CHAPTER 2

What's Moral Reasoning Got to Do with It? The Prosocial Decisions of Individuals and Groups

Moral reasoning and prosocial action

As was discussed in Chapter 1, the concept of moral reasoning has received special attention in developmental psychology. Usually, study participants are asked to find a solution to a hypothetical moral dilemma and to justify this solution from a moral point of view. It is argued that the sophistication of justifications proceeds through several distinct developmental stages (Kohlberg, 1969). These stages are characterized by the ability to differentiate and coordinate perspectives of self and others. Developmental levels of moral reasoning are characterized by different types of perspective taking. For instance, Level 1, called perspective differentiation, allows children to understand that others have a perspective different from that of the self but this is not yet considered. Level 2, perspective coordination, allows an individual to think about a problem from another's perspective and to coordinate the two perspectives. In Level 3, a person can think about a problem from the generalized perspective of all the people involved in a situation and can see all of them from a third-person observer perspective. So, in higher stages of moral reasoning, it is assumed that people come to see situations not only from their own perspective but also from the perspective of all others involved in the conflict. It has been argued that development into higher stages is crucial for prosocial behavior (Blasi, 1980; Lapsley, 1996) and for a moral sensitivity in terms of the awareness of others' needs in moral decision-making situations (Keller, 1996).

On the other hand, the predominance of moral reasoning for moral decision making has been questioned by various researchers (Blum, 1980; Hoffman, 2000; Turiel, 1983). More recently, Haidt (2001, 2003) has doubted the importance of moral reasoning as a *cause of prosocial behavior*. He argues that it is emotionally induced intuition that causes prosocial behavior and that moral reasoning is mere post hoc verbal justification of this behavior. His argument for the minor role of

moral reasoning *within the individual* is still the subject of controversy (Pizarro & Bloom, 2003; Saltzstein & Kasachkoff, 2004). In this study, we investigated Haidt's (2001) argument that moral reasoning plays an important role *outside the individual*, such as in group discussions, where people with different preferences argue with each other and try to change the others' minds. In such discourses, a highly developed reasoning ability in the moral domain may be of central importance for making influential prosocial arguments.

In this study, groups of three children who were either 9 (in third grade), 11 (in sixth grade), or 13 (in eighth grade) years old played the dictator game and the ultimatum game with another group of children. These age groups have been shown to differ in their level of moral reasoning (Gibbs & Widamann, 1982; Keller, 1996; Keller & Edelstein, 1991; Keller & Reuss, 1984), but not in their ability to show empathic distress or in their developmental level of justice reasoning (Damon, 1977; Hoffman, 2000).

Group decision making and moral reasoning

The group decision-making literature in social psychology has repeatedly shown that in situations where no logically correct answer exists or no one can demonstrate the correctness of an opinion, a majority model usually dominates group decisions (Kameda, Tindale, & Davis, 2003; Stasser, Kerr, & Davis, 1989). Imagine, as a simple example, that five members of a group discuss two alternatives. The majority model predicts that the group will select, with a probability of 1, an alternative that is supported by more than three members. Yet group decisions sometimes deviate from this majority model. In mock jury studies, it is a robust finding that even when the majority judges a defendant guilty, the jury will bring in a verdict of *not guilty* with some probability. The opposite is not true. In other words, the judgment of not guilty is much more influential than the judgment of guilty. Tindale, Smith, Thomas, Filkins, and Sheffey (1996) explained this so-called *leniency bias* by the standards of proof used in criminal trails. Common law has long stressed the importance of avoiding the error of falsely convicting an innocent defendant even if this would increase the chances of falsely acquitting a guilty one. Thus, the burden of proof is on the side of those who argue for guilty and this asymmetry produces an asymmetrical influence of the two alternatives in group discussion (cf. Baron, Kerr, & Miller, 1992; Laughlin, 1980; Laughlin & Ellis, 1986).

Egoistic behavior is often rational in terms of profit maximization. This principle of maximization is universal and can be easily understood. On the other hand, prosocial behavior is not economically rational and seems to require a different logic for arguing it (e.g. taking others' perspectives into account; see Habermas, 1983; Keller & Reuss, 1984, Keller & Edelstein, 1991). How can we make a selfish other understand the importance of considering others' welfare? Haidt (2001), who seriously questions the role of moral reasoning as a cause of *individual* prosocial behavior, argued that highly developed moral reasoning is used for changing others' moral intuition. A higher level of perspective taking might be necessary to make one's arguments appealing to others.

