table 5. Coding scheme for physical intervention, tattling, and prosocial behavior

Category	Coding	Behaviors
Physical intervention	Yes/No	Child protects the target object such that the actor puppet cannot destroy it, including by placing the object out of the actor's reach or physically withholding the actor puppet from reaching it
Tattling	Yes/No	Child tells recipient puppet or E2 with a complaining or disapproving tone of voice that the actor puppet or E1 destroyed the target object ("The cow tore it up" or "She tore up your beautiful flower!"), accompanied by explicit naming and/or pointing to the actor puppet or E1 (thus, simply neutrally informing the recipient about what had happened was not coded as tattling)
Prosocial behavior	Yes/No	Child comforts (e.g., strokes), makes suggestions ("You can draw a new flower" or "You can draw my Lion further"), helps (e.g., offers to draw another picture), or shares with recipient puppet or E2 (e.g., gives own drawing) (cf. Study 1, and Zahn-Waxler et al., 1992)

4.3 Results

Warm-up. During the warm-up puzzle game, children could correct the puppets' mistakes. All 32 children corrected each puppet at least once. Two children needed to be prompted to do so when the first puppet made a mistake but all children did so spontaneously when the second puppet made a mistake. Thus, all children knew that they could correct the puppets and were comfortable doing so.

Testing phases. I first report results of children's protest, followed by their physical intervention, tattling, and prosocial behavior. All reported *p*-values are two-tailed.

Protest. Children's protest was first analyzed as a binary variable. Thus, all three protest forms (imperative, normative, and hints) were pooled into one 'protest' code and compared against the 'no protest' code. As predicted, a chi-square analysis revealed that significantly more children protested in the harm than in the control condition, $\chi^2(1, N = 32) = 8.00, p = .005$. Specifically, 12 of 16 children (75%) in the harm condition protested at least once across the two trials, whereas only 4 of 16 children (25%) in the control condition did so. Since the 'hints of protest' code included relatively weak instances of protest, I conducted another analysis in which normative and imperative protest were pooled into a 'protest' category, and hints of protest and no protest were pooled into a 'no protest' category. This analysis again revealed a significant difference: 9 of 16 children (56%) in the harm condition showed imperative or normative protest whereas only 3 of 16 children (19%) in the control condition did so, $\chi^2(1, N = 32) = 4.80, p = .03$. A final analysis of protest kept the three protest codes separate and compared children's highest codes across conditions. This revealed significant condition differences in the highest codes assigned to children's protest responses, Mann-Whitney $U = 58.0$, Nharm = Ncontrol = 16, $p = .007$ (see Figure 4). Children thus clearly showed greater verbal protest when witnessing a puppet harming another puppet than when witnessing a physically similar but harmless situation.

Physical intervention. Physical intervention during the transgressions was infrequent: Only 5 children (4 in the harm condition and 1 in the control condition) showed such intervention. Statistical tests were thus not conducted on this measure. However, all of these 5 children also protested verbally using either imperative or normative protest and are thus represented in the analyses of protest reported above.

Tattling. A comparison of children's tattling across conditions revealed the predicted significant condition difference: Whereas 7 of 16 children (44%) in the harm condition tattled on the actor, none of the children in the control condition did so, Fisher's exact test, $p = .007$.

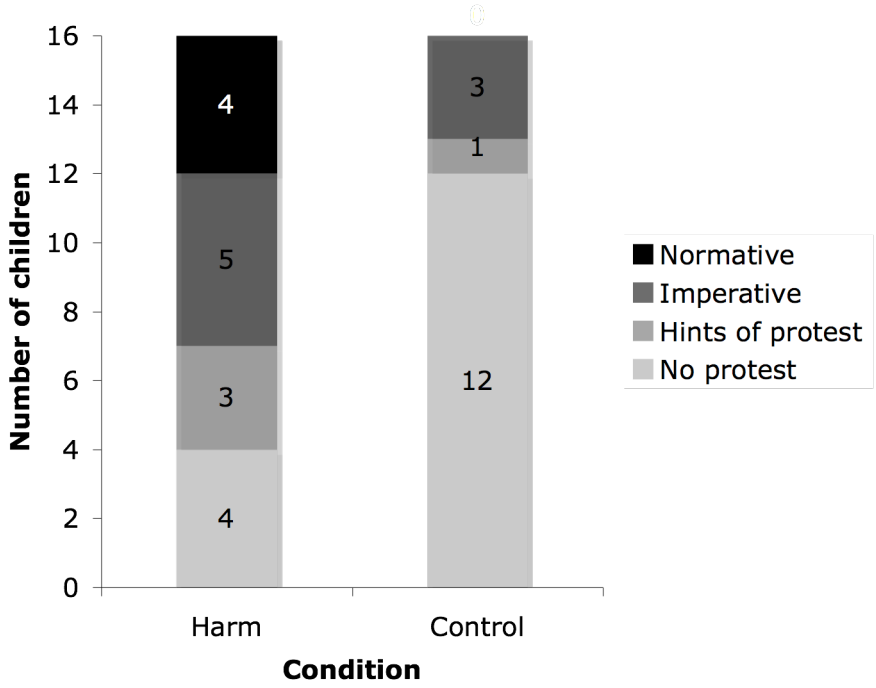

Figure 4. Number of children in each condition who showed each form of protest as their highest form of protest.

Prosocial behavior. Finally, children's prosocial behavior towards the recipient also revealed the expected result: 7 of 16 children (44%) behaved prosocially in the harm condition whereas none of the children behaved prosocially in the control condition, Fisher's exact test, $p = .007$.

In sum, children's responses to witnessing a puppet's harmful actions against another puppet differed strikingly from their responses to witnessing a puppet's neutral actions. Children protested more during the transgressions, tattled against the transgressor, and showed prosocial behavior towards the victim puppet, all behaviors that were either much reduced or entirely absent in the neutral case.

4.4 Discussion

Study 3 examined how children behave when they witness moral transgressions involving a third party as victim. After 3-year-old children witnessed a transgressor puppet harming another puppet by destroying her belongings, they actively intervened on behalf of

the absent victim by verbally protesting against the transgressions while they were occurring, and subsequently tattling on the transgressor to the victim and acting prosocially towards the victim. These behaviors were significantly reduced or entirely absent in a control condition in which the actor puppet behaved very similarly but did not harm the other puppet. Thus, children did not respond to the actions per se but only when the actions broke a moral norm. Beyond knowing what constitutes moral versus other types of transgressions (see Smetana, 2006; Turiel, 2006), responding to transgressions against themselves (e.g., Dunn & Munn, 1987) and withdrawing help from third-party transgressors (Study 2), 3-year-olds also seem to actively intervene against third-party transgressions, applying their moral norms in action and in an agent-neutral manner.

Children in this study verbally protested against moral transgressions to indicate their disagreement with them. Moreover, in the harm but not the control condition, some children's protest was normative; that is, children not only protested against moral transgressions but also did so with reference to relevant moral norms ("You're not supposed to do that") or moral reasons for why it was wrong ("[The victim] will be sad then"). Several children's protests were also imperative in nature. That is, children gave clear commands to stop the action or simply disagreed with a "No." Finally, some children showed hints of protest, clearly showing disagreement with the transgressions but less explicitly than with the other two forms of protest. The majority of children in the harm condition thus found some way to express disagreement with and attempt to prevent the moral transgressions, some more explicitly than others. This display of protest behavior is especially remarkable given that, unlike in Rakoczy et al.'s (2008) study in which children were taught game rules that they then saw broken, children in the present study were not taught any rules prior to testing. Children thus protested without being taught or reminded about any norms, suggesting that

they came into the testing situation in possession of the basic moral norms that they then saw violated.

The victim's absence during the transgressions allowed children to intervene on her behalf. However, these findings leave unclear whether children intervened out of concern for the victim or because the moral norm that prohibits destroying others' belongings was violated; still, children clearly did intervene against the moral transgressions. It also remains unclear whether children whose highest forms of protest were imperative or hints of protest grasped the normative nature of the moral transgressions but were simply unable to verbalize it, or whether they grasped that a transgression was being committed but did not fully grasp the norms that rendered it a moral transgression. To better address these issues, future studies might employ a hybrid design using both the design of the present study and the interview design used in prior studies (Nucci & Nucci, 1982a, 1982b). Thus, children's intervention in moral transgressions could be observed and children could then be interviewed about their judgments and reasoning about the transgressions they just witnessed. This would allow researchers to better assess how children who show different forms of responses during the transgressions reason about and evaluate the transgressions. Such an approach would not only provide a more complete picture of children's moral understanding but might also provide important information about individual differences in children's intervention and judgment-making.

Assessments of children's behavior other than protest revealed further interesting results. First, several children in the harm condition (but, importantly, not in the control condition) tattled about the transgressor to the victim. In contrast to prior work in which preschoolers were found to tattle primarily when they themselves had been harmed (Ingram & Bering, in press), the current results show that preschoolers also tattle when the transgression affects a third party. A rich interpretation of this finding is that children were

communicating about the moral violations so as to ensure that some punitive measure be taken against the transgressor, but it is also possible that children tattled to inform the victim (who was absent during the transgression) that *they* had not caused the damage and so avoid being punished by the victim. One way to assess this possibility in future work would be to have a new individual (one who is not the victim and who is unaware of the situation) enter the room after the transgression. If children tattle in this situation as well, then it would seem much more plausible that they tattle to enforce norms in an agent-neutral way rather than to avoid being punished. In any case, 3-year-olds do seem to tattle in third-party contexts, at least when the victim was absent during the transgression.

Children in the harm (but not the control) condition also acted prosocially towards the recipient upon her return. This fits with Study 1, in which toddlers sympathized with a victim even when she did not provide any emotional cues upon being harmed and they subsequently helped her more if she had been harmed than if she had not been harmed. I argued that children in Study 1 may have engaged in affective perspective-taking such that even in the absence of distress cues from the victim, they were nevertheless able to apprehend how the victim was feeling and were thereby able to sympathize with her. A similar process might have taken place in the current study whereby even in the victim's absence, children were able to apprehend how she would feel, sympathize with her, and thus act prosocially towards her upon her return. More broadly, prosocial behavior towards the victim is an important additional measure because the other two measures – protest and tattling – could conceivably be responses to the violation of a social norm more generally, but children's subsequent prosocial behavior towards the victim shows that they really did view the harm situation as a moral transgression against the victim.

Finally, it is important to consider when in ontogeny third-party intervention in moral transgressions might emerge. It has been argued that the foundations of morality are innate

(Wynn, 2008) and that an understanding of basic moral norms is in place very early in development (see Smetana, 2006; Turiel, 2006). Certainly, children much younger than 3 years of age monitor and evaluate third-party social interactions (Kuhlmeier et al., 2003) and seem to prefer helpful over hindering characters (Hamlin et al., 2007). In the second and third years of life, children protest and argue against moral violations against themselves (Dunn & Munn, 1987; Smetana, 1984, 1989). It is thus plausible that even earlier than 3 years of age, children would intervene in third-party moral transgressions. This idea gains support from the finding that 2-year-olds protest against others' social conventional transgressions, although in less explicit and normative ways than 3-year-olds (Rakoczy et al., 2008). Interestingly, prior work has shown that during the second year, children respond more frequently and in more differentiated ways to (second-party) moral violations than to social conventional violations (Dunn & Munn, 1987; Smetana, 1984, 1989). This raises the possibility that 2-year-olds would actually show more (or more sophisticated) intervention in third-party moral transgressions than in social conventional transgressions (Rakoczy et al., 2008). This possibility remains for future work to explore.

In sum, by age 3, children not only evaluate and judge simple moral transgressions but also actively protest against such transgressions, even when the victim is a third party. Moreover, they engage in third-party punishment by withdrawing help from (or shunning) the transgressor (Study 2) as well as tattling on the transgressor (Study 3). Also impressive is their behavior towards victims of third-party transgressions: Young children sympathize with victims even if those victims show no distress (Study 1) and they subsequently behave prosocially towards victims (Studies 1 and 3). Already by 3 years of age, then, children possess quite a sophisticated and multifaceted moral system as evident in their moral behavior.

Just a year or two later, children's judgments also reveal a multifaceted moral understanding. As discussed in Chapter 1, prior work has shown that by 4 or 5 years of age, children make quite refined moral judgments that take into account multiple factors about the transgressor and the context of the transgression. In the last study of this dissertation, I tested the flexibility of this moral understanding. Specifically, I asked whether 4- and 5-year-old children judge and subsequently cooperate with a remorseful transgressor differently from an unremorseful one.

5. STUDY 4: APPEASEMENT FUNCTIONS OF GUILT DISPLAYS

5.1 Introduction

Guilt is the aversive emotion that follows the realization that one has harmed another person (Nelissen & Zeelenberg, 2009). Like other self-conscious emotions, guilt likely plays a central role in socialization and adherence to conventions, norms, and morals. The aversive feeling of guilt, or the anticipation thereof, is a powerful mechanism that prevents individuals from transgressing and motivates individuals to rectify their transgressions (Hoffman, 1982; Keltner, 1995). Empirical work with adults provides support for this function of guilt (e.g., Ketelaar & Au, 2003), and work with young children also suggests a link between the experience of guilt and moral development (e.g., Kochanska, DeVet, Goldman, Murray, & Putnam, 1994; Kochanska et al., 2002; Zahn-Waxler & Kochanska, 1990).

Why, though, do transgressors display and verbalize their feelings of guilt to others? A prevailing view is that guilt displays serve critical appeasement functions[5]. Thus, the display of guilt shows others that the transgressor is also suffering, which evokes sympathy, concern, and forgiveness, and thus reduces the likelihood of punishment (Keltner & Anderson, 2000; Leary et al., 1996). Guilt displays also signal to others that the transgressor is aware of and committed to the norms of the group, and so the transgression is not reflective of the transgressor's personality, character, or ability. They may furthermore serve as a promise of more acceptable conduct in the future (Castelfranchi & Poggi, 1990; Goffman, 1967; Keltner et al., 1997; Leary et al., 1996; Nelissen & Zeelenberg, 2009). A remorseful transgressor is thus more likely to be seen as self-policing, dependable, and cooperative, and

[5] Note that displays of other self-conscious emotions (e.g., embarrassment and shame) are also thought to serve appeasement functions (see, e.g., Keltner, Young, & Buswell, 1997). For instance, adults report high levels of affiliative emotions such as amusement and sympathy in response to others' embarrassment (Keltner et al., 1997; R. S. Miller, 1987). Adults also help individuals who previously displayed an appropriate amount of embarrassment more than individuals who displayed too much embarrassment or none at all (Levin & Arluke, 1982; Semin & Manstead, 1981). Guilt displays are thus not alone in serving appeasement functions. Here, however, I focus on guilt displays alone.

to thus elicit more forgiveness, affiliation, and cooperation from the victim and from other group members, than an unremorseful transgressor (Darby & Schlenker, 1989).

Empirical research has shown that adults are indeed sensitive to these social and interpersonal functions of displaying guilt. For instance, adult victims' aggression is reduced towards an apologetic transgressor (Ohbuchi et al., 1989). Moreover, when judging third-party interactions, adults judge that transgressors who feel guilty need not make as many material sacrifices or restitutions to repair their misdeed as should transgressors who do not feel guilty (O'Malley & Greenberg, 1983).

The ontogenetic emergence of sensitivity to guilt displays is unknown, however. The extant literature suggests that by middle childhood, children do have some understanding of guilt. By around 7 years of age, for instance, children are aware of what kinds of situations elicit guilt versus other social emotions (Harris, Olthof, Terwogt, & Hardman, 1987). Children around this age also know that a person who feels guilt would and should apologize, formulate good intentions, and mend or substitute damaged objects (Berti, Garattoni, & Venturini, 2000). However, it remains unclear whether children understand the appeasement functions of guilt displays. Do they make appropriate inferences about the effects that guilt displays have on the victim? Do they themselves prefer to interact with transgressors who display guilt (and are thus more dependable and cooperative group members) than those who display no guilt? Finally, do they cooperate more with (by distributing more resources to) a transgressor who displays guilt than one who does not? These were the questions I addressed in the present study.

Similar questions have been asked in the related work on children's understanding of apologies, which are admissions of blameworthiness and regret and thus function as a stand-in for guilt. This work shows that quite young children do understand and draw appropriate inferences from apologies. Thus, apologetic actors are blamed less, forgiven more, liked

more, seen as more remorseful, and punished less by children as young as 6 years of age (Darby & Schlenker, 1982, 1989). Children around 4-5 years of age regard situations in which an actor apologizes as better and more just than ones in which the actor is unapologetic (Irwin & Moore, 1971; Wellman et al., 1979). Also around this age, children attribute negative feelings to an apologetic transgressor and improved feelings to a victim who has received an apology, though they seem not to make inferences about the transgressor's moral character based on whether she apologized or not (C. E. Smith et al., in press). Although this work on the understanding of apologies is interesting, it should be noted that children are regularly taught and prompted to apologize by their caregivers and teachers (see C. E. Smith et al., in press). Their ability to make judgments about transgressors who say "sorry" or who are described as having apologized may thus rely rather superficially on the use of key words such as 'sorry' or 'apologize.' It remains unclear how deeply children understand and what they can infer from the emotions behind apologies, namely, guilt and remorse. In the current study, therefore, I examined children's responses to displays of guilt without the confounding effects of apologies.

Prior work has also not explored whether children cooperate more with remorseful transgressors. In Study 2, I showed that young children are discriminating cooperators and helpers, as they reduce their helping towards harmful people and even towards people who intended to cause harm but were unable to do so. Moreover, 3-year-old children share more toys with a peer if that peer had previously shared toys with them (Levitt et al., 1985), and, when acting on behalf of a protagonist doll, they share more resources with dolls who have been generous to the protagonist doll or generous to another doll than they share with non-generous dolls (Olson & Spelke, 2008). Given that a guilt display communicates that the transgressor is otherwise a reliable and cooperative group-member, one may predict that children will also cooperate more with a remorseful than an unremorseful transgressor.

Children in the present study first watched videos of transgressors displaying guilt or no guilt about their transgressions, and were then asked a series of questions about the transgressors. The transgressors in the videos caused accidental rather than intentional harm. Although theories of guilt are unclear as to the role of intentionality (see, e.g., Baumeister et al., 1994; Fessler & Haley, 2003), prior empirical work shows that adults expect an accidental transgressor to experience more guilt than an intentional transgressor (McGraw, 1987). Also, apologies are effective in increasing forgiveness only following accidental harm, and actually seem *to decrease* forgiveness following intentional harm (Struthers, Eaton, Santelli, Uchiyama, & Shirvani, 2008). Thus, to ensure that the guilt displays in the present study would be believable and effective, I used accidental rather than intentional transgressions.

Guilt seems to have no single, clear facial expression but rather is associated with a variety of remorseful facial and vocal expressions, confessions, acceptance of responsibility, statements that the harm was accidental, apologies, and expressions of the desire to repair, or actual attempts to repair (Berti et al., 2000; Fessler & Haley, 2003; Keltner, 1995; Keltner & Buswell, 1996; Zahn-Waxler & Kochanska, 1990). However, I chose to limit the guilt display in this study to what might be the minimal components needed to clearly express guilt but without potential confounds. Thus, the guilt display in the present study consisted of a remorseful and concerned facial expression, confession, statement that the harm was accidental, and acceptance of responsibility. Apologies were excluded since I wanted to assess responses to guilt in the absence of apologies. Attempts to repair the damage were also excluded as these might bias children to judge the remorseful transgressor as more prosocial but for reasons other than the guilt display. Thus, the guilt display in the current study was restricted to a few core features.

I thus assessed 4- and 5-year-old children's grasp of the appeasement functions of guilt displays by assessing their judgments of accidental transgressors who displayed guilt versus no guilt. Based on prior work (e.g., C. E. Smith et al., in press; Wellman et al., 1979), I expected that at least by 5 years of age, children would expect a victim to be more upset with and to prefer a remorseful transgressor, and that they would themselves prefer to interact with a remorseful transgressor. I also predicted that children would distribute more resources to the remorseful transgressor. Following C. E. Smith et al. (in press), I also asked children about the moral character of the transgressors (i.e., whether the remorseful or unremorseful transgressor is meaner), but I did not have specific predictions given prior findings that 4- to 5-year-olds failed to make appropriate inferences about the moral character of transgressors (C. E. Smith et al., in press). Lastly, I also asked children to provide justifications for their judgments, as these are useful in revealing the criteria on which children base their judgments (Grueneich, 1982; Turiel, 1998).