Indirect evidence supporting this idea was obtained through the analyses of qualities of persuasive messages made by children. Clark and Delia and colleagues (e.g. Clark & Delia, 1976, 1977; Delia, Kline, & Burleson, 1979) showed that the quality of persuasive messages (i.e. asking a parent to buy something they wanted, asking to have an overnight birthday party, or urging a stranger to keep a lost puppy) increased with age and was further related to developmental level of moral reasoning. Interestingly, the puppy task, in which the children were purely altruistically motivated, elicited a higher quality of persuasive messages than either of the other two tasks.

Burleson and Fennelly (1981) investigated how levels of persuasive messages, which were designed by Clark and Delia (1976, 1977), influenced children's sharing behavior. They found that second-grade children who were exposed to the highest level of appeals donated significantly more candy to a group of unknown children. Though the exact processes await rigorous investigations, I hypothesize that level of moral reasoning and level of perspective taking will influence group decision making by increasing the power of prosocial arguments in a discussion.

Experiment and hypotheses

We (Takezawa, Gummerum, & Keller, in press) designed an experiment in which a group of three children played a dictator and an ultimatum game, both as proposer, with another group of three children. As proposers can unilaterally determine how much to offer to receivers, offers in the dictator game are driven by purely altruistic behavior that cannot be attributed to selfish motives. Thus, when a proposer group discusses dictator game offers, only the dispute between egoistic and altruistic perspectives will arise among group members. This is a situation where we expect that the level of moral reasoning will influence the result of group discussion.

As discussed in Chapter 1, offers in the ultimatum game are driven either by purely altruistic motives or by fear of rejection. Thus, when a proposer group discusses ultimatum game offers, they have to consider not only the issue of fairness but also how the responder group will play the game. As this presupposes some perspective-taking ability on the side of the proposers, making this game much more complex than the dictator game, we cannot predict how moral reasoning will influence group discussions. Thus, this game is exploratively investigated without setting specific hypotheses.

To demonstrate the effect of moral reasoning in group decision making, we decided to compare group decision-making processes between age groups that are thought to be different in levels of moral reasoning. We expected that prosocial arguments would be less influential in the groups of younger participants, whose members have not yet developed a higher level of moral reasoning. We formulated the following hypotheses based on previous studies:

Hypothesis 1: Altruistic propensity is correlated with individual offers in the dictator game.

Hypothesis 2: Developmental level of moral reasoning is not correlated with individual offers in the dictator game.

Hypothesis 3: Level of moral reasoning is higher in the older age groups, respectively.

Hypothesis 1 is derived from the research by Takezawa (2003) and Watabe and Hayashi (2001) showing that prosocial behavior in the dictator game is related to a kind of (spontaneous) altruistic preference. Hypothesis 2 is derived from previous studies that showed no relation between moral reasoning and prosocial behaviors (Haidt, 2001; Hart & Fegley, 1995). Hypothesis 3 is derived from numerous studies in developmental psychology, which have demonstrated differences in moral reasoning abilities for the age groups we study here (Gibbs & Widamann, 1982; Keller, 1996; Keller & Edelstein, 1991; Keller & Reuss, 1984).

We did not derive any particular hypothesis for the ultimatum game. As fair offers in the ultimatum game are motivated either by altruistic preference or by fear of rejection, a correlation between altruistic preference and the offers in the ultimatum game may be weakened depending on the proportion of egoistic people. We cannot make a particular prediction. The effect of moral reasoning is also uncertain and we examined it exploratively.

The last hypothesis is the following:

Hypothesis 4: Group offers in the dictator game with younger age groups deviate from the

majority model prediction in an egoistic direction.

This is an original hypothesis that will be tested if the above three hypotheses are supported. Briefly, if prosocial arguments are no more influential than egoistic arguments, group offers will deviate from the majority model prediction in an egoistic direction. We predict that such a deviation will occur only in the younger age groups, where moral reasoning is at a lower stage and therefore cannot make influential prosocial arguments against selfish arguments. For group offers in the ultimatum game, again, we did not derive a particular hypothesis (as for the study of the ultimatum game with adult groups, see Morgan & Tindale, 2002; Robert & Carnevale, 1997).

Method

Participants

Participants were 39 students from third grade (9 years old; 21 girls, 18 boys), 51 students from the sixth grade (11 years old; 27 girls and 24 boys) and 48 students from the eighth grade (13 or 14 years old; 36 girls and 12 boys). The sixth and eighth graders were recruited from a bilingual grammar school in Berlin, the third grade students from an elementary school in Berlin.[1] Participants came from middle- to upper-middle-class families. Only the students who brought signed consent forms were allowed to participate in the experiment.

Procedures

At the beginning of an experimental session, three same-sex students from the same class were randomly selected and taken by a female experimenter into a separate room. Upon arrival, instructions were read aloud by the experimenter. Besides the experimenter another female graduate student was in the room and took care of the video equipment.

Participants were instructed that they would play two consecutive games, each played with a different group of three anonymous same-sex and same-age students. They were told that neither group would know with whom it had played the games. Similarly, it was pointed out that their decisions would not be divulged to their classmates, parents, or teachers.

[1] I report on only the groups that played both games as proposer. We had the other groups play the ultimatum game as responder and the dictator game as proposer. When the payments were determined, responses from these groups were used.

Participants were instructed in the proposer role of the ultimatum or dictator game and given twenty 20-cent coins (4 euros) in sixth and eighth grade or twenty 10-cent coins (2 euros) in third grade. The order of games was counterbalanced. In one condition, participants played the ultimatum game first and the dictator game second (denoted U/D in the following sections), and in the other condition the order was reversed (i.e. D/U).

After their understanding of the game structure was confirmed by quiz questions, students were separated and were required to write down their individual preference for how much to offer to the other group. The group was then asked to discuss, with the aim of reaching an agreement, how much to offer to the other group. The groups had 10 minutes to come to a unanimous agreement. If groups had not come to a decision after 10 minutes, they were prompted to finish the discussion within 2 minutes by the experimenter. The group discussions were videotaped and the experimenter and the assistant left the room before the discussion started.

After the group discussion, the group decision was written on a sheet of paper by the experimenter. The second game was conducted with the same procedure. At the end of the second game, students were again separated and asked individually to answer post-session questions. At the end of the experimental session, students were debriefed and dismissed.

After a few weeks, experimenters visited the classrooms again and asked students to complete another post-session questionnaire measuring altruistic preference, developmental stage of moral reasoning, and some personality items. The payment for the two games (described below), sealed in an envelope for each student, was handed out at the end.

Game structures

In both games, students were instructed that all groups would make a decision independently and after all groups had completed the experiment, two groups would be randomly matched and the payment determined. In the dictator game, it was explained that the proposer group would determine how many of the 20 coins to offer as the proposer. In the ultimatum game, it was explained that groups in the role of proposer would determine how many of the 20 coins to offer to the other group and groups in the role of responder would determine the minimum acceptable offer. If the minimum acceptable offer exceeded an offer from a proposer group, neither group would receive any money. It was stressed that, for each game, a different group would be selected as their opponent and matched so that games were basically independent of each other. The payment was determined as the sum of the results of the two games. Each student in the same group received as

much money as the group earned in the games. For example, if a group earned 10 coins, each student received 10 coins.

Scoring

Altruistic preference was measured with two decomposed games in the second post-session questionnaire. In both games, participants had to select one of two alternatives, each allocating some money to oneself and another person. In the first game, alternative A gives self 15 (euros) and the other 5. Alternative B gives both 8. In the second game, alternative A was identical but alternative B gives both 12. After a choice was made in each game, desirability of the choice was questioned with a Likert scale (1: like it a little to 4: like it very much). Individual choice and desirability rating were combined and an 8-point scale was created for each game. Altruistic preference is the average of two 8-point scales and ranges from 1 (like selfish choice very much) to 8 (like altruistic choice very much).