5.2 Method

Participants

Participants were 4-year-old children (N = 20, 10 girls) between 48 months, 25 days and 53 months, 12 days (M = 51 months, 7 days; SD = 1 month, 12 days) and 5-year-old children (N = 20, 10 girls) between 60 months, 1 day and 66 months, 4 days (M = 63 months, 1 day; SD = 2 months, 8 days). Four additional children were tested but excluded due to experimenter error (two 5-year-olds), equipment failure (one 5-year-old), or unwillingness to participate (one 4-year-old). Children were recruited from and tested in their daycare centers in a medium-sized German city. All children were native German speakers whose parents had given permission for them to participate in child development studies. The ethnicity, parents' socioeconomic status, education, and occupation information was identical to the previous studies.

Design and Materials

During the experiment, children sat at a table on which two laptop computers (PowerBook G4 computers with 15-inch, 1440 x 960-pixel screens) were placed, one to the left and one to the right of the child. All videos were played using the full-screen option in Quicktime Player. A camera recorded a frontal view of the children and a microphone placed between the computers supplied sound to the camera. The procedure had two phases. In each phase, children saw one Guilt and one No Guilt video, about which they received comprehension probes (as manipulation checks, i.e., to make sure they grasped the content of the videos) and eight test questions. After the second phase, children received a distribution of resources task and one final test question about why they had distributed the resources in the way that they had. Thus, altogether, children watched four videos (two per phase) and answered 17 test questions (eight after each phase and one after the distribution of resources task).

Video Stimuli

Videos featuring three adult actresses (research assistants in the lab) served as stimuli. These videos featured a 'transgressor' accidentally harming a 'victim' and either displaying guilt or no guilt. One actress (Anya) always played the victim, and the other two actresses (Lisa and Susie) played the transgressors. Each video featured one target object: a doll, ball, clay bird, or picture.

All videos began with the three actresses seated around a table (see top panel of Figure 5). Anya excitedly told Lisa and Susie that she wanted to show them something, and then brought out and presented the target object for 45 s, as follows:

Doll. Anya said this was her favorite doll and happily showed off the doll's hair, eyes, etc.

Ball. Anya said this was her new ball and then happily played with it by throwing it in the air, rolling it on the table, etc.

Bird. Anya brought out a clay bird and happily talked about how she had made it and how pretty it was. After 20 s, she proudly added a tail feather with some more clay. When she was done, she again stated how pretty the bird was.

Picture. Anya brought out a colorful drawing of a butterfly and happily talked about how she had drawn it and how pretty it was. After 20 s, she proudly completed the drawing by adding the antennae. When she was done, she again stated how pretty the butterfly was.

The first 15 s of each video showed all three actresses, after which the camera zoomed in on Anya so as to show the target object more clearly (the transgressors were now out of view). Towards the end of Anya's 45-s presentation, the camera switched to a view of Anya and the transgressor assigned to that situation (the other transgressor remained out of view; see bottom panel of Figure 5). Anya now happily placed the target object on the table, and the transgressor then acted upon the object. For sake of simplicity, Lisa was always the transgressor in the doll and bird situations and Susie in the ball and picture situations. The transgressions thus proceeded as follows:

Doll. Lisa picked up the doll to admire it, but while she was playing with the doll's hair, its head came off and fell onto the table.

Ball. Susie played with the ball by throwing it up and catching it. The third time that she did this, she failed to catch the ball and it fell to the ground. When she reached under the table to retrieve it (out of view), the sound of something tearing was audible, and when she brought the ball back up, it was torn and the filling was spilling out.

Bird. Lisa picked up the clay bird to admire it, but as she touched one of the wings, it came detached and the rest of the bird fell onto the table.

Picture. Susie admired the picture but while returning it to Anya, accidentally tore it.

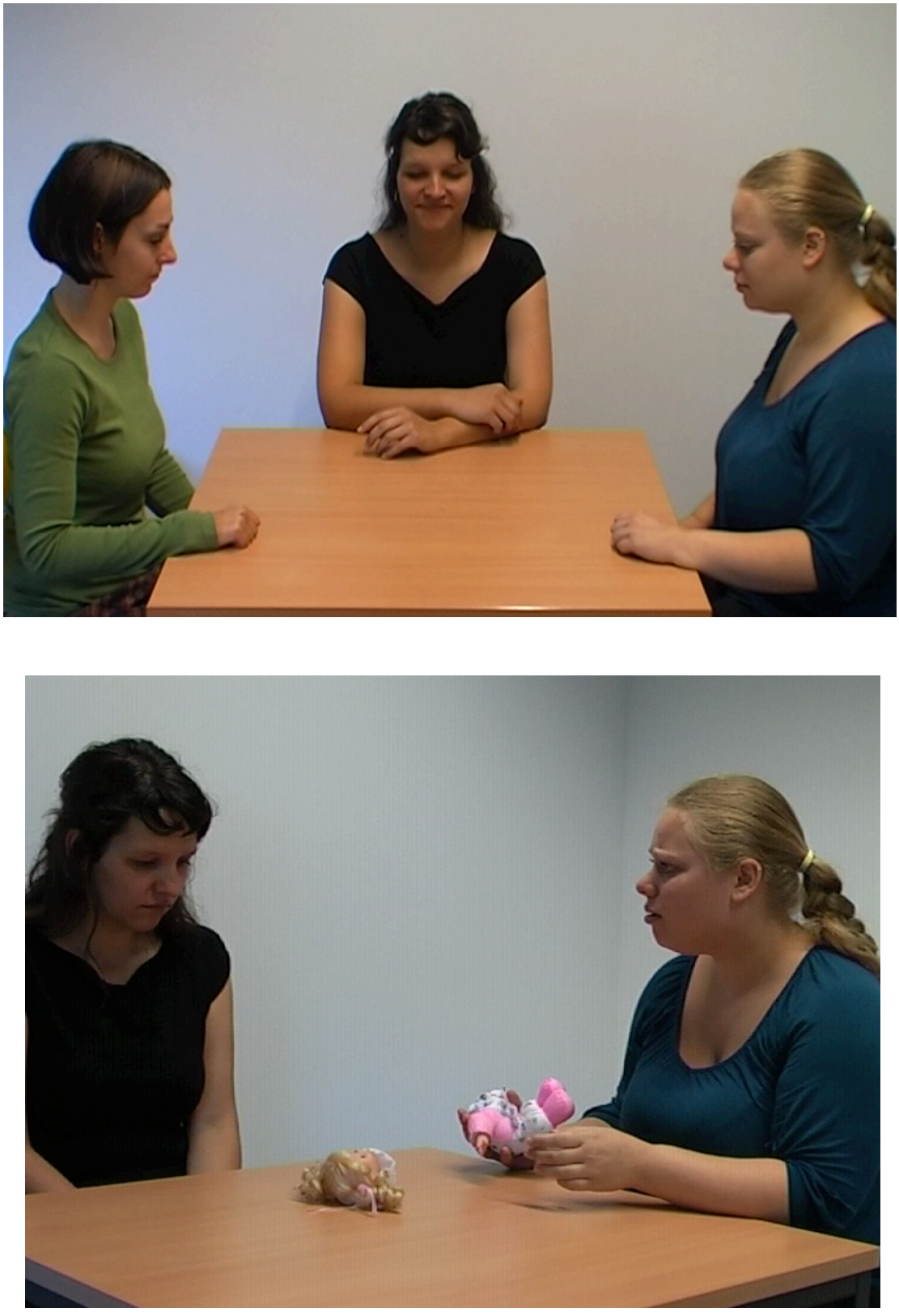

Figure 5. The camera angle used at the start of the video stimuli (top panel) and during the transgressors' actions (bottom panel; here, the transgressor [right] is showing guilt).

At the end of each incident, Anya said sadly, "Oh, my [target object]," and the transgressor then responded guiltily or non-guiltily. In the Guilt condition, the transgressor looked remorseful and concerned (with furrowed brow and concerned eyes; cf. Study 1, and Zahn-Waxler, Radke-Yarrow et al., 1992) and remorsefully said, "Oh, I've torn/broken your [target object]. I didn't want that to happen. It's my fault." While speaking, she alternated her gaze between Anya and the broken object, and placed her hand to her mouth once to display shock. She continued to look remorseful as she placed the object on the table. In the No Guilt condition, the transgressor looked neutral and said in a neutral tone of voice, "Yes, I've torn/broken your [target object]. Hmph [she shrugged], I don't care." She also alternated her gaze between Anya and the broken object while speaking, and then neutrally placed the object on the table. In both the Guilt and the No Guilt condition, Anya now picked up the object and looked at it sadly while the transgressor continued looking remorseful or neutral. At the end of the video (the average duration of each video was 2 min) was a still frame of this scene, which remained on the screen for 6 s. There were also shorter versions of all videos (starting just before Anya placed the object on the table), which ended with the same still frames.

For each target object, I created four videos in which the sides that the transgressors sat on and whether they showed guilt or no guilt were counterbalanced. For example, there were four videos of the doll situation: Lisa on the left showing guilt or no guilt, and on the right showing guilt or no guilt. There were thus 16 videos in all (four per target object), although each child only watched four of the 16 videos (one per target object). During testing, the doll and ball videos (both featuring the victim's possessions) were always presented together, as were the bird and picture videos (both featuring objects that the victim had created).

Procedure

Children were randomly assigned to one of 20 presentation orders (which counterbalanced the sides on which transgressors sat, which transgressor acted first, which transgressor showed guilt, and whether children saw the bird and picture or the ball and doll videos first). Other factors that were counterbalanced will be mentioned below. The computer on the left always showed the situations involving the transgressor on the left, and the computer on the right always showed the situations involving the transgressor on the right.

All children were tested by a female adult experimenter (E), who always sat to their left during the experiment (and who was not featured in the videos). E told children that she was going to show them videos of some people doing some things. They should watch carefully and she would then ask them some questions. E pulled up the first video assigned to the child, introduced the three characters, and played the video (e.g., a Guilt video of the ball situation on the left computer). At the end of the video, E paused the still frame and asked the first comprehension probe: "What did Lisa/Susie do to Anya's [target object]?" The child was expected to answer, "She broke it" or something similar. If the child answered with something less specific (e.g., "She dropped the ball"), E prompted the child further by asking, e.g., "And what happened to the ball?"

The first probe was to ensure that the child understood that the transgressor had damaged the object. Once the child's response indicated this, E said, "That's right" and asked the second comprehension probe: "How does she [pointing to relevant transgressor] feel now? Does she feel bad or does she not feel bad?" (Order of "feel bad" and "not feel bad" was counterbalanced across children). This second probe was to ensure that the child attended to and grasped how the transgressor felt, which was critical in order to test whether the child would draw any inferences on this basis. If the child answered correctly ("Bad" or something similar in the Guilt case; "Not bad" or something similar in the No Guilt case), E

said, "That's right. You've understood it correctly. Let's watch that last part again." E then played the shorter version of the video and paused the final still frame. If, however, the child answered the second probe incorrectly (e.g., "Bad" in the No Guilt case), E said, "Hmm, I'm not so sure about that. Let's watch that last part again and I'll ask you the questions again afterwards." E then played the shorter version of the video, paused the still frame, and repeated both comprehension probes as before. If the child still answered the second probe incorrectly, E corrected her by saying, "No, she felt/did not feel bad about breaking/tearing the [target object]." (Thus, regardless of whether or not children correctly answered the comprehension probes, all children saw each video entirely once and partially once.)

E then opened the second video, which was in the other condition and on the other computer (e.g., a No Guilt video of the doll situation on the right computer). She reminded the child of the characters' names, and then followed the same procedure as with the first video. Finally, after the child had seen both videos and answered the comprehension probes, E provided a reminder, for example: "So, Lisa/Susie [pointing to correct computer screen] broke Anya's [target object] and she feels bad about it, and Susie/Lisa broke Anya's [target object] and she doesn't feel bad about it." While providing this reminder (and throughout the procedure), E was careful to speak neutrally and not to nod or shake her head or in any other way provide evaluations of the transgressors. E then asked the following test questions:

1. Victim madder: "Whom is Anya madder at? – Susie or Lisa?" (pointing to each in turn)

1a. Victim madder-justification: "Why is she madder at her?"

2. Victim likes: "Whom does Anya like more? – Susie or Lisa?" (pointing to each)

2a. Victim likes-justification: "Why does she like her more?"

3. Child plays: "Whom would *you* prefer to play with? – Susie or Lisa?" (pointing to each)

3a. Child plays-justification: "Why would you like to play with her more?"

4. Meaner: "Who do you think is meaner? – Susie or Lisa?" (pointing to each)

4a. Meaner-justification: "Why do you think she is meaner?"

Questions 1, 2, 3, and 4 were forced-choice questions and children were expected to name and/or point to one transgressor. If a child responded "Both," E prompted her to choose one. Alternatively, if a child did not respond at all, E repeated the question once, but if the child still did not respond, E moved on to the next question. Questions 1a, 2a, 3a, and 4a were designed to elicit justifications for children's responses to the forced-choice questions. E thus let children respond freely to these questions and did not probe further. Following these eight test questions, E repeated the entire procedure and all of the questions with the second pair of videos (Phase 2).

For a given child, Phases 1 and 2 matched in terms of which transgressor showed guilt and which side each transgressor sat on. However, the order in which the transgressors' names appeared in the test questions and the order of the first three pairs of questions were counterbalanced across children and across the two phases for a given child. The fourth pair (the 'meaner' questions, similar to the moral character questions used by C. E. Smith et al., in press) appeared either first or last (counterbalanced across children). This was because, although it was feasible that children would not draw appropriate inferences about the transgressors' moral character (C. E. Smith et al., in press), if they did, it would be important to know that they were not influenced by their responses to the preceding questions.

Finally, after the second phase, E said that she would see Lisa and Susie soon and could take them something from the child. Then, in front of each computer, E placed a small container holding a photograph of the transgressor featured on the corresponding computer (the photographs featured the transgressors seated at the table and looking neutrally at the camera). E then gave the child three cloth flowers to distribute as she wanted. If the child did not distribute all the flowers or asked E for guidance, E encouraged her to decide for herself. When the child was done, E asked one final justification question, namely, why the child had

given two (or three) flowers to Lisa/Susie (whoever received more flowers). Again, E let children respond freely and did not probe further.

Coding and Reliability

A transcriber (blind to hypotheses) first transcribed children's verbal and/or pointing responses. From these transcriptions, the primary coder coded whether children responded correctly to the comprehension probes. Since E asked the second comprehension probe only after children answered the first probe satisfactorily, coding of responses to the first probe was only to make sure that E had followed this procedure and thus that all children understood that the target object was damaged. Coding of responses to the second probe assessed whether children grasped right away how the transgressor felt, whether they grasped it after re-watching the video, or whether E explicitly corrected them. For reliability, a second coder (blind to hypotheses) coded responses to the second comprehension probe for a random 25% of the sample. Reliability was perfect, $\kappa = 1$.

The primary coder also used the transcriptions to code children's responses to the forced-choice test questions (questions 1, 2, 3, and 4). Responses were scored '1' if they were consistent with the hypotheses that children should (1) judge that the victim is madder at the unremorseful transgressor, (2) judge that the victim prefers the remorseful transgressor, (3) themselves prefer to interact (play) with the remorseful transgressor, and (4) judge the unremorseful transgressor to have a worse moral character (i.e., to be meaner); responses not consistent with these hypotheses were scored '0.' A second coder (blind to hypotheses) coded a random 25% of the sample. Reliability was perfect, $\kappa = 1$.

The primary coder coded children's distribution of the three flowers from videotape and scored 0, 1, 2, or 3 to represent how many flowers children gave to the transgressor who displayed guilt. A second coder (blind to hypotheses) coded this measure for a random 25% of the sample. Agreement between coders was 100%.

Finally, the primary coder coded children's justifications (i.e., their responses to test questions 1a, 2a, 3a, and 4a, and their justification for the distribution of resources task) from the transcriptions and assigned scores of 0, 1, or 2 to each justification (see Table 6 for details of the coding scheme). The highest score (2) was assigned to justifications that referred to apologies or feelings of guilt, or involved moral evaluations. References to apologies received the highest score because in fact, neither transgressor apologized. Inferring an apology from a guilt display or no apology from the absence of a guilt display thus required an impressive grasp of the emotions behind apologies.

A '1' was assigned to all other justifications that were codable and relevant to the videos but were less sophisticated than the justifications that received a score of 2. A score of 0 was assigned to justifications that could not be placed into any other category, were irrelevant, or were uncodable. Justifications that could be assigned multiple scores were assigned the highest possible score. If a child did not provide a justification on a particular question, no score was assigned for that question. A second coder (blind to hypotheses) coded justifications of a random 25% of children. Reliability was excellent, $\kappa = .82$.

5.3 Results

I first report results of the comprehension probes in order to provide information about how well children understood the content of the videos. I then report children's performance on test questions and the distribution of resources task. Preliminary analyses revealed that for both age groups and for all measures, there were no significant effects of which side Lisa and Susie sat on, whether Lisa or Susie acted first, whether Lisa or Susie showed guilt, or whether a Guilt or No Guilt video was presented first. There were also no significant effects of gender but one: Among 4-year-olds, in Phase 1, 9 of 10 boys but only 3 of 9 girls chose the remorseful transgressor in response to the test question "Whom would you prefer to play with?" ($p = .01$). However, as this effect did not emerge in Phase 2, and as

no other gender effects emerged, gender was pooled for all subsequent analyses. All reported *p*-values are two-tailed.

Table 6. Coding scheme for justifications

Score	Category	Content
2	Apology	Transgressor did (or did not) apologize
	Guilt (re-described)	Transgressor did (or did not) feel bad about what she had done (child uses words other than those used in the videos or by E)
	Moral character, evaluation, or norm	Transgressor is a good (or bad) person, transgressor's act was good (or bad), or transgressor broke (or did not break) a moral norm
1	Guilt (repeated)	Transgressor did (or did not) feel bad about what she had done (child uses words that had been used in the videos or by E)
	Intention (re-described or repeated)	Transgressor intended (or did not intend) to cause harm or to damage the object (child may or may not use words that had been used in the videos or by E); e.g., "She didn't want that to happen" or "She didn't do it on purpose"
	Victim	Victim is sad or upset, or her belonging has been destroyed
	Own preference	Child expresses his or her own preference for the transgressor; e.g., "Because I like her better."
	Action	Transgressor damaged the [target object]
	Object	[Target object] is damaged or can no longer be repaired
0	Other, irrelevant, or uncodable	Response could not be put into any of the above categories (e.g., "Because that's how it is"), was irrelevant (e.g., "Because I always go to the zoo with my mother"), or was uncodable (e.g., because the child's speech could not be understood)

5.3.1 Comprehension probes

Comprehension probe 1

For all four videos, responses to the first comprehension probe (about what the transgressor had done) indicated that all children at both ages understood that the transgressors had damaged the target objects. Thus, all children grasped the basic premise of the videos.

Comprehension probe 2

5-year-olds. In Phase 1, when the transgressor showed guilt, 16 of 20 5-year-olds correctly identified her as feeling bad right away (binomial probability, $p = .012$). The remaining four children responded correctly after re-watching the video. When the transgressor showed no guilt, 11 of 20 5-year-olds correctly identified her as not feeling bad right away (binomial probability, using a test proportion of .50, $p = .824$). Eight of the 9 remaining children responded correctly after re-watching the video. Children's performance improved in Phase 2: When the transgressor displayed guilt, all 20 children correctly identified her right away as feeling bad, and when the transgressor showed no guilt, 17 of 20 children (binomial probability, $p = .003$) correctly identified her right away as not feeling bad and the other 3 responded correctly after re-watching the video. Altogether, in the Guilt case, 16 of 20 children responded correctly right away in both phases, and in the No Guilt case, 11 of 20 did so.

4-year-olds. The younger children struggled to grasp how the transgressors felt, especially when the transgressor was unremorseful. In Phase 1, when the transgressor showed guilt, a majority of 4-year-olds (14 of 20) correctly identified her as feeling bad right away (binomial probability, $p = .115$). Of the remaining 6 children, 5 responded correctly after re-watching the video. However, when the transgressor showed no guilt, a significant majority (16 of 20) initially responded *in*correctly (binomial probability, $p = .012$). Of those 16, 8

responded correctly after re-watching the video but 8 still responded incorrectly and had to be corrected by E. As with the 5-year-olds, performance improved in Phase 2: When the transgressor showed guilt, 16 of 20 children correctly identified her right away as feeling bad (binomial probability, $p = .012$), and 3 of the remaining 4 responded correctly after re-watching the video. When the transgressor showed no guilt, 9 of 20 children responded correctly right away (binomial probability, $p = .824$), and 7 of the remaining 11 responded correctly after re-watching the video. Altogether, in the Guilt case, 13 of 20 children responded correctly right away in both phases, but in the No Guilt case, only 2 of 20 children did so.

Age comparisons. I analyzed 4- and 5-year-olds' initial responses to the second comprehension probe (pooled across Phases 1 and 2) using chi-square tests. When the transgressor showed guilt, the number of children who responded correctly in both phases did not differ by age group, $\chi^2 [1, N = 40] = 1.13, p = .288$. However, when the transgressor showed no guilt, significantly more 5-year-olds than 4-year-olds responded correctly in both phases, $\chi^2 [1, N = 40] = 9.23, p = .002$. Still, the fact that a large number of children at both ages said that the transgressor felt bad even though the transgressor had shown no guilt points to the importance of checking children's comprehension and, when necessary, correcting them about the transgressors' feelings before going on to test their inferences about guilt-displaying versus non-guilt-displaying transgressors.

5.3.2 Test questions

Forced-choice questions

Preliminary analyses revealed no significant effects for either age group of whether or not children answered the comprehension probes correctly right away, or of whether the 'meaner' questions (questions 4 and 4a) appeared first or last. I thus did not include these variables in further analyses.

5-year-olds. The performance of the 5-year-olds was impressive: Children's responses to the 8 forced-choice test questions (4 per phase) indicate that they overwhelmingly drew all of the appropriate, hypothesized inferences in both the first and the second phase (the proportion of children who responded in the predicted way on each question ranged from .8 to .95; binomial probabilities, all *p*s < .013). For further analyses, I pooled together children's responses from both phases for each question. For example, responses to test question 1 in Phase 1 and test question 1 in Phase 2 were pooled, and a score of 1 was assigned for test question 1 if both responses drew the hypothesized inference; otherwise, a score of 0 was assigned. These more stringent analyses also indicated that the 5-year-olds drew all of the appropriate, hypothesized inferences (binomial probabilities, all *p*s < .013; see Figure 6). Furthermore, 19 of 20 children responded in the hypothesized way to a majority (5 or more) of the 8 forced-choice questions (binomial probability, $p < .0005$).

4-year-olds. In stark contrast to the 5-year-olds, the younger children's responses to the 8 forced-choice test questions indicate that they did not draw any of the hypothesized inferences in either phase (the proportion of children who responded in the predicted way on each question ranged from .35 to .65; binomial probabilities, all *p*s > .262). Pooling children's responses across phases for each question produced similar results (binomial probabilities, all *p*s > .166; see Figure 6). Altogether, only 7 of 20 children responded in the hypothesized way to 5 or more of the 8 forced-choice questions (binomial probability, $p = .263$).

Age comparisons. A chi-square analysis was conducted using pooled responses from both phases for each question. This analysis revealed that on all forced-choice test questions, significantly more 5- than 4-year-olds drew the hypothesized inference in both phases (all *p*s < .005; see Figure 6). Moreover, the proportion of the 8 test questions answered in the

hypothesized way was significantly higher among 5-year-olds (M = .91; SD = .20) than 4-year-olds (M = .51, SD = .28), $t(38) = 5.10$, $p < .0005$.

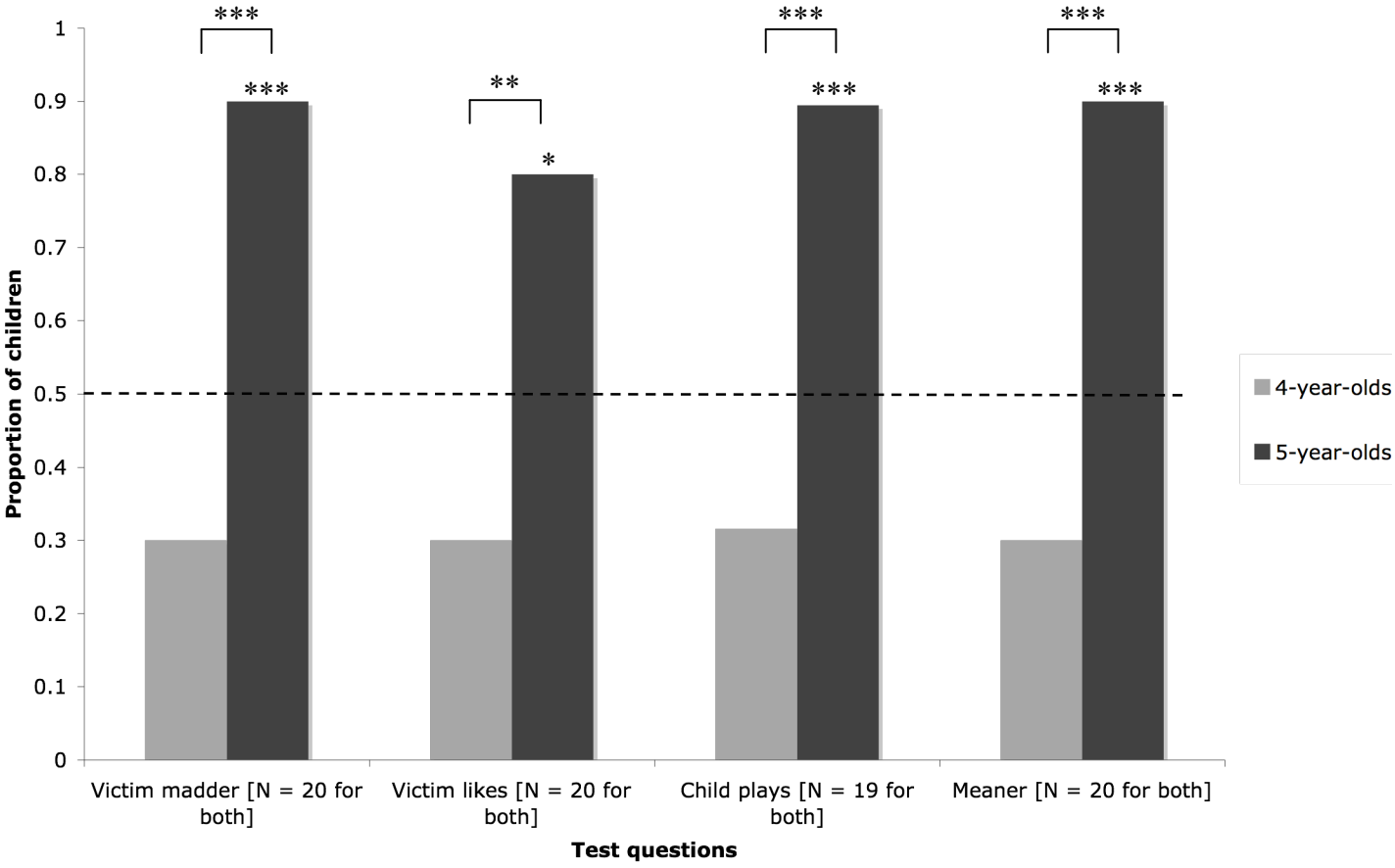

Figure 6. Proportion of children who answered each test question correctly in both phases. * $p < .05$; ** $p < .01$; *** $p < .0005$.

Justifications

Levels of children's justifications were compared across the two age groups. Justifications were only included in analyses if children had answered the preceding forced-choice test question in the hypothesized way (but including all justifications produced the same results). The majority of 5-year-old children (15 of 20) provided at least one highest-level (score of 2) justification across the 8 justification questions, indicating a sophisticated level of understanding and reasoning about the transgressors and transgressions. In the first phase, 11 of 20 children provided at least one such justification, and in the second phase, 13

of 19 children did so. Unlike the 5-year-olds, very few 4-year-olds (only 2 of 18) provided any level-2 justifications across the 8 justification questions. In the first phase, 2 of 15 children provided at least one such justification, and in the second phase, 1 of 16 children did so. (For both age groups, the details of each question in each phase are provided in Figure 7.)

The difference between the number of 5- versus 4-year-olds who provided at least one level-2 justification was highly significant, $\chi^2 [1, N = 38] = 15.64, p < .0005$. This was reflected in both the first and the second phases ($p = .012$ and $p < .0005$, respectively). Also, on average, 5-year-olds provided higher levels of justifications ($M = 1.75$, $SD = .44$) than did 4-year-olds ($M = 1.00$, $SD = .49$), Mann-Whitney $U = 60.0$, N5-year-olds = 20, N4-year-olds = 18, $p < .0005$.

5.3.3 Distribution of resources

Distribution

5-year-olds. Three of the 20 5-year-olds did not receive the distribution of resources task due to experimenter error. Of the remaining 17 children, 16 gave more flowers (2 or 3 out of 3) to the guilt-displaying transgressor (binomial probability, $p < .0005$).

4-year-olds. Similar to the 5-year-olds, a majority of 4-year-olds (12 of 20) gave more flowers to the guilt-displaying transgressor, but this proportion was not different from chance (binomial probability, $p = .503$).

Age comparisons. A chi-square analysis revealed that more 5- than 4-year-olds gave more flowers to the remorseful person, Fisher's exact test (due to small *N*s in some cells), $p = .023$. The mean number of flowers given to the remorseful transgressor was also higher among 5-year-olds ($M = 2.06$; $SD = .43$) than 4-year-olds ($M = 1.55$, $SD = .61$), $t(35) = 2.90$, $p = .005$.

Justifications

Levels of children's justifications were compared across the two age groups. Children's justifications were only included in analyses if children had distributed the resources as predicted (i.e., given more flowers to the remorseful person), although the results remained the same even when all justifications were included. Of the 16 5-year-olds who distributed the resources as predicted, one child was not asked the corresponding justification question due to experimenter error. Nine of the 15 remaining children provided level-2 justifications. Among the 12 4-year-olds who responded as predicted on the distribution of resources task, 5 provided level-2 justifications (see Figure 7 for details for both age groups). Unlike with the justifications for the test questions, the two age groups did not differ in the number of children who provided level-2 justifications on the distribution of resources task (p = .343). This was also reflected in an analysis comparing mean levels of justifications provided by the two age groups, p = .336.

5.4 Discussion

In this study, I tested the flexibility of young children's moral judgments and behavior. Specifically, I asked whether 4- and 5-year-old children judge and respond to a third-party transgressor who displays guilt differently from one who displays no guilt. The results show that 5-year-olds do so robustly: They know that the victim will be more upset with a transgressor who displayed no guilt than with one who did, and conversely, that the victim will like a transgressor who displayed guilt more than one who did not. Most 5-year-olds also say that they themselves would prefer to affiliate (in the form of play) with a remorseful transgressor than with an unremorseful one. Moreover, contrary to C. E. Smith et al. (in press), who found that 4- to 5-year-olds were not able to make inferences about the moral characters of apologetic versus non-apologetic transgressors, the results from the present study show that 5-year-olds are able to make such inferences about remorseful versus

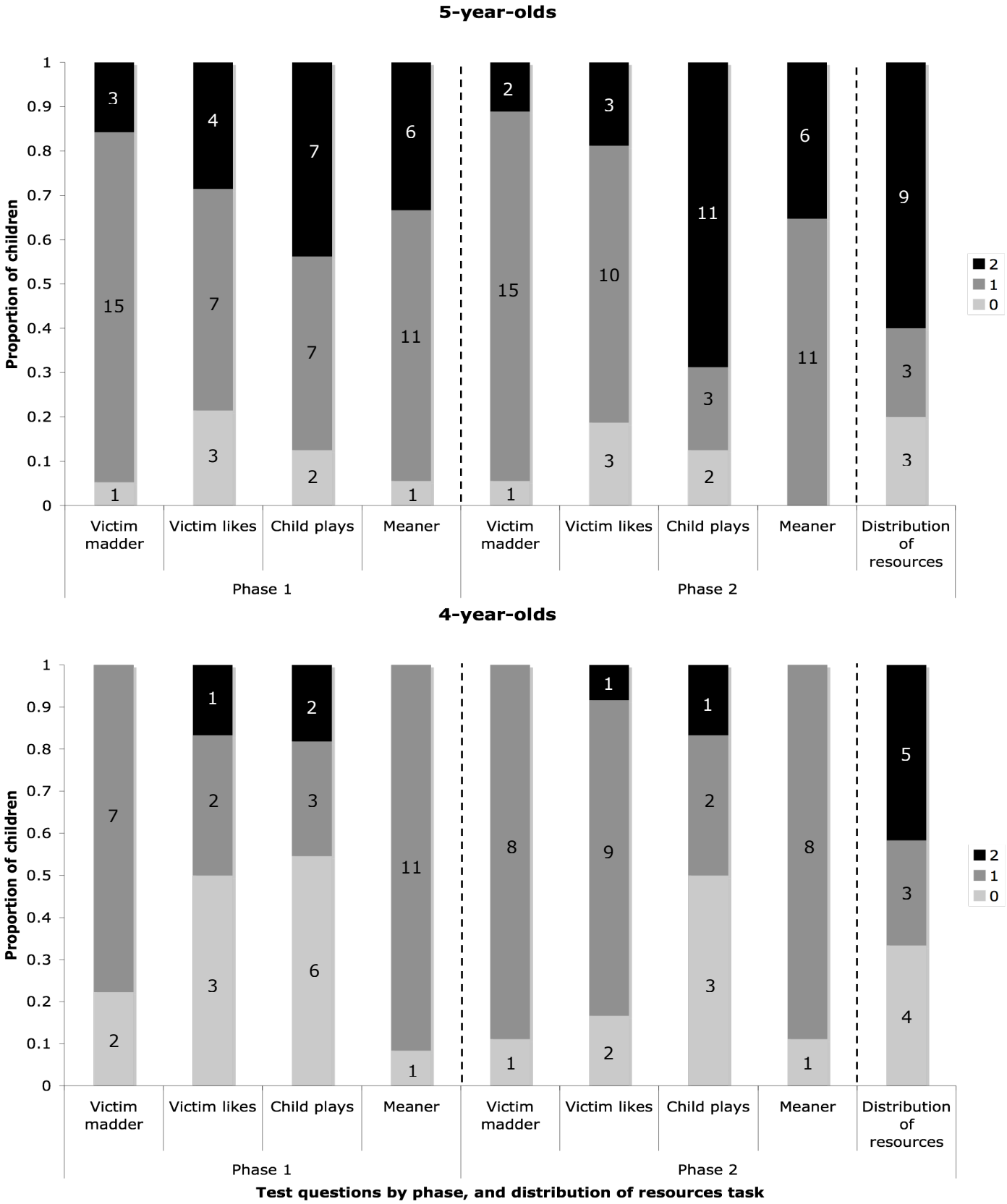

Figure 7. Proportion (y-axis) and number (within the bars) of children who gave each level of justification for the test questions in each phase and the distribution of resources task.

unremorseful transgressors. Importantly, 4-year-olds did not draw any of these inferences. Finally, 5-year-olds, but not 4-year-olds, showed a strong preference to cooperate with (by distributing more resources to) the remorseful than the unremorseful transgressor. These results show that between the ages of 4 and 5 years, children's moral judgments and behavior become sensitive to the appeasement functions of guilt displays.

Note that the actors in the videos did not apologize, so children in the present study could not have relied on apologies to draw their inferences. Thus, 5-year-olds seem to truly grasp the functions of displaying guilt rather than simply relying on whether an apology script was followed or not. Children's justifications also provide evidence for this idea, as several 5-year-olds stated that the guilt-displaying transgressor had apologized when in fact she had not offered an explicit apology. This suggests that at least some 5-year-olds grasp the true meaning behind apologizing, namely, to convey guilt and the appeasing information associated with it.

One rather surprising finding in the present study was that 5-year-olds were able to draw appropriate inferences about the moral character of the transgressors (test question 4), given C. E. Smith et al.'s (in press) finding that 4- to 5-year-olds were unable to do so based on whether a transgressor apologized or not. There are many possible reasons for this discrepancy. First, I gave children a forced-choice between two transgressors and asked them who was meaner. This may have been easier for children to respond to than C. E. Smith et al.'s question about whether their single transgressor was "nice or naughty most of the time." Second, since I did not ask who was meaner "most of the time," it is possible that children were in fact responding more specifically about who was meaner in the situations that they had witnessed rather than more generally (which would be related but not identical to evaluating overall moral character). Finally, my use of videos may have helped children understand the characters (the victim and transgressor) and their interactions more clearly

than the drawings and stories used in prior work did (e.g., Darby & Schlenker, 1982; C. E. Smith et al., in press; Wellman et al., 1979). This may have made it easier for children to make more sophisticated evaluations. All in all, it seems that at least under some circumstances, 5-year-olds (but apparently not 4-year-olds) can judge moral character at least to some degree based upon whether a transgressor has displayed guilt or not.

A question that arises is why 4-year-olds did not draw the same inferences as the 5-year-olds. One possible explanation emerges from the fact that the predominant response among 4-year-olds was to attribute guilty feelings to a transgressor who displayed no guilt (as seen in their responses to the second comprehension probe). It is plausible that children of this age are aware that they themselves would feel guilty in a similar situation and they project this onto an unremorseful transgressor (Darby & Schlenker, 1982). Relatedly, children may have a script that predicts that transgressors experience guilt (Arsenio, 1988; Harris, 1985), and perhaps the younger children in the present study were unable to inhibit this script enough to use the actual information presented to them. These ideas are compatible with research on the executive functioning of 3- and 4-year-olds, which shows domain-general difficulties including representational inflexibility and trouble switching between incompatible perspectives (e.g., Zelazo et al., 2003). In the present study, such difficulties may have prevented the 4-year-olds from grasping that one of the transgressors did not experience guilt, which would naturally make it difficult to distinguish between or draw inferences about a remorseful versus an unremorseful transgressor. Thus, the 4-year-olds' performance in the current study may be explained by general difficulties with complex executive functions rather than more specific difficulties with understanding guilt displays.

It is also plausible that the 4-year-olds would have drawn similar inferences as the 5-year-olds if the guilt display in the current study had included additional components that typically accompany guilt displays, such as apologies or attempts to repair the damage.

Indeed, 4-year-olds do judge stories in which a transgressor apologizes as better and more just than stories in which a transgressor does not apologize (Irwin & Moore, 1971; Wellman et al., 1979). Thus, perhaps the guilt displays employed in the present study were simply too subtle for these younger children to use effectively. Future studies could vary the strength of the guilt response to explore this question further.

A caveat regarding the current results concerns the potential confounding factor of intentionality. Specifically, the transgressor who displayed guilt stated that she "didn't want [the damage] to happen," whereas the transgressor who displayed no guilt did not say anything about her intentions. It is thus possible that 5-year-old children relied on this difference between the Guilt versus No Guilt situations rather than the guilt displays per se. Note, though, that acknowledging that one did not intend to cause harm is in fact a fundamental part of displaying guilt (Fessler & Haley, 2003). Moreover, given that guilt seems not to have a single, clear facial expression (Berti et al., 2000; Zahn-Waxler & Kochanska, 1990) and given that I excluded apologies and reparation attempts from the guilt displays, it seemed imperative to retain at least some verbal cues that are critical for conveying guilt. Also, an examination of children's justifications shows that children likely did not rely primarily on intentionality information. Specifically, only a small number of children (six 5-year-olds and one 4-year-old) referred to the transgressor's intentions in their justifications, and five of those six 5-year-olds and the one 4-year-old also produced higher-level justifications (e.g., apology), suggesting that these children drew their inferences based not only on intentionality information but also on other aspects of the transgressors' reactions. Finally, note that in Study 2, 3-year-olds did take a transgressor's intentions into account, as seen in their prosocial behavior. The fact that 4-year-olds in the current study did not differentiate between the transgressors in their judgments or their prosocial behavior suggests that they likely did not view them as differing in terms of their intentions.