Developmental level of moral reasoning was measured with open-ended questions in the second post-session questionnaire. Participants were first asked whether it is important to keep a promise to an unfamiliar person and to give a reason for their answer. Second, they were asked whether it is important to help an unfamiliar person and why. Answers to the open-ended questions were scored from stage 1 to stage 3 reasoning with transitional stages (1.5, 2.5) according to the manual by Keller (1991, 1996). Inter-rater reliability for two independent raters was 87%. In line with standard procedures in developmental research on moral reasoning, the highest of the two scores was chosen as indicating the overall moral reasoning score.

Results

Because offers were often not normally distributed, both parametric and non-parametric tests were used in the following analyses.

Individual-level analyses of offers

Figure 2.1 shows the distribution of offers in the dictator game (DG) and the ultimatum game (UG) divided by grade, gender, and task order. In both games, a Mann–Whitney U test showed no significant effect of gender (DG: U = 2,038.5, p = 0.30; UG: U = 1,754.0, p = 0.53) and no significant effect of task order in UG (U = 1,784.0, p = 0.34). The effect of task order was

significant in DG (U = 1,663.5, p = 0.00). Figure 2.1f indicates that the offers in both games tended to increase when DG preceded UG. A Kruskal–Wallis test revealed no significant effect of grade on the individual offers in either game (DG: H = 2,766, p = 0.25; UG: H = 2,485, p = 0.30).

Next, analysis of variance (ANOVA) was conducted on each DG and UG to examine the interactions between these three independent variables, grade, gender, and task order. ANOVA on DG showed a significant main effect of task order, F (1, 138) = 8.40, p = 0.00 and Gender × Task order, F (1, 138) = 4.02, p = 0.05. The main effect of task order mirrors the findings of the non-parametric test: Participants gave more when DG was played before UG than the other way around. This effect was especially pronounced in boys (U = 186.0, p = 0.00) but not in girls. ANOVA on UG showed significant two-way interaction effects of Grade × Task order, F (2, 126) = 3.93, p = 0.02 and Gender × Task order, F (1, 126) = 6.49, p = 0.01. When the data was split according to grade, the effect of task order was only significant in sixth grade (U = 138.0, p = 0.00), where offers were higher when DG was played before UG. Similar to DG, the effect of task order was only significant for the boys (U = 190.5, p = 0.03) who gave more when DG preceded UG.

As DG offers have been found to be lower than UG offers (Forsythe, Horowitz, & Savin, 1994; Watabe & Hayashi, 2001), we compared the offers between games through within-individual comparison. A Wilcoxon signed-rank test showed, in general, that the offers were lower in DG than in UG (Z = −6.39, p = 0.00)**.** It was found that only 11 of 87 (7.9%) students made higher offers in DG than in UG.

Individual-level analyses of correlates to offers

In this section, I examine correlates of individual offers. First, we calculated altruistic preference and developmental level of moral reasoning for each individual. We could measure altruistic propensity for 122 individuals (35 from third, 44 from sixth and 43 from eighth grade), and moral reasoning for 117 individuals (31 from third, 45 from sixth and 41 from eighth grade). This means about 15% of the data on moral reasoning was lost.

As predicted, a Kruskal–Wallis test on altruistic preference with grade as independent variable revealed no significant effect (χ^2 = 0.14, p = 0.93). Mann–Whitney U tests showed significant results for gender (U = 1290.0, p = 0.01) and task order (U = 1427.5, p = 0.05); girls were more altruistic than boys and participants in the D/U condition were slightly more altruistic than participants in the U/D condition. An ANOVA on altruistic preference with grade, sex, and task order as independent variables showed only the significant effect of sex, F (1, 122) = 6.56, p = 0.01. Again, the effect of grade was not significant, F (2, 122) = .26, p = 0.77. Altruistic preference

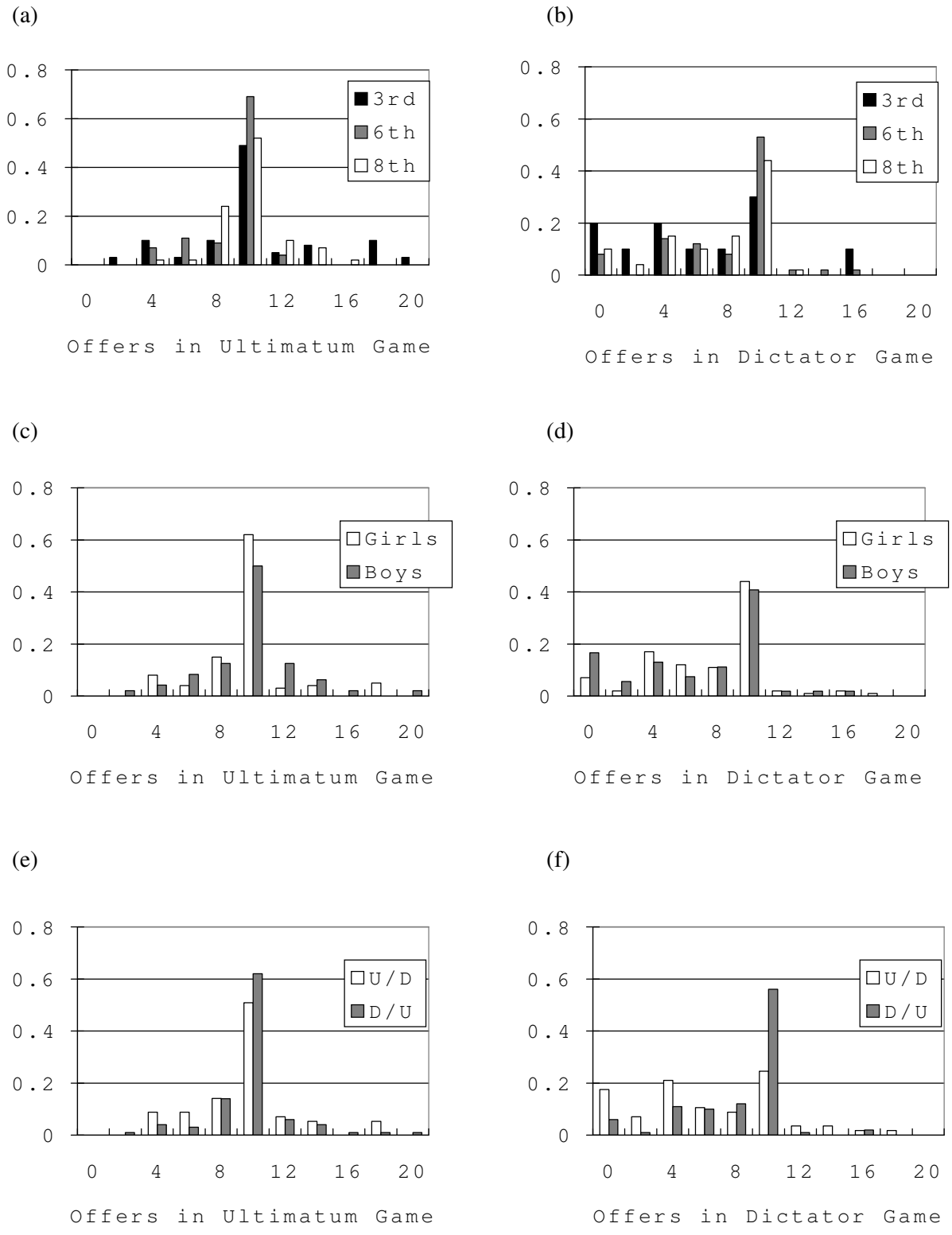

Figure 2. 1. Distribution of individual offers in (a, c, e) the ultimatum game and (b, d, f) the dictator game divided by grade (third, sixth, eighth), sex, and task order (U/D: ultimatum game first; D/U: dictator game first).

correlated significantly with individual offers in both games (DG: $r = 0.35$, $p < .01$; UG: $r = 0.23$, $p < .05$).

We then conducted a Kruskal–Wallis test on moral reasoning for each grade and Mann–Whitney *U* tests for possible effects of gender and task order. As we hypothesized, a significant difference was found for grade; developmental level of moral reasoning was significantly higher in older than in younger participants ($\chi^2 = 18.65$, $p = 0.00$, one tailed).