Altogether, then, it seems unlikely that children in the present study relied solely on intentionality information to draw their inferences.

Let us briefly consider the evolutionary roots of appeasement displays. Nonhuman species, especially great apes and monkeys, produce appeasement displays following agonistic encounters. These displays include submissive behaviors such as postural contraction, gaze aversion, and submissive vocalizations, and affiliative behaviors such as lip smacking and kissing, and they serve appeasement functions by eliciting embracing, grooming, and sexual play from conspecifics, which counteract aggressive tendencies and increase social reconciliation (e.g., de Waal, 1986, 1988). Such displays are thus considered homologues of human appeasement displays, although among humans, these displays have become more differentiated and self-conscious, and have evolved into more symbolic concepts that guide social practices and behavior rather than being tied to physical, concrete events (Keltner & Anderson, 2000; Keltner & Buswell, 1997; Keltner & Haidt, 1999). Still, the fact that the foundations of appeasement displays seem to be present in other primates suggests deep evolutionary roots of such displays, and points to the critical functions that these displays likely serve in restoring cooperative relationships and thus in maintaining cooperation in groups.

In conclusion, displaying guilt serves important functions: It appeases the victim and other group members, restores interpersonal relationships, and indicates a willingness on the part of the transgressor to perform better and be a more cooperative group member in the future (e.g., Baumeister et al., 1994; Castelfranchi & Poggi, 1990; Darby & Schlenker, 1989; Goffman, 1967; Keltner et al., 1997). A remorseful transgressor should thus elicit more forgiveness, affiliation, and cooperation from the victim and from other group members than an unremorseful transgressor. As such, displays of guilt (and other self-conscious emotions) are mechanisms that enable people to maintain the stability of moral communities and

relationships, and to restore cooperation in groups, but to do so flexibly and discriminately, i.e., while still being able to identify non-cooperative individuals that deserve to be punished. The results of the present study show that already by 5 years of age, children are sensitive to some of these functions of guilt displays (even in the absence of explicit apologies), and this sensitivity does indeed impact their moral judgments and cooperative behavior. However, just a year earlier, children do not show the same sensitivity. Why this understanding emerges around 5 years of age remains an open question, but the present study makes clear that even preschoolers are already sophisticated, flexible, and discriminating moral beings.

6. GENERAL DISCUSSION AND FUTURE DIRECTIONS

The purpose of this dissertation was to explore the ontogenetic origins of some essential aspects of human morality. In four studies, I explored young children's understanding of and their responses to moral transgressions and the victims and transgressors therein. Together, the studies revealed the emergence of a sophisticated, flexible, and differentiated morality early in human ontogeny, as evident in children's moral emotions, behavior, and judgments. Below, I will discuss the findings and point to future research directions.

6.1 Concern for and prosocial behavior towards victims

One of the major findings of the present dissertation is that early in development, children are concerned about and attempt to help the victims of moral transgressions. In Study 1, 18-month-old and 2-year-old children were found to sympathize with a victim even though she displayed no emotional cues, arguably through some form of affective perspective-taking. Moreover, children subsequently showed greater prosocial behavior towards the victim, and children's sympathy for the victim correlated with their prosocial behavior towards the victim. Similar findings also emerged in Study 3, wherein, despite the victim's absence during the moral transgressions (and thus despite the lack of distress cues from the victim during the transgressions), 3-year-old children subsequently acted prosocially towards her upon her return.

These findings are impressive for two reasons. First, they show that even very young children are concerned about victims and act prosocially towards them. This conclusion is in line with prior work, which shows that infants and young children are generally prosocial: They empathize and sympathize with those in distress, as well as instrumentally help, comfort, and share resources and information with others (Brownell et al., 2009; Dunn, 1988; Eisenberg & Fabes, 1998; Warneken & Tomasello, 2007, 2009b; Zahn-Waxler, Radke-

Yarrow et al., 1992). The evidence is thus abundant now for an early ontogenetic emergence of prosocial motives and behavior in humans.

It has been proposed that prosocial motives and behavior not only have deep ontogenetic roots but also deep evolutionary roots (see de Waal, 2009; Warneken & Tomasello, 2009b). Certainly, the evidence for prosocial behavior in our nearest primate relatives is accumulating fast. Chimpanzees retrieve out-of-reach objects for humans and assist unrelated conspecifics in obtaining food (Warneken et al., 2007; Warneken & Tomasello, 2006), they show reciprocity by being somewhat more likely to help those who had previously helped them (de Waal, 1997; Melis et al., 2008), and they may be sensitive to some aspects of fairness (Brosnan & de Waal, 2003; Brosnan et al., 2005).

Regarding prosocial motives such as empathy and sympathy, however, systematic evidence is still needed. Some researchers claim that monkeys and apes possess the core elements of empathy, such as emotional contagion, and that apes further possess some of the cognitive elements, such as perspective-taking, that allow them to engage in true, other-oriented empathy (e.g., de Waal, 2009). However, these claims are based primarily on anecdotal evidence; to my knowledge, systematic experimental evidence for these claims is missing. In an ongoing study, Katja Liebal, Michael Tomasello, and I are attempting to fill precisely this gap by assessing sympathy in all four non-human great ape species. In a set-up very similar to Study 1, we presented a 'recipient' ape with food and allowed him to eat some of it but then pulled the rest away (harm condition), or we allowed him to eat all the available food (neutral condition). An 'observer' ape (equivalent to the child participant in Study 1) watched this interaction. Subsequently, in a separate situation, we presented the recipient with food that he could only reach using sticks. However, we only provided sticks to the observer, not to the recipient, thus creating an opportunity for the observer to act prosocially towards the recipient. The working hypothesis was that if observers sympathize with the

recipients in the harm condition more than in the neutral condition, and if sympathy motivates their prosocial behavior (as in humans), then observers should show greater prosocial behavior following the harm situation than following the neutral situation. Although data analyses for this study are still underway, preliminary results suggest that at least some ape species do engage in greater prosocial behavior in the harm condition than in the neutral condition. All in all, then, critical elements of morality (prosocial behavior and prosocial motives) might indeed have deep ontogenetic and evolutionary roots, supporting the claim that they are biologically-based human endowments (Hoffman, 2000; Warneken & Tomasello, 2009a, 2009b).

The findings that young children are concerned about victims even when the victims show no distress cues are remarkable for a second reason, namely, that they demonstrate the flexible nature of empathy-related processes even in young children. As noted by Hoffman (2000), empathic arousal is a reliable prosocial motivator because it is multi-determined: The many modes of empathic arousal enable observers to respond empathically to whatever distress cues are available. Typically, victims are present and providing clear distress signals; in this case, any or several of the arousal mechanisms may be operating, ranging from the most basic and automatic forms of mimicry to the most advanced forms of cognitive reappraisal. However, when a victim is absent or is not providing distress signals for some reason, the cognitively advanced modes of empathic arousal still enable one to empathize, thus adding scope to one's empathic capability (Hoffman, 2000). Indeed, adults do empathize and sympathize with victims, both when they have direct access to the victims' distress cues and when they must engage more sophisticated cognitive processes, and these empathic processes do motivate their prosocial behavior (Batson et al., 1991; Batson et al., 1981; Decety & Jackson, 2006; Ruby & Decety, 2004; Singer & Lamm, 2009). The findings of Studies 1 and 3 show that this multi-determined nature of empathy-related responses is

already functional in early development. This suggests that empathy-related responses are indeed reliable proximate mechanisms underlying moral behavior even early in human ontogeny. An interesting and important question that arises here is, *why* do empathy-related responses motivate prosocial behavior? I consider this question next.

6.1.1 Why do empathy-related responses motivate prosocial behavior?

From the perspective of ultimate mechanisms (i.e., how did this phenomenon contribute to our survival and reproduction evolutionarily?), this question has a rather straightforward response: To the degree that prosocial motives result in prosocial behavior, they benefit the self by inducing reciprocity (direct and indirect), benefit one's genes (inclusive fitness and kin selection), or benefit one's group (group selection). From the perspective of proximate mechanisms, however, the question is more challenging. One proposal in this regard is that empathic responses are unpleasant and we intervene prosocially in order to alleviate our own rather than the other's unpleasant state. An alternative but similarly egoistic proposal is that acting prosocially feels good, i.e., it is a self-rewarding behavior. However, these egoistic explanations overlook the distinction between the consequences of an act and its ultimate goal: Just because a person feels better after acting prosocially does not mean that she acted prosocially with the ultimate goal of feeling better (see Batson & Shaw, 1991; Hoffman, 1981). The philosopher David Hume (1776/1965) similarly argued that benevolent acts and sentiments cannot simply be reduced to self-interest because we may derive joy from making others happy and we may even act from the combined motives of benevolence and enjoyment, but that does not make our benevolence identical to our self-enjoyment (see Chapter 1). Indeed, empirical work has convincingly shown that egoistic gain is not typically the ultimate goal of empathy-based prosocial behavior. For instance, empathically aroused adults choose to help even when they could easily escape the aversive situation both physically and psychologically, and even when there

is a cost-free alternative that would enhance their mood (Batson et al., 1981, 1988; Stocks, Lishner, & Decker, 2009).

Furthermore, as Hoffman (1975, 2000) points out, the fact that our own welfare so profoundly depends on the welfare of others might actually be the critical mechanism underlying prosocial motivation. That is, empathic responses might function as prosocial motivators precisely because they transform another person's misfortune into one's own feeling of distress, thus intricately connecting the alleviation of others' distress to the alleviation of one's own distress. Indeed, at the level of the individual, this may well be the way in which the empathy-prosocial behavior link is propagated and maintained.

All in all, empirical and theoretical work suggests that empathic responses motivate adults to act prosocially out of a genuine concern for the welfare of the other. Although this is hypothesized to be true of young children's prosocial behavior as well (Hoffman, 2000; Warneken & Tomasello, 2009a), the corresponding empirical work (e.g., presenting children with a distressed individual but providing them with the option of an easy escape from the aversive distress cues) has not yet been conducted. Future work should thus assess *why* children help those with whom they sympathize: primarily for egoistic or primarily for prosocial reasons.

6.2 Involvement and intervention in third-party transgressions

A second major finding of the present dissertation is that young children are deeply concerned with and involved in third-party moral interactions. In all four studies, children were witnesses to (rather than active participants in) moral transgressions. That is, in none of the studies were children harmed in any way, nor did they harm anyone. Yet they not only paid close attention to the interactions but they also became actively involved in those interactions, affectively, behaviorally, and through their evaluations.

The most striking finding in this regard was that even young children demonstrated some forms of third-party punishment. Young children thus not only defend themselves against moral violations by protesting, retaliating, tattling on the perpetrator, and so on (Dunn, 1988; Dunn & Hughes, 1998; Dunn & Munn, 1987; Ingram & Bering, in press), but they also engage in such behaviors in the context of third-party transgressions, when they themselves have not been harmed. This was evident in children's selective prosocial behavior: When given a choice between instrumentally helping a perpetrator and a neutral person, 3-year-old children generally chose not to help the perpetrator (Studies 2a and 2b). Children thus engaged in a common form of third-party punishment, namely, withdrawing cooperation from perpetrators (Boyd & Richerson, 2005). Two other forms of third-party intervention emerged in Study 3: (i) When a victim puppet was absent during a moral transgression, 3-year-old children actively intervened (by verbally protesting) against the transgression, presumably to prevent it from occurring, and (ii) upon the victim's return, they tattled on the perpetrator to the victim.

It is, of course, difficult to pinpoint why children intervened in these various ways. For instance, the reduced prosocial behavior in Study 2 may have been the result of active withdrawal of cooperation or of more passive shunning (i.e., punishment through partner choice and thus at no cost to the punisher; Boyd & Richerson, 2005). Similarly, in Study 3, children may have protested against the transgression out of concern for the victim or because the moral norm that prohibits destroying others' belongings was being violated. Also in Study 3, children may have tattled about the transgressor to the victim in order to ensure that the transgressor would be punished or simply to inform the victim that they were not responsible for the transgression. Although more work is needed to understand the mechanisms behind children's interventions, what becomes clear when considering Studies 2 and 3 together is that third-party intervention, which is a critical proximate mechanism for

maintaining large-scale human cooperation (Boyd & Richerson, 1992; Fehr & Fischbacher, 2003, 2004), is functional quite early in development.

As discussed previously (see Chapter 1), involvement and intervention in third-party (rather than dyadic) transgressions are thought to be strong evidence for moral understanding, as they demonstrate agent-neutral (rather than self-interested) applications of moral norms. Indeed, philosophers such as Hume, Kant, Nagel, and Rawls have long argued that agent-neutral morality, which is achieved by moving away from an egocentric perspective and assuming an impartial point of view, is the essence of true morality (see Chapter 1). The studies in this dissertation suggest that fairly early in ontogeny, humans already possess at least some degree of agent-neutral morality.

Third-party intervention is also considered essential in evolutionary analyses of the origins of morality because (as discussed in Chapter 1) it is argued that intervention in dyadic interactions alone would likely not sustain the kind of large-scale cooperation seen among humans (Boyd & Richerson, 1992; Fehr & Fischbacher, 2003, 2004). Indeed, third-party intervention (especially punishment) is thought to be unique to humans (Gintis et al., 2003; Krebs, 2008; Tomasello, 2009). As part of her doctoral work, for instance, Katrin Riedl (in collaboration with Keith Jensen, Josep Call, and Michael Tomasello) has been comparing third-party punishment in chimpanzees and children. In one study, chimpanzees could punish food theft committed by a conspecific directed either towards another group member or towards themselves. Riedl and colleagues replicated an earlier finding that chimpanzees retaliate against theft that directly affects them (K. Jensen et al., 2007), but they found no evidence for third-party punishment, even when the victim was related to the punisher, and when the punisher outranked the thief. On the other hand, in an ongoing study using a similar procedure with 3-year-old children, Riedl and colleagues are finding (consistent with the results of Studies 2 and 3) that children respond to third-party theft by returning stolen toys or

food to a victim puppet. These results suggest that fairly early in ontogeny, humans respond to transgressions that affect others, whereas chimpanzees do not, even as adults. Third-party punishment may thus indeed be uniquely human[6]. In this dissertation, I demonstrated that this essential (and possibly unique) aspect of human morality is evident in early ontogeny.

The ontogenetic emergence of third-party intervention still needs further research, however. The studies in this dissertation show that as early as 3 years of age, children engage in multiple forms of third-party intervention. Whether children younger than age 3 also engage in similar intervention remains a question for future research. There is some reason to believe that younger children at least pay attention to and evaluate third-party interactions. For instance, 6-month-old infants discriminate helpful from hindering characters and prefer the helpful characters (Hamlin et al., 2007), and according to a more recent study, even 3-month-old infants evaluate negative (but not positive) social interactions (Hamlin, Wynn, & Bloom, in press). These findings could be seen as evidence of an extremely early (or even innate; see Wynn, 2008) human tendency to evaluate third-party interactions, but whether infants this young punish third-party transgressors (by withdrawing help, for instance) has not been explored thus far. Given that punishing in such contexts requires understanding not only what constitutes transgressions but also what constitutes punishment as well as having the motivation and physical capability to punish in some way, I would not expect third-party punishment to be present this early in human ontogeny. This remains for future work to assess, however.

[6] However, third-party intervention of other kinds does seem to be present in other species. For instance, some researchers have found evidence for "third-party policing" (i.e., physically impartial intervention into conflicts) among primates, and argue that such policing serves important functions for the group including preventing the injuries and damaged relationships that typically result from conflicts (Flack, de Waal, & Krakauer, 2005; Flack, Girvan, de Waal, & Krakauer, 2006). Open questions thus remain about why non-human species police but seem not to punish (although alternatively, perhaps they do punish but researchers have not yet found the appropriate way to test this).

6.3 Young children demonstrate flexible moral understanding

The studies in this dissertation point to a third important conclusion, namely, that morality is not rigid but surprisingly flexible, even early in development. That is, even young children's morality is not driven by simple rules such as "Always help people who need help" or "People who cause harm are always bad." Rather, early morality seems to be sophisticated, and if it is driven by rules, then it at least seems to be driven by complex and multifaceted rules. The evidence for this was manifold across studies.

6.3.1 Flexible prosocial behavior

First, it emerged that children's prosocial behavior is selective fairly early in ontogeny. Thus, at 18 months, 2 years, and 3 years of age, children are more prosocial towards a victim than a non-victim (Studies 1 and 3). At 3 years of age, children also help a moral transgressor less than a non-transgressor (Study 2). What's more, this reduced helping is not just a function of whether an actor transgressed by causing a harmful outcome; 3-year-old children also reduce their helping towards actors who intended to cause harm, even if they were unable to do so (Study 2b), and 5-year-old children selectively help a remorseful transgressor over an unremorseful one (Study 4). These findings are consistent with some prior work on children's prosocial behavior. For instance, 3-year-olds share toys with a peer more if that peer had previously shared toys with them, suggesting a sensitivity to reciprocity by this age (Levitt et al., 1985; see also Olson & Spelke, 2008). By about 4 years of age, children share with another child (even at a cost to themselves) if the other child is their friend rather than a non-friend or a stranger (Birch & Billman, 1986; Moore, 2009). The picture emerging from these prior findings and the current findings makes evolutionary sense. As Krebs (2006) notes: "Evolutionary theory leads to the expectation that dispositions to engage in indiscriminate altruism should not evolve" (p. 48). That is, for cooperation to evolve and be maintained, safety mechanisms must also exist that prevent one from being

exploited by others and that bias altruism towards certain individuals under certain circumstances (Warneken & Tomasello, 2009a).

Interestingly, however, it has previously been argued that children start out as indiscriminate helpers, and only as they develop preferences, become sensitive to context, and acquire norms do their early prosocial impulses slowly transform into more deliberate, selective, and morally informed choices (Hay & Cook, 2007). This idea fits with the evolutionary story that children are naturally and indiscriminately cooperative while they are still under the care and protection of their close kin who, even if not always trustworthy, at least ensure that any cooperation will have inclusive fitness benefits. According to this claim, selective cooperation becomes important later in childhood, when children begin to independently deal with strangers and must thus distinguish between cooperative and non-cooperative partners (Tomasello, 2009; Warneken & Tomasello, 2009a).

The present studies cannot speak to the issue of whether very young children (below 18 months of age) are indeed indiscriminate in their cooperation or not. Prior work suggests that at least in the domain of simple, instrumental helping, 14- to 18-month-old infants are rather indiscriminate helpers (see Warneken & Tomasello, 2009a). At the same time, 6-month-old infants seem able to discriminate a helpful from a hindering character and to prefer the helpful character (Hamlin et al., 2007), although whether infants this young would selectively help the hindering character less has not thus far been tested. However, the studies in this dissertation show that certainly by the second half of the second year, children's prosocial behavior is becoming selective. The ontogeny of this selectivity needs further examination, from the perspective of the social and cognitive skills required to be selectively cooperative (Hay, 1994) as well as from the perspective of evolutionary theory (Warneken & Tomasello, 2009a).

6.3.2 Flexible moral evaluations of and behavior towards transgressors

Early moral flexibility was also revealed in terms of the factors that young children took into account in their moral evaluations and behaviors. Study 2a showed that 3-year-olds understand basic moral transgressions and they selectively reduce their helping behavior towards transgressors. Study 2b, however, showed that this understanding is more sophisticated than a simple response to whether or not harm was caused, as 3-year-olds in that study took into account a transgressor's intentions, only reducing their prosocial behavior towards a transgressor who intended but was unable to cause harm, not towards a transgressor who unintentionally caused harm. Several prior studies have shown that Piaget (1932/1997) greatly underestimated the ability of children younger than 11 to 12 years of age to consider a transgressor's intentions in their evaluations (Grueneich, 1982; Imamoğlu, 1975; Yuill & Perner, 1988; Zelazo et al., 1996). However, Study 2b provided the first evidence that children even younger than 5 years of age take into account not only the outcomes of but also the intentions behind moral transgressions.