We found that 51% of third-grade children were classified in the lower stages (1 or 1.5) but only 31 and 32% of sixth- and eighth-grade students, respectively. Moreover, 23% of eighth-grade and 17% of sixth-grade students were in the higher stages (2.5 or 3) but only 3% of third-graders. The *U* tests revealed a significant effect of gender on moral reasoning ($U = 1161.5$, $p = 0.01$): 32% of girls but 46% of boys were classified in the lower stages (1 or 1.5), but 20% of girls and only 4% of boys were at higher moral reasoning stages (2.5 and 3). The *U* test for task order ($U = 1450.0$, $p = 0.24$) was not significant. Next, ANOVA was conducted on moral reasoning with the same three independent variables. As we hypothesized, the main effect for grade was significant, F (2, 117) = 6.30, $p = 0.00$. Post hoc Games–Howell tests showed that the moral reasoning of participants from third grade was significantly lower than the moral reasoning of both sixth and eighth graders. Between the two older age groups, no difference in moral reasoning emerged. No other main and interaction effects were significant. As predicted, level of moral reasoning did not correlate significantly with individual offers in either game (DG: $r = 0.03$; UG: $r = -0.03$).

We conducted four separate regression analyses on individual offers in DG or UG as a function of altruistic preference or moral reasoning, respectively. Altruistic preference predicted individual offers in both games (DG: $\beta = 0.33$, $t = 3.06$, $p = 0.00$, $R^2 = 0.11$; UG: $\beta = 0.26$, $t = 2.27$, $p = 0.03$, $R^2 = 0.07$); the higher the participants scored concerning their altruistic preferences, the more they contributed in both games. Moral reasoning ability could not predict either DG or UG individual offers (DG: $\beta = 0.01$, $t = 0.12$, $p = 0.90$, $R^2 = 0.00$; UG: $\beta = -0.07$, $t = -0.59$, $p = 0.56$, $R^2 = 0.01$).

Summary of individual-level analyses

Analyses of the post-session questionnaire revealed that altruistic preference was not significantly different between the three age groups but moral reasoning level was higher in the higher grades. We further found that moral reasoning was not related to the individual offers in DG and UG. The results are congruent with previous research that found no relationship between moral reasoning

and prosocial behaviors (Hart & Fegley, 1995). On the other hand, altruistic preference was related to offers in both games. This result also replicates previous findings (Takezawa, 2003).

DG offers in our experiment seem to be higher than in previous experiments (cf. Forsythe et al., 1994; for an experiment using DG and UG with children, see Harbaugh, Kraus, & Liday, 2003). As the current experimental setting was not comparable to the previous studies where participants played games individually under perfect anonymity, however, it is logically impossible to evaluate the current results just by comparing the average offers. Instead, within-individual comparison between the games successfully replicated a typical finding that DG offers were lower than UG offers. This assures us that participants in our experiment understood the strategic difference between the games and responded to it as theoretically expected. We could replicate all the major findings and all hypotheses were supported so far. Thus, an ideal situation was provided for examining the hypotheses on group decision-making processes.

Group-level analyses: Simple comparisons

Figure 2.2 shows the distribution of offers by the groups (see also Table 2.1). Mann–Whitney U tests revealed that gender (U = 142.5, p = 0.13) and task order (U = 196.0, p = 0.72) had no significant effects on the offers in UG. A Kruskal–Wallis test also showed no significant effect for grade (H = 1,335, p = 0.51). Figure 2.2b and d gives the impression that DG offers were smaller in third and sixth compared to eighth grade and in boys compared to girls.

These effects, however, did not reach significance in the Kruskal–Wallis and U test (grade: H = 4.274, p = 0.12; sex: U = 182.5, p = 0.18) probably because of small sample size. Task order had a significant effect with DG offers being lower when played after UG (U = 141.0, p = 0.01).
Figure 2.2 shows that group offers were remarkably different between conditions. The mere comparison of group offers, however, is not theoretically meaningful to us. As has been long discussed in group decision-making literature, group-level data is a product of (1) group decision-making processes that work as a kind of function, and (2) individual-level data that work as input to the function. For instance, if for two groups individual preferences are aggregated with the identical majority principle function, very small differences in individual preference among the two groups are likely to be amplified at a group level (for details, see Tindale et al., 1996). Thus, without analyzing group decision-making processes, we cannot identify how a difference at the group level is produced. This led us into the analyses detailed in the next sections.