Moreover, Study 4 revealed that children also take a transgressor's subsequent reaction into account in their moral evaluations and behaviors. Displays of guilt (and other self-conscious emotions) are thought to be mechanisms that enable people to maintain the stability of moral communities and relationships, and to restore cooperation in groups, but to do so flexibly and discriminately, i.e., while still being able to identify non-cooperative individuals that deserve to be punished. The results of Study 4 show that a transgressor's display of guilt does indeed appease 5-year-old children in just these ways: They judge the victim as being more upset at the unremorseful transgressor and as liking the remorseful transgressor more, and they themselves prefer to affiliate with the remorseful transgressor and judge the unremorseful transgressor as being meaner. Moreover, they cooperate more with (by distributing more resources to) the remorseful transgressor.

Several other studies have also assessed the role of various factors in children's moral judgments, and have revealed impressive results. Children around 4 to 5 years of age have been found to take into account factors such as whether a transgressor apologized and whether she had already been punished (D. T. Miller & McCann, 1979; C. E. Smith et al., in press; Wellman et al., 1979). Children also take into account factors in the psychological domain such as the victim's and the children's own relationship with the perpetrator (Slomkowski & Killen, 1992; Wellman et al., 1979). In line with and extending this prior work, Studies 2b and 4 clearly show that children's moral evaluations and behavior are sophisticated and flexible. All in all, then, across the four studies in this dissertation, I found strong evidence for flexible moral understanding in early ontogeny.

6.4 The negativity bias in moral development

In this dissertation, I largely focused on children's understanding of and responses to moral transgressions rather than morally positive situations, such as people helping or sharing with others. However, in the one study in which I did assess a third-party helping situation (Study 2a), I found that children's prosocial behavior did not increase towards the helper when compared to their prosocial behavior towards a neutral person. This finding is consistent with the well-documented phenomenon of a negativity bias, that is, a greater impact of negative than of positive information. In adults, this bias has been documented in numerous domains (e.g., Baumeister et al., 2001) and more recently, it has also been documented in infants' and children's social-emotional development (Vaish et al., 2008). It is therefore quite plausible that it also exists in the moral domain.

Indeed, research on children's moral judgments indicates that children correctly identify "bad" acts substantially earlier than "good" acts (Hill & Hill, 1977; Rhine et al., 1967). Young children also seem to judge the valence of ill-intentioned acts in a sophisticated manner before being able to do so for well-intentioned acts. For example, L. C. Jensen and

Hughston (1973) presented preschoolers with stories of children whose good, neutral, or bad acts had either a negative social consequence (spanking from the mother) or a positive social consequence (receiving a pleasant surprise from the mother). The results indicated that children judged bad acts as bad irrespective or whether they were punished or rewarded, whereas children's evaluations of good and neutral acts relied to a large extent on the social consequences that followed those acts (see also Costanzo, Coie, Grumet, & Farnill, 1973; Nobes, Panagiotaki, & Pawson, 2009). In addition, Leslie, Knobe, et al. (2006) found that children judge the side effect of an action to have been caused intentionally if that side effect is morally bad but not if it is morally good (see also Leslie, Mallon, et al., 2006). Aloise (1993) also showed a negativity bias in trait attribution wherein children required fewer negative behaviors to infer negative traits (such as rude or rough) about other people than positive behaviors to infer positive traits (such as polite or gentle). These findings, as well as the findings of Study 2a, point to an exciting new direction for research: the examination of a negativity bias in children's social-moral development. In this research, it will be important not only to trace and describe the domains of social-moral development in which the negativity bias is present (or absent) but also its ontogenetic emergence and its functions.

In a recent review paper (Vaish et al., 2008), my colleagues and I argued that the negativity bias may have its ontogenetic roots in the early positive experiences that most young infants have with their caregivers from birth onwards. Such experiences establish a positive background against which negative experiences stand out, demand more attention, and have a greater impact than positive experiences. Similarly, the ontogeny of a negativity bias in social-moral development might also lie in the primarily positive and prosocial interactions that most infants and young children have. Such interactions likely create a positive background against which subsequent negative, harmful behaviors stand out and have a greater influence on children's emotions, behaviors, and judgments. In addition, it has

been proposed that since parents tend to be more concerned with inhibiting undesirable behavior than with promoting commendable behavior, they are more likely to punish than to reward. Children thus have more experience with the social consequences of negative acts than those of positive acts, and therefore learn the notion of what is good later than the notion of what is bad (see Karniol, 1978; Piaget, 1997).

Regarding the functions of the negativity bias, it has been proposed that this bias serves the evolutionarily adaptive purpose of allowing individuals to safely explore the environment while appropriately avoiding harmful situations (e.g., Cacioppo, Gardner, & Berntson, 1997, 1999; Vaish et al., 2008). I submit that the negativity bias in the social-moral domain likely serves a similar function. For instance, it seems critical to be able to effectively identify harmful or non-cooperative individuals (especially if most people are predominantly helpful and cooperative) such that one can avoid harmful or negative interactions with them in the future. In sum, I propose that future work should aim not only to systematically describe the negativity bias in social-moral development but also to consider the ontogenetic roots and functions of this bias.

6.5 Individual- and group-level differences

It is noteworthy that in none of the studies in this dissertation did children perform uniformly. For instance, in Study 1, 40% of children showed concern for the victim in the harm condition, but the rest were not coded as showing concern. Similarly, in Study 2a, 78% of children did not help the transgressor, but 22% still did. Individual differences were thus apparent in children's morality. Furthermore, the samples included in this dissertation were all drawn from the same population of middle- to upper-class Germans. This raises the question of whether the moral emotions, judgments, and behavior observed in the present studies, and morality more generally, generalize across groups and cultures. Obviously, the

issue of individual- and group-level differences cannot be addressed by the studies in this dissertation. I would nevertheless like to briefly consider the issue here.

6.5.1 Individual-level differences

Individual differences in morality are evident quite early in child development. For instance, individual differences in empathic concern and prosocial responses to others' distress can be detected as early as 14 months of age (Zahn-Waxler, Robinson et al., 1992). Individual differences in spontaneous sharing with peers consolidate between 30 and 36 months of age, and individual differences in sharing with mothers may emerge even earlier (see Hay, 1979; Hay & Cook, 2007). Such individual differences are especially relevant because they predict children's later prosocial tendencies and general social adjustment (e.g., Eisenberg et al., 1999). The interesting question, of course, is, whence this individual variation?

Numerous factors account for individual differences in morality. Biological factors such as genes, temperament, brain development, and neuroendocrine development have all been shown to significantly impact prosocial development (e.g., Knafo, Zahn-Waxler, Van Hulle, Robinson, & Rhee, 2008). Gender has also been found to play a role such that girls show greater empathy and prosocial behavior than boys, although the evidence for this is equivocal (see, e.g., Eisenberg & Lennon, 1983; Zahn-Waxler, Robinson et al., 1992). Several developmental processes, such as the emergence of language, increasing understanding of others' mental states, as well as the development and consolidation of a self, also contribute to moral development (see, e.g., Baird & Astington, 2004; Bischof-Köhler, 1991).

Additionally, parenting and disciplining styles are critical in shaping moral development. For example, Hoffman (2000) has proposed that when a child harms someone, the most effective way to discipline that child is by induction, wherein parents highlight the

other's perspective, point to the other's distress, and make it clear that the child's action caused it. This induces the child to feel empathic distress and guilt, helps the child realize the harm her behavior may cause another, and weigh that harm against her own desires; over time, this leads the child to construct an internalized norm of considering others. Note that this technique takes as a given that from very early on, infants and children are naturally empathic and cooperative; parents thus simply need to draw children's attention to the appropriate aspects of the situation in order to initiate the processes of empathy, guilt, and moral internalization (see also Warneken & Tomasello, 2009b). This, Hoffman argues, is more effective than techniques that rely on extrinsic motivators, power-assertion, or love-withdrawal in bringing about moral internalization of the sort desired in society.

Relationships with siblings and peers may also contribute to individual differences in prosocial and cooperative behavior (e.g., Brownell, Zerwas, & Balaraman, 2002; Ross, 1996), although the nature of this influence is not clear-cut. For instance, children with older siblings show an accelerated development of a theory of mind (Ruffman, Perner, Naito, Parkin, & Clements, 1998). Based on these findings, one might predict that they should also show accelerated moral development (see Baird & Astington, 2004). However, children with older siblings seem to be less likely to share with peers and more likely than other toddlers to respond negatively to (e.g., by aggravating or being amused by) peers' distress (Demetriou & Hay, 2004; Hay, Castle, Davies, Demetriou, & Stimson, 1999). More work is thus needed to determine how and via what mechanisms sibling and peer relationships impact moral development, and indeed, to determine how all of the factors discussed here and other factors interact with each other to influence an individual's moral development.

6.5.2 Group-level differences

Let us now consider differences in morality *between* groups. For sake of simplicity, I will limit my discussion to just two domains of morality relevant to this dissertation: prosocial behavior and third-party punishment.

It has been proposed that the human capacity for prosocial behavior is universal (Hoffman, 2000; Tomasello, 2009). The findings that even infants help others in the absence of rewards or praise, that children's instrumental helping actually *decreases* with extrinsic rewards, and that chimpanzees also instrumentally help others, all suggest that humans are naturally (rather than culturally) endowed with the capacity for prosocial behavior (see Warneken & Tomasello, 2009b). However, a universal capacity for altruism need not imply universality of prosocial behavior across cultural contexts because prosocial behavior also includes evaluations based on cultural values and moral beliefs (Trommsdorff, Friedlmeier, & Mayer, 2007). One may predict, then, that as children are socialized into the values, beliefs, and norms of their specific culture, their prosocial behavior will change to reflect this (see also Hay & Cook, 2007; Tomasello, 2009; Warneken & Tomasello, 2009b).

To my knowledge, no work thus far has systematically traced prosocial behavior from early to middle childhood across cultures, and so evidence for or against these developmental predictions is still needed. Researchers have, however, documented cultural differences in children's prosocial behavior (see Eisenberg, 1989, 1992). For instance, a recent study of 5-year-old children's spontaneous prosocial behavior (helping, comforting, and advising) revealed that German and Israeli children displayed more prosocial behavior towards a distressed adult as compared to Indonesian and Malaysian children (Trommsdorff et al., 2007). The researchers propose that in cultures that promote face-saving values and respect for hierarchical relations (such as Indonesia and Malaysia), ignoring the mishap of another person (especially an authority figure) can be more valued than attempting to help and

thereby risking that the other person lose face. The authors suggest that varying the experimental setting such that children are directly asked for help may evoke social responsibility and a feeling of obligation and thereby reduce the cultural constraints on prosocial behavior.

Interesting cross-cultural similarities and differences are also emerging in another domain of prosocial behavior: sharing. Rochat and colleagues (2009) recently compared distributive justice among 3- and 5-year-old children across several cultures, and found that in all cultures, children at both ages tended to maximize their own gain. However, these authors also found cross-cultural differences such that generally speaking, self-interest was significantly less pronounced in cultures that promote collective values, such as Peru, than in cultures that promote individualistic values, such as the United States (see Birch & Billman, 1986; Rao & Stewart, 1999, for similar results; see also Carson & Banuazizi, 2008, for evidence of differences in the distributive justice displayed by children in collectivist versus individualistic societies).

Cultural influences on sharing were also recently observed in a simple sharing task used by Liebal, Haun, and Tomasello (D. Haun, personal communication, February 2010) with 5- and 10-year-olds from Germany and from the ≠Akhoe Haillom group (a hunter-gatherer group in northern Namibia). Children in this study were allowed to take as many pieces of fruit as they wanted out of a basket but were asked to keep in mind that the next child to be tested would also like some fruit (the experimenter claimed not to know how many pieces of fruit were in the basket). The majority of German children responded fairly by taking half the fruits, and their egoistic responses (taking more than half) decreased with age. On the other hand, the majority of Haillom children responded egoistically, and their egoistic responses increased with age. The researchers argue that because the Haillom are a demand-share society in which sharing typically occurs in response to others' demands

(rather than spontaneously), Haillom children likely did not feel obliged to share spontaneously. Thus, already by age 5, children in the two cultures behaved according to their culture-specific sharing norms.

Another interesting case for cross-cultural comparisons is third-party punishment. As discussed previously, third-party punishment is thought to have been crucial for the emergence of human morality. Based on this assumption, one might predict cross-cultural uniformity regarding this phenomenon and indeed, third-party punishment (in response to unfair offers in economic games) has been documented in a diverse range of populations (Henrich et al., 2006). At the same time, there is substantial variation across populations, with some societies rather unwilling to punish stingy offers, some quite willing, and others willing to punish offers that are either too stingy or too generous (Henrich et al., 2006). There is also cross-cultural variation in the effectiveness of punishment for maintaining cooperation: The threat of costly punishment increases cooperation among university students in Boston (Rand et al., 2009) but not among university students in China (Wu et al., 2009). These cross-cultural variations are generally explained with reference, at least in part, to variations in societal norms. For instance, Henrich et al. (2006) argue that "local learning dynamics generate between-group variation as different groups arrive at different 'cultural' equilibria" (p. 1770). Although the ontogeny of third-party punishment across cultures has not yet been examined, based on the adult findings, one may predict that very young children should engage in third-party punishment in similar ways across cultures due to certain natural endowments, but as children acquire their culture-specific norms of punishment, their behavior should accordingly change by middle or late childhood.

A similar proposal comes from Richard Shweder, who questions Turiel's hypothesis that children universally differentiate moral versus conventional obligations. Recall from Chapter 1 that in support of their theory, Turiel and colleagues showed that children in

several cultures differentiate between moral and social conventional issues (e.g., Nucci et al., 1996). However, Shweder and colleagues found that 5- to 7-year-olds in India and the United States differed significantly in terms of which issues they considered to be moral, conventional, and personal (Shweder, Mahapatra, & Miller, 1987). For instance, children in India considered it wrong to address one's father by first name, whereas children in the United States considered this to be acceptable. Shweder et al. (1987) argue that a culture's worldview has a significant bearing on the ontogenesis of moral understanding in that culture, and not all cultures consider social practices to be social conventions that derive their authority from group consensus. Furthermore, harm, rights, and justice form just one (rather than the only) moral code, the one predominantly used in Western cultures; other moral codes, such as duty, hierarchy, and interdependency, or sacred order, tradition, and personal sanctity, predominate in other cultures (see also Haidt, 2007; Shweder, 1990). Shweder et al. thus propose that although the capacity to make judgments about moral versus conventional issues and to experience various moral sentiments (such as empathy or guilt) is likely universal, the content of those judgments and the objects of the moral sentiments are culturally acquired. Based on their findings, they conclude: "… if children do subscribe to a universal moral code spontaneously generated independently of participation in social practices and socialization experiences, then we must search for it within the first four years of life. By age 5, children around the world do not typically agree with each other about what is morally right or wrong" (Shweder et al., 1987, p. 60).

All in all, morality is clearly both similar and variable across cultures. The challenge is to determine which aspects are similar and why, and which aspects are variable and why. One plausible hypothesis is that certain tendencies or capacities are universally present (and perhaps evolutionarily shaped) but are modified significantly by socialization and culture such that they might eventually look quite distinct (see, e.g., Tomasello, 2009). This

modification, it is argued, is where chimpanzees and humans are distinct: Both chimpanzees and humans are naturally endowed with the capacity for altruism and both become more selective altruists based upon direct social experiences with others (e.g., both prefer reciprocators), but only humans seem to reshape their natural altruistic tendencies according to the norms of their social group (Warneken & Tomasello, 2009b). The extant empirical work suggests that certainly by around 5 years of age (if not earlier), many aspects of children's moral behavior and understanding (including those assessed in the present studies) are already culture-specific. However, much more work is needed to demonstrate the extent of this culture-specificity as well as the mechanisms underlying its ontogenetic emergence.

All in all, both individual- and group-level factors play critical roles in determining any individual's morality (and thus likely also go some way towards explaining why, despite the natural prosocial and moral proclivity that humans possess, antisocial and immoral behavior also exists in human societies; see, e.g., Baumeister, 1999). Considering such differences in the future is thus absolutely essential for a comprehensive understanding of human morality and moral development.

6.6 Affect, cognition, or both?

I turn now to a brief discussion about whether affect or cognition serves as the basis for human (in particular, young children's) morality. The affect versus cognition (or sentiment versus reason) debate has long occupied philosophers and psychologists alike (see Chapter 1). More recently, however, there has been a move towards more integrative approaches to morality wherein both emotional intuition and conscious reasoning play important roles and influence each other. My approach in this dissertation was similarly integrative, and I thus did not attempt to tease apart the influences of affect versus cognition. Moreover, I propose that the studies in this dissertation provide support for such integrative approaches. The clearest example of this comes from Study 1, in which young children were

affectively involved in (and concerned about) a victim's plight. However, since the victim provided no observable emotional cues, children likely had to rely on cognitive processes such as affective perspective-taking or relevant scripts in order to engage affectively with the victim. Moreover, they would also have needed to rely on the basic cognitive processes required for sympathizing and empathizing, such as self-other discrimination. This shows the integral link between affective and cognitive processes, and, in line with the hybrid approach of researchers such as Jean Decety and Tania Singer, suggests that cognitive (top-down) processes influence affective (bottom-up) processes (Decety & Jackson, 2006; Decety & Lamm, 2006; Singer & Lamm, 2009).

It is also worth considering the third-party intervention findings in this light. As mentioned previously, third-party punishment is considered critical to (human) morality because it demonstrates agent-neutral application of moral norms. This obviously requires knowledge of the relevant norms, which falls most naturally into the domain of cognition. However, third-party intervention is also thought to rely on complex (and perhaps human-specific) social emotions. In dyadic interactions, when an individual is harmed, she experiences anger, which motivates her to retaliate against the transgressor (Izard, 1977). In a third-party context, the emotional response to another's harmful behavior is moral outrage, which is thought to occur when the harmful behaviors against third parties are experienced as if they were transgressions against the self (Fessler & Haley, 2003). Thus, like anger, moral outrage (or moral indignation) leads an individual to inflict costs on transgressors, which reduces the attractiveness of future transgressions. Moral outrage is thus thought to contribute to third-party punishment and to be a powerful mechanism for maintaining morality in the group.

Although I did not attempt to assess whether children's intervention in third-party transgressions had cognitive or affective roots, I believe that the most likely scenario is a

hybrid one in which children were (cognitively) aware of the norms being broken, but they also had an affective response (something like moral outrage) to the transgressions, and the cognition and affect together motivated them to intervene and punish the transgressor[7]. This corresponds nicely to the Affect-Backed Normative Theory put forward by Shaun Nichols (2002, 2004). As mentioned in Chapter 1, Nichols proposes that moral judgments rely on a Normative Theory about which actions are prohibited as well as an affective mechanism that produces strong emotional responses to those actions. Together, these cognitive and affective processes give rise to moral judgments and behavior. I submit that it is plausible that already early in human ontogeny, morality results from both of these processes working closely together and influencing each other. More generally, I would argue that the classic debate about whether cognition or emotion dictates morality is most likely to be resolved by taking an integrative approach, although this would be a descriptive resolution, not a prescriptive one; whether cognition or emotion *should* dictate morality must be resolved in the philosophical domain (see Blasi, 1990).

6.7 Morality: Judgments, behavior, and emotions

An integrative approach is needed not only towards the bases of morality, but also towards the elements that comprise morality: judgments, behavior, and emotions. Theoreticians and researchers often focus on just one element to the exclusion of the others. My approach in this dissertation has been to include all three elements under the 'morality'

[7] A similar role of affect has also been proposed for positive, beneficial interactions. When an individual in a dyadic interaction receives a benefit, she feels gratitude, which motivates her to reciprocate and to defend the interests of her benefactor (Trivers, 1971). The equivalent of gratitude in a third-party interaction is moral approbation, which leads individuals to be positively inclined towards, and seek to reward, virtuous, prosocial actors (Fessler & Haley, 2003). This emotion can thus increase prosociality and cooperation by increasing the attractiveness of adhering to moral norms. Interestingly, in Study 2a, children did not reward an actor who acted prosocially in a third-party interaction. This may be because young children are not yet as cognizant of the norms of prosociality as they are of the norms of not causing harm, or because the emotion of moral approbation is not yet as powerful as the emotion of moral outrage, or both. Note, however, that in experiments with adults, punishment or the threat thereof seems to be more salient than reward as an incentive for cooperation or generosity (Andreoni et al., 2003; but see Rand et al., 2009). It is thus possible that the bias towards punishing rather than rewarding, and the bias against being punished rather than towards being rewarded, persists from childhood into adulthood (see discussion about negativity bias above).

umbrella and to explore one or more of these elements in each study depending on the question being addressed in that study. This broad approach led to some important findings that may not have emerged using a narrower focus.

The prime example of this comes from Study 2b, in which 3-year-olds were found to take into account a transgressor's intentions. This result might seem surprising given that decades of work on moral development, beginning with Piaget (1932/1997), had suggested that children below 5 years of age do not take transgressors' intentions into account but rather focus primarily on the consequences of the transgressors' actions (see Chapter 1 for a brief review of this literature). Importantly, almost all of this prior work had relied on children's verbal judgments of hypothetical situations, which may not be optimal for very young children whose language skills are limited, and which do not always correspond to children's actual behavior (Astington, 2004; Darley & Shultz, 1990; Wainryb et al., 2005). My use of prosocial behavior as a dependent measure revealed a sensitivity to transgressors' intentions at a significantly younger age than prior work had indicated. Moreover, my findings fit nicely with work from outside the moral domain that suggests that even children younger than 3 years of age clearly grasp others' intentions and mental states (see Harris, 2000; Tomasello et al., 2005).

More generally, it seems critical to consider all the various facets of morality, not only in order to assess how they diverge from one another (e.g., behavior and verbal judgments may not always coincide) but also how they provide converging evidence (e.g., the verbal judgments of the 5-year-olds in Study 4 corresponded to their cooperative behavior in the distribution of resources task). Moreover, the various facets of morality are theoretically separable but they do impact each other (e.g., moral affective responses such as sympathy seem to mediate subsequent moral behavior towards victims, as in Study 1) and they are intricately linked with each other (e.g., the moral reasoning abilities of 4- to 5-year-

old children are related to their empathic concern and prosocial behavior; P. A. Miller, Eisenberg, Fabes, & Shell, 1996). An integrative approach in which all facets are considered thus promises to be more informative than an approach that focuses on just one aspect of morality, and should thus be employed in future research.

CONCLUSION

I began this dissertation by posing some timeless questions about the nature of morality: Does morality emerge from reason or emotions? Is morality innate or socially determined? Are humans basically good, basically evil, or blank slates? Can other animals be moral? If answers to these questions remain open after millennia of discussions and debates, I might be forgiven for not providing any definitive answers in this dissertation. Still, by means of an ontogenetic approach in which multiple aspects of morality and a relatively wide age range were included, and by linking the present findings to phylogeny and to the ultimate and proximate motivators underlying morality, this dissertation has made some headway into getting at the roots and the nature of human morality, the mechanisms that drive it, and the influences upon it. The present findings contribute to the growing literature that suggests deep ontogenetic and phylogenetic roots of human morality. Yet, they also point to the many ways in which the nature of morality is flexible and sophisticated, and hint that many aspects of morality are still developing during childhood and are likely susceptible to socialization and cultural influences. The present findings thus paint a complex and multilayered picture, which seems only appropriate for a subject as vast and as fascinating as the origins of human morality.

REFERENCES

Alexander, R. D. (1987). *The biology of moral systems*. New York: Aldine de Gruyter.

Aloise, P. A. (1993). Trait confirmation and disconfirmation: The development of attribution biases. *Journal of Experimental Child Psychology, 55*, 177-193.

Andreoni, J., Harbaugh, W., & Vesterlund, L. (2003). The carrot or the stick: Rewards, punishments, and cooperation. *American Economic Review, 93*(3), 893-902.

Arsenio, W. F. (1988). Children's conceptions of the situational affective consequences of sociomoral events. *Child Development, 59*, 1611-1622.

Arsenio, W. F., & Lover, A. (1995). Children's conceptions of sociomoral affect: Happy victimizers, mixed emotions, and other expectancies. In M. Killen & D. Hart (Eds.), *Morality in everyday life: Developmental perspectives* (pp. 87-128). Cambridge, UK: Cambridge University Press.

Astington, J. W. (2004). Bridging the gap between theory of mind and moral reasoning. *New Directions for Child and Adolescent Development, 103*, 63-72.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*(3), 110-118.

Baird, J. A., & Astington, J. W. (2004). The role of mental state understanding in the development of moral cognition and moral action. *New Directions for Child and Adolescent Development, 103*, 37-49.

Barrett, K. C. (1995). A functionalism approach to shame and guilt. In J. P. Tangney & K. W. Fischer (Eds.), *Self-conscious emotions: The psychology of shame, guilt, embarrassment, and pride* (pp. 25-63). New York: Guilford Press.

Batson, C. D. (1987). Prosocial motivation: Is it ever truly altruistic? In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 20, pp. 65-122). San Diego: Academic.

Batson, C. D. (1991). *The altruism question: Toward a social-psychological answer*. Hillsdale, NJ: Erlbaum.

Batson, C. D. (1998). Altruism and prosocial behavior. In D. T. Gilbert, S. T. Fiske & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2, pp. 282-316). Boston: McGraw Hill.

Batson, C. D., Batson, J. G., Slingsby, J. K., Harrell, K. L., Peekna, H. M., & Todd, R. M. (1991). Empathic joy and the empathy-altruism hypothesis. *Journal of Personality and Social Psychology, 61*(3), 413-426.

Batson, C. D., Duncan, B. D., Ackerman, P., Buckley, T., & Birch, K. (1981). Is empathic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology, 40*(2), 290-302.

Batson, C. D., Dyck, J. L., Brandt, J. R., Batson, J. G., Powell, A. L., McMaster, M. R., et al. (1988). Five studies testing two new egoistic alternatives to the empathy-altruism hypothesis. *Journal of Personality and Social Psychology, 55*(1), 52-77.

Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Adults' emotional reactions to the distress of others. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp. 163-185). Cambridge, UK: Cambridge University Press.

Batson, C. D., Sager, K., Garst, E., Kang, M., Rubchinsky, K., & Dawson, K. (1997). Is empathy-induced helping due to self-other merging? *Journal of Personality and Social Psychology, 73*(3), 495-509.

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. (1999). *Evil: Inside human violence and cruely*. New York: W. H. Freeman.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*, 323-370.

Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: An interpersonal approach. *Psychological Bulletin, 115*(2), 243-267.

Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: Infants' understanding of intentional action. *Developmental Psychology, 41*(2), 328-337.

Berti, A. E., Garattoni, C., & Venturini, B. (2000). The understanding of sadness, guilt, and shame in 5-, 7-, and 9-year-old children. *Genetic, Social, and General Psychology Monographs, 126*(3), 293-318.

Birch, L. L., & Billman, J. (1986). Preschool children's food sharing with friends and acquaintances. *Child Development, 57*(2), 387-395.

Bischof-Köhler, D. (1991). The development of empathy in infants. In M. E. Lamb & H. Keller (Eds.), *Infant development: Perspectives from German speaking countries* (pp. 245-273). Hillsdale, NJ: Lawrence Erlbaum Associates.

Blair, R. J. R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition, 14*, 698-718.

Blasi, A. (1990). How should psychologists define morality? Or, the negative side effects of philosophy's influence on psychology. In T. E. Wren (Ed.), *The moral domain: Essays on the ongoing discussion between philosophy and the social sciences* (pp. 38-70). Cambridge, MA: The MIT Press.

Blount, S. (1995). When social outcomes aren't fair: The effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes, 63*(2), 131-144.

Boesch, C., & Boesch, H. (1989). Hunting behavior of wild chimpanzees in the Taï National Park. *American Journal of Physical Anthropology, 78*, 547-573.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.

Boyd, R., & Richerson, P. J. (1989). The evolution of indirect reciprocity. *Social Networks, 11*, 213-236.

Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology, 13*, 171-195.

Boyd, R., & Richerson, P. J. (2005). Solving the puzzle of human cooperation. In S. Levinson (Ed.), *Evolution and culture* (pp. 105-132). Cambridge, MA: MIT Press.

Bräuer, J., Call, J., & Tomasello, M. (2006). Are apes really inequity averse? *Proceedings of the Royal Society of London B, 273*, 3123-3128.

Brosnan, S. F., & de Waal, F. B. M. (2003). Monkeys reject unequal pay. *Nature, 425*, 297-299.

Brosnan, S. F., Schiff, H., & de Waal, F. B. M. (2005). Tolerance for inequity may increase with social closeness in chimpanzees. *Proceedings of the Royal Society of London B, 272*(1560), 253-258.

Brownell, C. A., Svetlova, M., & Nichols, S. R. (2009). To share or not to share: When do toddlers respond to another's needs? *Infancy, 14*(1), 117-130.

Brownell, C. A., Zerwas, S., & Balaraman, G. (2002). Peers, cooperative play, and the development of empathy in children. *Behavioral and Brain Sciences, 25*, 28-29.

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition, 112*, 337-342.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review, 1*, 3-25.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology, 76*, 839-855.

Carson, A. S., & Banuazizi, A. (2008). "That's not fair": Similarities and differences in distributive justice reasoning between American and Filipino children. *Journal of Cross-Cultural Psychology, 39*(4), 493-514.

Castelfranchi, C., & Poggi, I. (1990). Blushing as discourse: Was Darwin wrong? In W. R. Crozier (Ed.), *Shyness and Embarrassment: Perspectives from Social Psychology* (pp. 230-254). Cambridge, UK: Cambridge University Press.

Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgment. Vol. 1.* Cambridge, UK: Cambridge University Press.

Cole, P. M., Barrett, K. C., & Zahn-Waxler, C. (1992). Emotion displays in two-year-olds during mishaps. *Child Development, 63*, 314-324.

Costanzo, P. R., Coie, J. D., Grumet, J. F., & Farnill, D. (1973). A reexamination of the effects of intent and consequence on children's moral judgments. *Child Development, 44*(1), 154-161.

Cunningham, M. R., Steinberg, J., & Grev, R. (1980). Wanting to and having to help: Separate motivations for positive mood and guilt-induced helping. *Journal of Personality and Social Psychology, 38*(2), 181-192.

Cushman, F. A. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition, 108*(2), 353-380.

Cushman, F. A., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgments: Testing three principles of harm. *Psychological Science, 17*(12), 1082-1089.

Damon, W., & Killen, M. (1982). Peer interaction and the process of change in children's moral reasoning. *Merrill-Palmer Quarterly, 28*(3), 347-367.

Darby, B. W., & Schlenker, B. R. (1982). Children's reactions to apologies. *Journal of Personality and Social Psychology, 43*(4), 742-753.

Darby, B. W., & Schlenker, B. R. (1989). Children's reactions to transgressions: Effects of the actor's apology, reputation and remorse. *British Journal of Social Psychology, 28*, 353-364.

Darley, J. M., & Shultz, T. R. (1990). Moral rules: Their content and acquisition. *Annual Review of Psychology, 41*, 525-556.

Darwin, C. (1874). *The descent of man and selection in relation to sex*. New York: Rand, McNally & Company.

de Hooge, I. E., Zeelenberg, M., & Breugelmans, S. M. (2007). Moral sentiments and cooperation: Differential influences of shame and guilt. *Cognition and Emotion, 21*(5), 1025-1042.

de Waal, F. B. M. (1986). The integration of dominance and social bonding in primates. *Quarterly Review of Biology, 61*, 459-479.

de Waal, F. B. M. (1988). The reconciled hierarchy. In M. R. A. Chance (Ed.), *Social fabrics of the mind* (pp. 105-136). Hillsdale, NJ: Erlbaum.

de Waal, F. B. M. (1989). Food sharing and reciprocal obligations in chimpanzees. *Journal of Human Evolution, 18*, 433-459.

de Waal, F. B. M. (1997). The chimpanzee's service economy: Food for grooming. *Evolution and Human Behavior, 18*, 375-386.

de Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology, 59*, 279-300.

de Waal, F. B. M. (2009). *The age of empathy: Nature's lessons for a kinder society*. New York: Harmony Books.

Decety, J., & Jackson, P. L. (2006). A social-neuroscience perspective on empathy. *Current Directions in Psychological Science, 15*(2), 54-58.

Decety, J., & Lamm, C. (2006). Human empathy through the lens of social neuroscience. *The Scientific World Journal, 6*, 1146-1163.

Decety, J., & Sommerville, J. A. (2003). Shared representations between self and other: A social cognitive neuroscience view. *Trends in Cognitive Sciences, 7*, 527-533.

Demetriou, H., & Hay, D. F. (2004). Toddlers' reactions to the distress of familiar peers: The importance of context. *Infancy, 6*(2), 299-318.

Denham, S. A. (1986). Social cognition, prosocial behavior, and emotion in preschoolers: Contextual validation. *Child Development, 57*, 194-201.

Dunn, J. (1988). *The beginnings of social understanding*. Cambridge, MA: Harvard University Press.

Dunn, J., Cutting, A. O., & Demetriou, H. (2000). Moral sensibility, understanding others, and children's friendship interactions in the preschool period. *British Journal of Developmental Psychology, 18*(2), 159-177.

Dunn, J., & Hughes, C. (1998). Young children's understanding of emotions within close relationships. *Cognition and Emotion, 12*, 171-190.

Dunn, J., & Munn, P. (1987). Development of justification in disputes with mother and sibling. *Developmental Psychology, 23*(6), 791-798.

Eisenberg, N. (1986). *Altruistic emotion, cognition, and behavior*. Hillsdale, NJ: Erlbaum.

Eisenberg, N. (1989). *The roots of prosocial behavior in children*. Cambridge, UK: Cambridge University Press.

Eisenberg, N. (1992). *The caring child*. Cambridge, MA: Cambridge University Press.

Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology, 51*, 665-697.

Eisenberg, N. (2005). The development of empathy-related responding. In G. Carlo & C. P. Edwards (Eds.), *Nebraska symposium on motivation, Vol. 51: Moral motivation through the life span* (pp. 73-117). Lincoln, Nebraska: University of Nebraska 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. (1998). Prosocial development. In N. Eisenberg (Ed.), *Handbook of child psychology, Vol. 3: Social, emotional, and personality development* (5th ed., pp. 701-778). New York: Wiley & Sons.

Eisenberg, N., Fabes, R. A., Miller, P. A., Fultz, J., Shell, R., Mathy, R. M., et al. (1989). Relation of sympathy and personal distress to prosocial behavior: A multimethod study. *Journal of Personality and Social Psychology, 57*, 55-66.

Eisenberg, N., Fabes, R. A., Murphy, B., Karbon, M., Maszk, P., Smith, M., et al. (1994). The relations of emotionality and regulation to dispositional and situational empathy-related responding. *Journal of Personality and Social Psychology, 66*, 776-797.

Eisenberg, N., Guthrie, I. K., Murphy, B. C., Shepard, S. A., Cumberland, A., & Carlo, G. (1999). Consistency and development of prosocial dispositions: A longitudinal study. *Child Development, 70*, 1360-1372.

Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin, 94*, 100-131.

Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin, 101*, 91-119.

Eisenberg, N., Shea, C. L., Carlo, G., & Knight, G. P. (1991). Empathy-related responding and cognition: A "chicken and the egg" dilemma. In W. Kurtines & J. Gewirtz (Eds.), *Handbook of moral behavior and development, Vol. 2: Research* (pp. 63-88). Hillsdale, NJ: Lawrence Erlbaum Associates.

Eisenberg, N., Spinrad, T. L., & Sadovsky, A. (2006). Empathy-related responding in children. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 517-549). Mahwah: Lawrence Erlbaum Associates.

Eisenberg, N., & Strayer, J. (1987). Critical issues in the study of empathy. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp. 3-13). Cambridge, UK: Cambridge University Press.

Eisenberg-Berg, N., & Lennon, R. (1980). Altruism and the assessment of empathy in the preschool years. *Child Development, 51*, 552-557.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature, 425*, 785-791.

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior, 25*, 63-87.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*(6868), 137-140.

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

Feshbach, N. D. (1978). Studies of empathic behavior in children. In B. A. Maher (Ed.), *Progress in experimental personality research* (Vol. 8, pp. 1-47). New York: Academic Press.

Feshbach, N. D., & Roe, K. (1968). Empathy in six- and seven-year-olds. *Child Development, 39*, 133-145.

Fessler, D. M. T., & Haley, K. J. (2003). The strategy of affect: Emotions in human cooperation. In P. Hammerstein (Ed.), *Genetic and Cultural Evolution of Cooperation* (pp. 7-36). Cambridge, MA: MIT Press.

Flack, J. C., de Waal, F. B. M., & Krakauer, D. C. (2005). Social structure, robustness, and policing cost in a cognitively sophisticated species. *American Naturalist, 165*, E126-E139.

Flack, J. C., Girvan, M., de Waal, F. B. M., & Krakauer, D. C. (2006). Policing stabilizes construction of social niches in primates. *Nature, 439*, 426-429.

Gergely, G., Bekkering, H., & Kiraly, I. (2002). Rational imitation in preverbal infants. *Nature, 415*, 755.

Gergely, G., Nadasdy, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*, 165-193.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior, 24*(3), 153-172.

Goffman, E. (1967). *Interaction Ritual: Essays on Face-to-Face Behavior*. Garden City, NY: Anchor.

Gomes, C. M., Boesch, C., & Mundry, R. (2009). Long-term reciprocation of grooming in wild West African chimpanzees. *The Proceedings of the Royal Society of London B, 276*, 699-706.

Greene, J. D., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences, 6*(12), 517-523.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*, 2105-2108.

Grueneich, R. (1982). Issues in the developmental study of how children use intention and consequence information to make moral evaluations. *Child Development, 53*, 29-43.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814-834.

Haidt, J. (2007). The new synthesis in moral psychology. *Science, 316*, 998-1002.

Hamilton, W. D. (1964). The genetical evolution of social behavior I and II. *Journal of Theoretical Biology, 7*, 1-52.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450*(22), 557-560.

Hamlin, J. K., Wynn, K., & Bloom, P. (in press). 3-month-olds show a negativity bias in their social evaluations. *Developmental Science*.

Harris, P. L. (1985). What children know about the situations that provoke emotions. In M. Lewis & C. Saarni (Eds.), *The socialization of emotion* (pp. 161-186). New York: Plenum.

Harris, P. L. (1995). From simulation to folk psychology: The case for development. In M. Davies & T. Stone (Eds.), *Folk psychology: The theory of mind debate* (pp. 207-231). Oxford: Blackwell.

Harris, P. L. (2000). *The work of the imagination*. Oxford, UK: Blackwell Publishers.

Harris, P. L., Núñez, M., & Brett, C. (2001). Let's swap: Early understanding of social exchange by British and Nepali children. *Memory & Cognition, 29*(5), 757-764.

Harris, P. L., Olthof, T., Terwogt, M. M., & Hardman, C. E. (1987). Children's knowledge of the situations that provoke emotion. *International Journal of Behavioral Development, 10*(3), 319-343.

Harwood, M. D., & Farrar, M. J. (2006). Conflicting emotions: The connection between affective perspective taking and theory of mind. *British Journal of Developmental Psychology, 24*, 401-418.

Hastings, P. D., Zahn-Waxler, C., & McShane, K. (2006). We are, by nature, moral creatures: Biological bases of concern for others. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 483-516). Mahwah: Lawrence Erlbaum Associates.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion*. Cambridge, UK: Cambridge University Press.

Hauser, M. (2006). *Moral minds: How nature designed our universal sense of right and wrong*. New York: Harper Collins.

Hay, D. F. (1979). Cooperative interactions and sharing between very young children and their parents. *Developmental Psychology, 15*(6), 647-653.

Hay, D. F. (1994). Prosocial Development. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 35*(1), 29-71.

Hay, D. F., Castle, J., Davies, L., Demetriou, H., & Stimson, C. (1999). Prosocial action in very early childhood. *Journal of Child Psychology & Psychiatry, 40*, 905-916.

Hay, D. F., & Cook, K. V. (2007). The transformation of prosocial behavior from infancy to childhood. In C. A. Brownell & C. B. Kopp (Eds.), *Socioemotional Development in the Toddler Years* (pp. 100-131). New York: The Guilford Press.

Helwig, C. C. (2006). Rights, civil liberties, and democracy across cultures. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Development* (pp. 185-210). Mahwah, NJ: Lawrence Erlbaum Associates.

Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior, 19*, 215-242.

Henrich, J., & Boyd, R. (2001). Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology, 208*, 79-89.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., et al. (2006). Costly punishment across human societies. *Science, 312*(5781), 1767-1770.

Hill, K., & Hill, C. (1977). Children's concepts of good and bad behavior. *Psychological Reports, 41*, 955-958.

Hobbes, T. (1994). *Leviathan*. Indianapolis: Hackett Publishing Company. (Original work published 1651).

Hobson, J. A., Harris, R., García-Pérez, R., & Hobson, P. (2009). Anticipatory concern: A study in autism. *Developmental Science, 12*(2), 249-263.

Hoffman, M. L. (1975). Developmental synthesis of affect and cognition and its implications for altruistic motivation. *Developmental Psychology, 11*(5), 607-622.

Hoffman, M. L. (1976). Empathy, role taking, guilt, and development of altruistic motives. In T. Lickona (Ed.), *Moral development and behavior* (pp. 124-143). New York: Holt, Rinehart, & Winston.

Hoffman, M. L. (1981). Is altruism part of human nature? *Journal of Personality and Social Psychology, 40*(1), 121-137.

Hoffman, M. L. (1982). Development of prosocial motivation: Empathy and guilt. In N. Eisenberg (Ed.), *The development of prosocial behavior* (pp. 281-338). New York: Academic Press.

Hoffman, M. L. (1984). Interaction of affect and cognition in empathy. In C. E. Izard, J. Kagan & R. B. Zajonc (Eds.), *Emotion, cognition, and behavior* (pp. 103-131). Cambridge, UK: Cambridge University Press.

Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. Cambridge, UK: Cambridge University Press.

Holmgren, R. A., Eisenberg, N., & Fabes, R. A. (1998). The relations of children's situational empathy-related emotions to dispositional prosocial behavior. *International Journal of Behavioral Development, 22*(1), 169-193.

Hornik, R., Risenhoover, N., & Gunnar, M. (1987). The effects of maternal positive, neutral, and negative affective communications on infant responses to new toys. *Child Development, 58*(4), 937-944.

Hume, D. (1965). *An enquiry concerning the principles of morals*. Notre Dame: University of Notre Dame Press. (Original work published 1776).

Iannotti, R. J. (1985). Naturalistic and structured assessments of prosocial behavior in preschool children: The influence of empathy and perspective taking. *Developmental Psychology, 21*(1), 46-55.

Imamoğlu, E. O. (1975). Children's awareness and usage of intention cues. *Child Development, 46*, 39-45.

Ingram, G. P. D., & Bering, J. M. (in press). Children's tattling: The reporting of everyday norm violations in preschool settings. *Child Development*.

Irwin, D. M., & Moore, S. G. (1971). The young child's understanding of justice. *Developmental Psychology, 5*(3), 406-410.

Izard, C. E. (1977). *Human emotions*. New York: Plenum Press.

Jensen, K., Call, J., & Tomasello, M. (2006). What's in it for me? Self-regard precludes altruism and spite in chimpanzees. *Proceedings of the Royal Society of London B, 273*, 1013-1021.

Jensen, K., Call, J., & Tomasello, M. (2007). Chimpanzees are vengeful but not spiteful. *Proceedings of the National Academy of Sciences, 104*(32), 13046-13050.

Jensen, L. C., & Hughston, K. (1973). The relationship between type of sanction, story content, and children's judgments which are independent of sanctions. *Journal of Genetic Psychology, 122*, 49-54.

Kant, I. (1959). *Foundations of the metaphysics of morals (Lewis White Beck, Trans.)*. Indianapolis: Bobbs-Merrill. (Original work published 1785).

Karniol, R. (1978). Children's use of intention cues in evaluating behavior. *Psychological Bulletin, 85*(1), 76-85.

Karniol, R. (1982). Settings, scripts, and self-schemata: A cognitive analysis of the development of prosocial behavior. In N. Eisenberg (Ed.), *The development of prosocial behavior* (pp. 251-278). New York: Academic Press.

Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality and Social Psychology, 68*(3), 441-454.

Keltner, D., & Anderson, C. (2000). Saving face for Darwin: The functions and uses of embarrassment. *Current Directions in Psychological Science, 9*(6), 187-192.

Keltner, D., & Buswell, B. N. (1996). Evidence for the distinctness of embarrassment, shame, and guilt: A study of recalled antecedents and facial expressions of emotion. *Cognition and Emotion, 10*(2), 155-171.

Keltner, D., & Buswell, B. N. (1997). Embarrassment: Its distinct form and appeasement functions. *Psychological Bulletin, 122*(3), 250-270.

Keltner, D., & Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition and Emotion, 13*(5), 505-521.

Keltner, D., Young, R. C., & Buswell, B. N. (1997). Appeasement in human emotion, social practice, and personality. *Aggressive Behavior, 23*, 359-374.

Ketelaar, T., & Au, W. T. (2003). The effects of feelings of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: An affect-as-information interpretation of the role of emotion in social interaction. *Cognition and Emotion, 17*(3), 429-453.

Killen, M. (1991). Social and moral development in early childhood. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

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.

Kochanska, G., DeVet, K., Goldman, M., Murray, K., & Putnam, S. P. (1994). Maternal reports of conscience development and temperament in young children. *Child Development, 65*, 852-868.

Kochanska, G., Gross, J. N., Lin, M.-H., & Nichols, K. E. (2002). Guilt in young children: Development, determinants, and relations with a broader system of standards. *Child Development, 73*(2), 461-482.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347-480). Chicago: Rand McNally.

Kohlberg, L. (1976). Moral stage and moralization: The cognitive-developmental approach. In T. Lickona (Ed.), *Moral development and behavior: Theory, research, and social issues* (pp. 84-107). New York: Holt, Rinehart, & Winston.

Kopp, C. B. (2001). Self regulation in childhood *Encyclopedia of the Social and Behavioral Sciences*. London: Elsevier.

Korsgaard, C. M. (1996). *The sources of normativity*. Cambridge, UK: Cambridge University Press.

Krebs, D. L. (2006). The nature of altruism: Ultimate determinants and proximate mechanisms. *Psychological Inquiry, 17*(1), 48-50.

Krebs, D. L. (2008). Morality: An evolutionary account. *Perspectives on Psychological Science, 3*(3), 149-172.

Krebs, D. L., & Russell, C. (1981). Role-taking and altruism: When you put yourself in the shoes of another, will they take you to their owner's aid? In J. P. Rushton & R. M. Sorrentino (Eds.), *Altruism and helping behavior* (pp. 137-165). Hillsdale, NJ: Erlbaum.

Kuhlmeier, V. A., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science, 14*(5), 402-408.

Leary, M. R., Landel, J. L., & Patton, K. M. (1996). The motivated expression of embarrassment following a self-presentational predicament. *Journal of Personality, 64*(3), 619-636.

Leslie, A. M., Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect. *Psychological Science, 17*, 421-427.

Leslie, A. M., Mallon, R., & Dicorcia, J. A. (2006). Transgressors, victims, and cry babies: Is basic moral judgment spared in autism? *Social Neuroscience, 1*(3-4), 270-283.

Levin, J., & Arluke, A. (1982). Embarrassment and helping behavior. *Psychological Reports, 51*, 999-1002.

Levitt, M. J., Weber, R. A., Clark, M. C., & McDonnell, P. (1985). Reciprocity of exchange in toddler sharing behavior. *Developmental Psychology, 21*, 122-123.

Lewis, M., Sullivan, M. W., Stanger, C., & Weiss, M. (1989). Self development and self-conscious emotions. *Child Development, 60*(1), 146-156.

Lickona, T. (1976). Research on Piaget's theory on moral development. In T. Lickona (Ed.), *Moral development and behavior: Theory, research, and social issues* (pp. 219-240). New York: Holt, Rinehart, and Winston.

Mascolo, M. F., & Fischer, K. W. (2007). The codevelopment of self and sociomoral emotions during the toddler years. In C. A. Brownell & C. B. Kopp (Eds.), *Socioemotional Development in the Toddler Years* (pp. 66-99). New York: The Guilford Press.

McCullough, M. E., Kimeldorf, M. B., & Cohen, A. D. (2008). An adaptation for altruism?: The social causes, social effects, and social evolution of gratitude. *Current Directions in Psychological Science, 17*(4), 281-285.

McGraw, K. M. (1987). Guilt following transgression: An attribution of responsibility approach. *Journal of Personality and Social Psychology, 53*(2), 247-256.

Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees recruit the best collaborators. *Science, 311*, 1297-1300.

Melis, A. P., Hare, B., & Tomasello, M. (2008). Do chimpanzees reciprocate received favors? *Animal Behaviour, 76*(3), 951-962.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31*, 838-850.

Mikhail, J. (2007). Universal moral grammar: Theory, evidence, and the future. *Trends in Cognitive Sciences, 11*(4), 143-152.

Miller, D. T., & McCann, C. D. (1979). Children's reactions to the perpetrators and victims of injustices. *Child Development, 50*, 861-868.

Miller, P. A., & Eisenberg, N. (1988). The relation of empathy to aggressive and externalizing/antisocial behavior. *Psychological Bulletin, 103*, 324-344.

Miller, P. A., Eisenberg, N., Fabes, R. A., & Shell, R. (1996). Relations of moral reasoning and vicarious emotion to young children's prosocial behavior toward peers and adults. *Developmental Psychology, 32*(2), 210-219.

Miller, R. S. (1987). Empathic embarrassment: Situational and personal determinants of reactions to the embarrassment of another. *Journal of Personality and Social Psychology, 53*, 1061-1069.

Moore, C. (2009). Fairness in children's resource allocation depends on the recipient. *Psychological Science, 20*(8), 944-948.

Mumme, D. L., Fernald, A., & Herrera, C. (1996). Infants' responses to facial and vocal emotional signals in a social referencing paradigm. *Child Development, 67*(6), 3219-3237.

Nagel, T. (1979). *Mortal questions*. Cambridge, UK: Cambridge University Press.

Nagel, T. (1986). *The view from nowhere*. Oxford: Oxford University Press.

Nelissen, R. M. A., & Zeelenberg, M. (2009). When guilt evokes self-punishment: Evidence for the existence of a Dobby effect. *Emotion, 9*(1), 118-122.

Nichols, S. (2002). Norms with feelings: Towards a psychological account of moral judgment. *Cognition, 84*, 221-236.

Nichols, S. (2004). *Sentimental rules: On the natural foundations of moral judgment*. New York: Oxford University Press.

Nobes, G., Panagiotaki, G., & Pawson, C. (2009). The influence of negligence, intention, and outcome on children's moral judgments. *Journal of Experimental Child Psychology, 104*, 382-397.

Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature, 393*, 573-577.

Nucci, L. P., Camino, C., & Sapiro, C. M. (1996). Social class effects on northeastern Brazilian children's conceptions of areas of personal choice and social regulation. *Child Development, 67*(3), 1223-1242.

Nucci, L. P., & Nucci, M. S. (1982a). Children's responses to moral and social conventional transgressions in free-play settings. *Child Development, 53*(5), 1337-1342.

Nucci, L. P., & Nucci, M. S. (1982b). Children's social interactions in the context of moral and conventional transgressions. *Child Development, 53*(2), 403-412.

Núñez, M., & Harris, P. L. (1998). Psychological and deontic concepts: Separate domains or intimate connection? *Mind & Language, 13*(2), 153-170.

Nussbaum, M. C. (2001). *Upheavals of thought: The intelligence of emotions*. New York: Cambridge University Press.

O'Malley, M. N., & Greenberg, J. (1983). Sex differences in restoring justice: The down payment effect. *Journal of Research in Personality, 17*, 174-185.

Ohbuchi, K., Kameda, M., & Agarie, N. (1989). Apology as aggression control: Its role in mediating appraisal of and response to harm. *Journal of Personality and Social Psychology, 56*(2), 219-227.

Olson, K. R., & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition, 108*, 222-231.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*, 255-258.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience, 9*, 148-158.

Piaget, J. (1997). *The moral judgment of the child*. New York: The Free Press. (Original work published 1932).

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*, 1-72.

Radke-Yarrow, M., & Zahn-Waxler, C. (1984). Roots, motives, and patterns in children's prosocial behavior. In E. Staub, D. Bar-Tal, J. Karylowski & J. Reykowski (Eds.), *Development and maintenance of prosocial behavior: International perspectives on positive behavior* (pp. 81-99). New York: Plenum Press.

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.

Rand, D. G., Dreber, A., Ellingsen, T., Fudenberg, D., & Nowak, M. A. (2009). Positive interactions promote public cooperation. *Science, 325*(5945), 1272-1275.

Rao, N., & Stewart, S. M. (1999). Cultural influences on sharer and recipient behavior: Sharing in Chinese and Indian preschool children. *Journal of Cross-Cultural Psychology, 30*, 219-241.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Regan, D. T., Williams, M., & Sparling, S. (1972). Voluntary expiation of guilt: A field experiment. *Journal of Personality and Social Psychology, 24*(1), 42-45.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33*, 12-21.

Rhine, R. J., Hill, S. J., & Wandruff, S. E. (1967). Evaluative responses of preschool children. *Child Development, 38*(4), 1035-1042.

Rochat, P., Dias, M. D. G., Liping, G., Broesch, T., Passos-Ferreira, C., Winning, A., et al. (2009). Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *Journal of Cross-Cultural Psychology, 40*(3), 416-442.

Ross, H. S. (1996). Negotiating principles of entitlement in sibling property disputes. *Developmental Psychology, 32*(1), 90-101.

Ruby, P., & Decety, J. (2004). How would *you* feel versus how do you think *she* would feel? A neuroimaging study of perspective-taking with social emotions. *Journal of Cognitive Neuroscience, 16*(6), 988-999.

Ruffman, T., Perner, J., Naito, M., Parkin, L., & Clements, W. A. (1998). Older (but not younger) siblings facilitate false belief understanding. *Developmental Psychology, 34*, 161-174.

Sagi, A., & Hoffman, M. L. (1976). Empathic distress in newborns. *Developmental Psychology, 12*, 175-176.

Schopenhauer, A. (1995). *On the basis of morality (E. F. J. Payne, Trans.)*. Providence: Berghahn Books. (Original work published 1841).

Semin, G. R., & Manstead, A. S. R. (1981). The beholder beheld: A study of social emotionality. *European Journal of Social Psychology, 11*, 253-265.

Shweder, R. A. (1990). In defense of moral realism: Reply to Gabennesch. *Child Development, 61*(6), 2060-2067.

Shweder, R. A., Mahapatra, M., & Miller, J. G. (1987). Culture and moral development. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children*. Chicago: The University of Chicago Press.

Silk, J. B., Brosnan, S. F., Vonk, J., Henrich, J., Povinelli, D. J., Richardson, A. S., et al. (2005). Chimpanzees are indifferent to the welfare of unrelated group members. *Nature, 437*(7063), 1357-1359.

Simner, M. L. (1971). Newborns' response to the cry of another infant. *Developmental Psychology, 5*, 136-150.

Singer, T., & Lamm, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences, 1156*, 81-96.

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., & Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature, 439*, 466-469.

Slomkowski, C. L., & Killen, M. (1992). Young children's conceptions of transgressions with friends and nonfriends. *International Journal of Behavioral Development, 15*(2), 247-258.

Smetana, J. G. (1984). Toddlers' social interactions regarding moral and conventional transgressions. *Child Development, 55*, 1767-1776.

Smetana, J. G. (1989). Toddlers' social interactions in the context of moral and conventional transgressions in the home. *Developmental Psychology, 25*(4), 499-508.

Smetana, J. G. (2006). Social-cognitive domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen & J. Smetana (Eds.), *Handbook of Moral Development* (pp. 119-153). Mahwah, NJ: Lawrence Erlbaum Associates.

Smetana, J. G., Schlagman, N., & Adams, P. W. (1993). Preschool children's judgments about hypothetical and actual transgressions. *Child Development, 64*(1), 202-214.

Smith, A. (2006). *The theory of moral sentiments*. Mineola, NY: Dover Publications, Inc. (Original work published 1759).

Smith, C. E., Chen, D., & Harris, P. L. (in press). When the happy victimizer says sorry: Children's understanding of apology and emotion. *British Journal of Developmental Psychology*.

Sober, E., & Wilson, D. S. (1998). *Unto others*. Cambridge, MA: Harvard University Press.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science, 18*(7), 587-592.

Stein, N. L., & Levine, L. J. (1989). The causal organisation of emotional knowledge: A developmental study. *Cognition & Emotion, 3*(4), 343-378.

Stocks, E. L., Lishner, D. A., & Decker, S. K. (2009). Altruism or psychological escape: Why does empathy promote prosocial behavior? *European Journal of Social Psychology, 39*, 649-665.

Struthers, C. W., Eaton, J., Santelli, A. G., Uchiyama, M., & Shirvani, N. (2008). The effects of attributions of intent and apology on forgiveness: When saying sorry may not help the story. *Journal of Experimental Social Psychology, 44*, 983-992.

Tangney, J. P. (1991). Moral affect: The good, the bad, and the ugly. *Journal of Personality and Social Psychology, 61*, 598-607.

Tangney, J. P. (1992). Situational determinants of shame and guilt in young adulthood. *Personality and Social Psychology Bulletin, 18*, 199-206.

Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology, 58*, 345-372.

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift fur Tierpsychologie, 20*, 410-433.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2009). *Why we cooperate*. Cambridge, MA: MIT Press.

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*, 675-691.

Tracy, J. L., & Robins, R. W. (2004). Show your pride: Evidence for a discrete emotion expression. *Psychological Science, 15*(3), 194-197.

Trivers, R. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology, 46*, 35-57.

Trommsdorff, G., Friedlmeier, W., & Mayer, B. (2007). Sympathy, distress, and prosocial behavior of preschool children in four cultures. *International Journal of Behavioral Development, 31*(3), 284-293.

Turiel, E. (1998). The development of morality. In W. D. S. Ed.) & N. E. V. Ed.) (Eds.), *Handbook of child psychology. Vol. 3. Social, emotional, and personality development* (pp. 863-932). New York: Wiley.

Turiel, E. (2006). Thought, emotions, and social interactional processes in moral development. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Development* (pp. 7-35). Mahwah, NJ: Lawrence Erlbaum Associates.

Turiel, E., Killen, M., & Helwig, C. (1987). Morality: Its structure, functions, and vagaries. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children* (pp. 155-243). Chicago: The University of Chicago Press.

Ueno, A., & Matsuzawa, T. (2004). Food transfer between chimpanzee mothers and their infants. *Primates, 45*, 231-239.

Ungerer, J. A., Dolby, R., Waters, B., Barnett, B., Kelk, N., & Lewin, V. (1990). The early development of empathy: Self-regulation and individual differences in the first year. *Motivation and Emotion, 14*, 93-106.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin, 134*(3), 383-403.

Vaish, A., & Striano, T. (2004). Is visual reference necessary? Contributions of facial versus vocal cues in 12-month-olds' social referencing behavior. *Developmental Science, 7*(3), 261-269.

van der Mark, I. L., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2002). Development of empathy in girls during the second year of life: Associations with parenting, attachment, and temperament. *Social Development, 11*(4), 451-468.

Wainryb, C., Brehl, B., & Matwin, S. (2005). Being hurt and hurting others: Children's narrative accounts and moral judgments of their own interpersonal activities. *Monographs of the Society for Research in Child Development, 70*.

Warneken, F., Hare, B., Melis, A. P., Hanus, D., & Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young children. *PLoS Biology, 5*(7), 1414-1420.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science, 311*, 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.

Warneken, F., & Tomasello, M. (2009a). The roots of human altruism. *British Journal of Psychology, 100*(3), 455-471.

Warneken, F., & Tomasello, M. (2009b). Varieties of altruism in children and chimpanzees. *Trends in Cognitive Sciences, 13*(9), 397-402.

Weiner, B., & Peter, N. (1973). A cognitive-developmental analysis of achievement and moral judgments. *Developmental Psychology, 9*(3), 290-309.

Wellman, H. M., Larkey, C., & Somerville, S. C. (1979). The early development of moral criteria. *Child Development, 50*, 879-873.

Wellman, H. M., Phillips, A. T., & Rodriguez, T. (2000). Young children's understanding of perception, desire, and emotion. *Child Development, 71*, 895-912.

Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition, 35*, 245-275.

Wispé, L. (1986). The distinction between sympathy and empathy: To call forth a concept, a word is needed. *Journal of Personality and Social Psychology, 50*, 314-321.

Wispé, L. (1987). History of the concept of empathy. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp. 17-37). New York: Cambridge University Press.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*, 1-34.

Woodward, A. L. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development, 22*, 145-160.

Wu, J.-J., Zhang, B.-Y., Zhou, Z.-X., He, Q.-Q., Zheng, X.-D., Cressman, R., et al. (2009). Costly punishment does not always increase cooperation. *Proceedings of the National Academy of Sciences, 106*(41), 17448-17451.

Wynn, K. (2008). Some innate foundations of social and moral cognition. In P. Carruthers, S. Laurence & S. Stich (Eds.), *The innate mind: Foundations and the future*. Oxford, UK: Oxford University Press.

Yamamoto, S., & Tanaka, M. (2009). Do chimpanzees (Pan troglodytes) spontaneously take turns in a reciprocal cooperation task? *Journal of Comparative Psychology, 123*(3), 242-249.

Young, G., & Lewis, M. (1979). Effects of familiarity and maternal attention on infant peer relations. *Merrill-Palmer Quarterly, 25*, 105-119.

Young, S. K., Fox, N. A., & Zahn-Waxler, C. (1999). The relations between temperament and empathy in 2-year-olds. *Developmental Psychology, 35*(5), 1189-1197.

Yuill, N., & Perner, J. (1988). Intentionality and knowledge in children's judgments of actor's responsibility and recipient's emotional reaction. *Developmental Psychology, 24*(3), 358-365.

Zahn-Waxler, C., & Kochanska, G. (1990). The origins of guilt. In R. A. Thompson (Ed.), *The Nebraska Symposium on Motivation 1988: Socioemotional development* (Vol. 36, pp. 183-258). Lincoln: University of Nebraska Press.

Zahn-Waxler, C., & Radke-Yarrow, M. (1982). The development of altruism: Alternative research strategies. In N. Eisenberg (Ed.), *The development of prosocial behavior* (pp. 109-137). New York: Academic Press.

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology, 28*(1), 126-136.

Zahn-Waxler, C., & Robinson, J. L. (1995). Empathy and guilt: Early origins of feelings of responsibility. In J. P. Tangney & K. W. Fischer (Eds.), *Self-conscious emotions: The psychology of shame, guilt, embarrassment, and pride* (pp. 143-174). New York: Guilford.

Zahn-Waxler, C., Robinson, J. L., & Emde, R. N. (1992). The development of empathy in twins. *Developmental Psychology, 28*, 126-136.

Zelazo, P. D., Helwig, C. C., & Lau, A. (1996). Intention, act, and outcome in behavioral prediction and moral judgment. *Child Development, 67*, 2478-2492.

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., et al. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development, 68*(3), i-151.

ACKNOWLEDGEMENTS

This dissertation would have been impossible to produce without the many forms of morality shown by many people. First and foremost, I have had the privilege of working with Mike Tomasello and Malinda Carpenter, who provided me with just enough direction to not get lost but just enough freedom to feel like I was finding my own way. I am eternally grateful for their guidance, open-mindedness, and patience. I was also lucky enough to have amazing teachers in my pre-dissertation life. In particular, Angel Lillard and Amanda Woodward both contributed significantly to my intellectual development and supported me in just the right ways at just the right times.

I benefited greatly from discussions with members of the Psychology department at the MPI-EVA, especially Keith Jensen, Hannes Rakoczy, Amanda Seed, and Felix Warneken, and from discussions with Doris Bischof-Köhler, Rob Boyd, Özlem Ece Demir, Tobias Großmann, Kalina Michalska, and Joan Silk. I am also grateful to Tobias Großmann, Alicia Melis, Katja Liebal, and Marie Schäfer for helpful and prompt feedback on the dissertation manuscript.

I thank the remarkable people that make the Child Lab work like a well-oiled machine, and especially Katharina Haberl, Angela Loose, Eva Siegert, Elvira Portner, and Manuela Missana for help with planning studies, data collection, and coding. Many thanks to Petra Jahn, Manuel Reinartz, Waldemar Beser, and Vladislava Nadova for technical assistance, Roger Mundry for statistical assistance, and Annett Witzmann and Henriette Zeidler for organizing my life and for regularly going beyond the call of duty.

My deepest gratitude to the children who participated in the studies and to their parents and teachers, without all of whose friendly cooperation this dissertation would not exist.

I thank the Großmanns (and associates) for being my family away from home. I thank Tobias for his support and enthusiasm for my work, and for keeping me as balanced as he possibly could. And finally, I thank my family for waiting out my student days with patience and unrelenting love and support, and for being my very own moral compass.

SUMMARY

Humans are, for the most part, moral. We behave in moral ways: We help and share with others, comfort those in distress, and cooperate with others to achieve far more than any one of us could achieve alone. We also have a sense of morality, which consists of thoughts and feelings about rights and duties, good and bad character traits, and right and wrong motives and behaviors. Moreover, our moral apparatus contains feelings of entitlement and obligation, and moral emotions such as guilt and shame. Together, these thoughts and feelings lead us to monitor, judge, and react to our own motives and behaviors but also to others' motives and behaviors. The current thesis explored the ontogenetic emergence of some basic aspects of this morality. Specifically, in four studies, I examined young children's understanding of and their responses to moral transgressions and the victims and transgressors therein. All of the studies featured third-party moral transgressions, which are considered the litmus test of moral understanding because they tap into agent-neutral (rather than self-interested) applications of moral norms, and because intervention in third-party moral interactions is considered essential to human morality.

In Study 1, I explored 18- and 25-month-old children's sympathetic and prosocial responses to a victim. In most research on the early ontogeny of sympathy, young children are presented with an overtly distressed person and their sympathetic responses are observed. This work leaves unclear whether, when the victim's distress is not perceptible, young children can nevertheless sympathize with the victim. Children in Study 1 saw an adult either harming another adult by destroying or taking away her possessions (harm condition) or else doing something similar that did not harm her (neutral condition). The "victim" expressed no emotions in either condition. Nevertheless, in the harm as compared with the neutral condition, children showed more concern and subsequent prosocial behavior towards the victim. Children's concerned looks during the harmful event were also positively correlated

with their subsequent prosocial behavior towards the victim. Thus, by 18 months of age, children can sympathize with the victims of moral transgressions even in the absence of overt emotional cues from the victims, possibly by some form of affective perspective-taking, and they subsequently behave prosocially towards the victims of such transgressions.

Studies 2 and 3 assessed various forms of third-party intervention. The two parts of Study 2 assessed third-party intervention in the form of selective prosocial behavior towards actors in third-party interactions. In Study 2a, 3-year-old children watched one adult (the actor) harming or helping another adult. Children subsequently helped the harmful actor less often than a third (previously neutral) adult, but helped the helpful and neutral adults equally often. In Study 2b, 3-year-old children helped an actor who intended but failed to harm another adult less often than a neutral adult, but helped an accidentally harmful and a neutral adult equally often. Young children thus selectively avoid helping those who cause – or even intend to cause – others harm.

Study 3 examined whether children actively enforce agent-neutral moral norms by intervening in and attempting to prevent third-party moral transgressions. Three-year-old children and two puppets each created a picture or clay sculpture, after which one puppet left the room. In the harm condition, the remaining (actor) puppet then destroyed the absent (recipient) puppet's picture or sculpture. In a control condition, the actor acted similarly but did not harm the recipient. Children protested during the actor's actions, and, upon the recipient's return, tattled on the actor and behaved prosocially towards the recipient more in the harm than in the control condition. Together, Studies 2 and 3 show that young children actively intervene in third-party moral transgressions.

Study 4 examined the flexibility of children's moral understanding. I asked whether 4- and 5-year-old children, who have previously been shown to make quite sophisticated moral judgments about transgressors, have the flexibility to modify those judgments based

upon the transgressor's subsequent remorse. After children watched videos of transgressors either displaying or not displaying guilt, 5-year-olds appropriately inferred that the victim would be more upset with the transgressor who had not displayed guilt and would prefer the guilt-displaying transgressor. The 5-year-olds also said that they would prefer to interact with the guilt-displaying transgressor, judged the transgressor who had not displayed guilt to be meaner, and, in a distribution of resources task, gave more resources to the guilt-displaying transgressor. The 4-year-olds did not draw any of these inferences and distributed the resources equally to the two transgressors. Thus, between 4 and 5 years of age, children's moral judgments and behaviors become flexible enough to vary based upon whether or not the transgressor displayed remorse.

Together, the present studies contribute to the growing literature that suggests deep ontogenetic roots of human morality, and demonstrate the early ontogenetic emergence of a potentially human-unique aspect of morality, namely, agent-neutral morality. Yet, they also point to the many ways in which the nature of morality is flexible and sophisticated, and hint that many aspects of morality are still developing during childhood.

KURZFASSUNG

Menschen sind moralische Wesen. Unsere Handlungen sind durch Moralität geprägt: wir helfen und teilen mit anderen, stehen denen bei, die sich in Not befinden und kooperieren mit anderen um Dinge zu erreichen, die weit über das hinausgehen, was man als Einzelner schaffen kann. Dies gilt ebenso für unsere Gedanken und Intuitionen, die stark von moralischen Inhalten wie Gut und Böse, Recht und Unrecht bestimmt sind. Darüber hinaus verfügen wir über moralische Emotionen wie Schuld und Scham. Diese moralischen Gedanken und Gefühle dienen nicht nur dazu, unsere eigenen Motivationen und Verhaltensweisen zu überwachen und zu beurteilen, sondern auch die Anderer.

Die vorliegende Arbeit untersucht die ontogenetische Entstehung grundlegender Aspekte von Moralität. In vier Studien untersuchte ich, wie Kinder auf moralische Transgressionen und insbesondere auf die daran beteiligten Täter und Opfer reagieren. In allen Studien waren die Kinder nicht direkt an den Transgressionen beteiligt, sondern lediglich Beobachter. Diese Vorgehensweise gewährleistete, dass die Kinder nicht aus reinem Selbstinteresse handelten, und damit ihre Reaktion tatsächlich auf der moralischen Beurteilung der Situation beruhte. Die von einem unbeteiligten Beobachter gezeigten moralischen Verhaltensweisen werden als ein grundlegendes Merkmal einer sogenannten ‚agent-neutral' Moralität erachtet, welche wiederum ein entscheidendes Kriterium reifer menschlicher Moralität darstellt.

In Studie 1 untersuchte ich die Reaktionen (Mitgefühl und Hilfeverhalten) von 18 - und 25-Monate alten Kindern auf ein Opfer. In vorangegangenen Studien wurde vom Opfer immer eine klare emotionale Reaktion gezeigt und das Verhalten der Kinder auf die vom Opfer präsentierten emotionalen Signale beobachtet. Daher wissen wir nicht, ob Kinder auch ohne die Darbietung emotionaler Signale durch das Opfer Mitgefühl empfinden können und Hilfeverhalten zeigen. In Studie 1 sahen Kinder, wie ein Erwachsener einem anderen

Erwachsenen entweder Schaden zufügte (‚Harm'-Bedingung), oder eine ähnliche Handlung ausführte, die aber niemandem schadete (neutrale Bedingung). In keiner der beiden Bedingungen zeigte das Opfer Emotionen. Obwohl die Kinder keine emotionalen Signale zur Verfügung hatten, zeigten sie ein erhöhtes Mitgefühl und mehr Hilfeverhalten in der ‚Harm'-Bedingung als in der neutralen Bedingung. Außerdem zeigte sich eine Korrelation zwischen dem im Gesicht des Kindes ausgedrückten Mitgefühl und dem später gezeigtem Hilfeverhalten gegenüber dem Opfer. Diese Befunde deuten darauf hin, dass Kinder bereits im Alter von 18 Monaten in der Lage sind, auch ohne offensichtliche emotionale Signale mit einer Person Mitgefühl zu empfinden. Dies ist eine Fähigkeit, die sehr wahrscheinlich auf einer Form von affektiver Perspektivenübernahme basiert.

Studien 2 und 3 untersuchten verschiedene Formen der ‚third-party intervention' (Intervention aufgrund der Beobachtung der Interaktion Anderer). Die beiden Teile der Studie 2 untersuchten ‚third-party intervention' in der Form selektiven Hilfeverhaltens. In Studie 2a beobachteten 3-jährige Kinder, wie eine Person eine andere Person verletzte oder ihr half. Die Kinder halfen der Person weniger, die der anderen Person Schaden zugefügt hatte. Aber sie halfen der Person, die der anderen Person half, nicht mehr als einer neutralen Person. In Studie 2b halfen 3-jährige Kinder einer Person, welche die Absicht geäußert hatte, eine andere Person zu verletzen, weniger als einer neutralen Person. Sie halfen jedoch einer Person, die einer anderen versehentlich Schaden zugefügt hatte, nicht weniger häufig als einer neutralen Person. Diese Befunde zeigen, dass es Kinder im Alter von 3 Jahren gezielt vermeiden, jemandem zu helfen, der einer anderen Person Schaden zufügt oder beabsichtigt ihr Schaden zuzufügen.

In der dritten Studie untersuchte ich, ob Kinder aktiv eingreifen, wenn moralische Normen gebrochen werden und versuchen diesen Normbruch zu verhindern. Kinder im Alter von drei Jahren und zwei Puppen malten jeweils entweder ein Bild oder kneteten eine Figur.

Danach verließ eine der Puppen den Raum. In der ‚Harm'-Bedingung zerstörte die noch im Raum verbliebene Puppe das gemalte Bild oder die geknetete Figur der anderen Puppe. In der ‚Control'-Bedingung führte die im Raum verbliebene Puppe eine ähnliche Handlung aus, die aber niemandem schadete. In der ‚Harm'-Bedingung protestierten die Kinder mehr gegen die Handlungen der im Raum verbliebenen Puppe als in der ‚Control'-Bedingung. Außerdem „petzten" die Kinder und halfen der Puppe, deren Sachen zerstört worden waren, in der ‚Harm'-Bedingung mehr als in der ‚Control'-Bedingung. Insgesamt deuten die Befunde von Studien 2 und 3 darauf hin, dass Kinder aktiv eingreifen, wenn es zu einer moralischen Normverletzung zwischen anderen kommt.

Studie 4 untersuchte die Flexibilität des moralischen Verständnisses bei Kindern. Insbesondere beschäftigte ich mich in dieser Studie mit der Frage, ob die moralischen Urteile von Kindern im Alter von 4 und 5 Jahren durch die Schuldbekennung und Reue einer Person nach einer moralischen Normverletzung beeinflusst werden. Kinder sahen Videos, in denen eine Person (Täter) das Eigentum einer anderen Person (Opfer) kaputtmachte. Wenn der Täter daraufhin seine Schuld bekannte und Reue zeigte, schlussfolgerten die 5-jährigen Kinder richtig, dass das Opfer dann weniger verärgert sein würde. Außerdem sagten sie, dass sie im direkten Vergleich lieber mit dem Täter der sich schuldig gezeigt hatte, als dem Täter der sich nicht schuldig gezeigt hatte spielen würden und beurteilten den nicht bekennenden Täter als boshafter. In einer weiteren Aufgabe wurden die Kinder gefragt, wie sie drei kleine Geschenke zwischen den beiden Tätern aufteilen würden. Hier zeigte sich wiederum, dass 5-jährige Kinder die Mehrzahl der Geschenke (2-3) an den Reue bekennenden Täter verteilten. Im Alter von 4 Jahren zeigten sich keine solchen Effekte. Diese Befunde zeigen, dass die moralischen Urteile und Verhaltensweisen in diesem Alter entscheidend davon abhängen, ob ein Transgressor (Täter) sich schuldig zeigt und Reue bekennt, was auf eine erhöhte

Flexibilität und Kontextabhängigkeit des sich entwickelnden moralischen Verständnisses hindeutet.

Die hier vorgelegten Studien zeigen, dass menschliche Moralität tiefe ontogenetische Wurzeln hat und dass die für den Menschen vermutlich einzigartigen Aspekte einer so genannten ‚agent-neutral' Moralität bereits sehr früh entstehen. Außerdem deuten diese Studien darauf hin, dass moralisches Verhalten von Natur aus flexibel und vielschichtig ist und sich bestimmte Aspekte der menschlichen Moralität erst im Laufe des Heranwachsens entwickeln.

Publications based upon this dissertation:

Vaish, A., Carpenter, M., & Tomasello, M. (in press). Young children selectively avoid helping people with harmful intentions. *Child Development*.

Vaish, A., Carpenter, M., & Tomasello, M. (2009). Sympathy through affective perspective-taking and its relation to prosocial behavior in toddlers. *Developmental Psychology, 45,* 534-543.

Other publications by the author:

Vaish, A., & Woodward, A. (2010). Infants use attention but not emotions to predict others' actions. *Infant Behavior and Development, 33,* 79-87.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134,* 383-403.

Grossmann, T., & Vaish, A. (2008). Reading faces: Developing a multi-level analysis of a social stimulus. In T. Striano & V. Reid (Eds.), *Social cognition: Development, neuroscience and autism* (pp. 167-180). Blackwell.

Lillard, A., Nishida, T., Massaro, D., Vaish, A., Ma, L., & McRoberts, G. (2007). Signs of pretense across age and scenario. *Infancy, 11,* 1-30.

Striano, T., & Vaish, A. (2006). Seven- to 9-month-old infants use facial expressions to interpret others' actions. *British Journal of Developmental Psychology, 24,* 753-760.

Striano, T., Vaish, A., & Benigno, J. (2006). The meaning of infants' looks: Information-seeking and comfort-seeking? *British Journal of Developmental Psychology, 24,* 615-630.

Striano, T., Henning, A., & Vaish, A. (2006). Selective looking by 12-month-olds to a temporally contingent partner. *Interaction Studies, 7,* 233-250.

Vaish, A., & Woodward, A. (2005). Baby steps on the path to understanding intentions. *Behavioral and Brain Sciences, 28,* 717-718. [Commentary]

Vaish, A., & Striano, T. (2004). Is visual reference necessary? Contributions of facial versus vocal cues in 12-month-olds' social referencing behavior. *Developmental Science*, *7,* 261-269.

CURRICULUM VITAE

For reasons of data protection,

the curriculum vitae is not included in the online version.

SELBSTÄNDIGKEITSERKLÄRUNG

I, Amrisha Vaish, herewith declare that I have conceived and written the current dissertation without any help and/or materials unless explicitly stated otherwise. All studies herein are my original ideas, conceived in collaboration with my supervisors Michael Tomasello (Studies 1-4) and Malinda Carpenter (Studies 1, 2, and 4). The data for Study 3 were collected by Manuela Missana, under my supervision, to be used in her Magisterarbeit. This dissertation has not been submitted elsewhere.

9th April, 2010