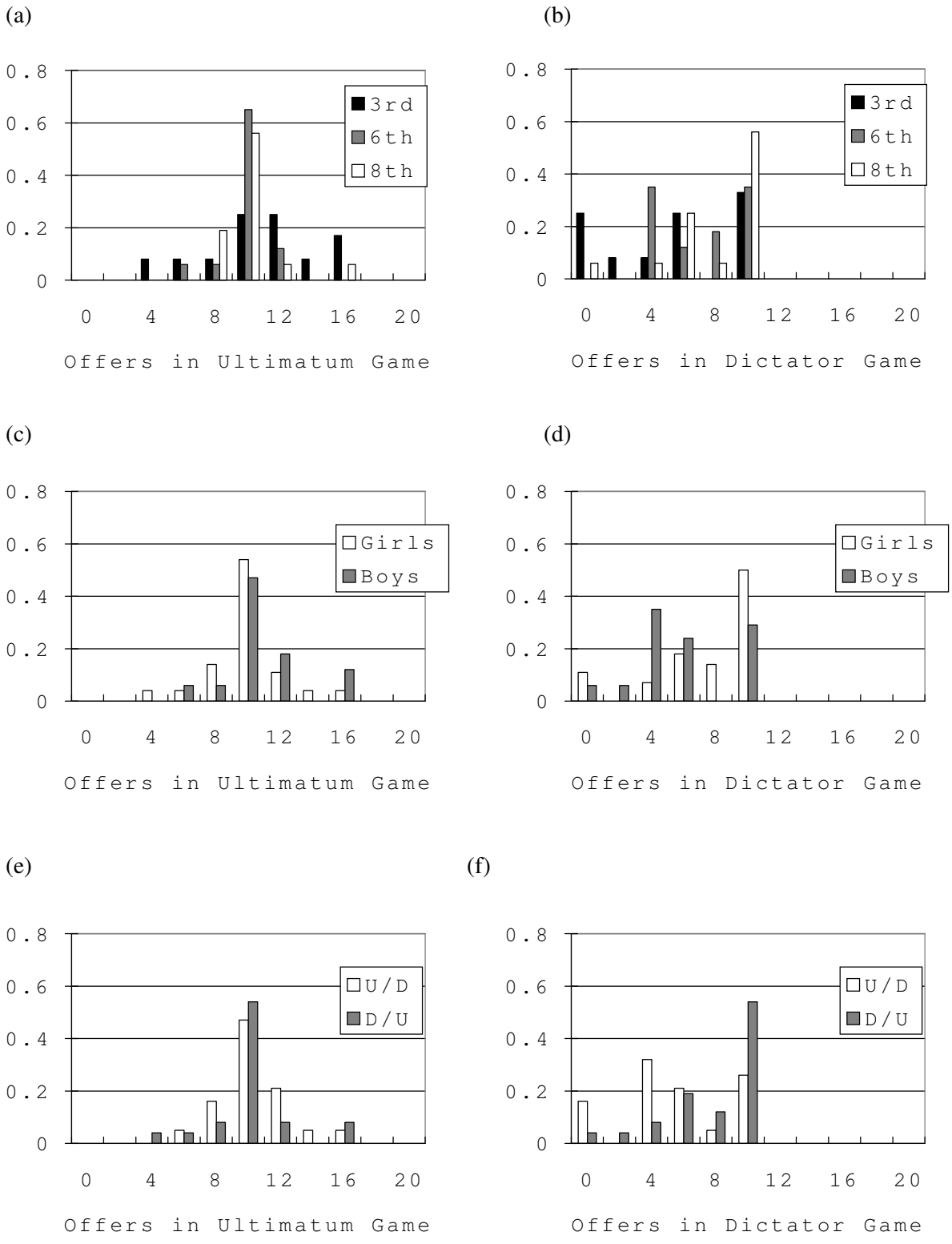

Figure 2.2. Distribution of group offers in (a, c, e) the ultimatum game and (b, d, f) the dictator game divided by grade (third, sixth, eighth), sex, and task order (U/D: ultimatum game first; D/U: dictator game first)

Table 2.1. Average offers by groups

		Model	Dictator game	Ultimatum game
Grade	Third	Actual	*N* = 12	*N* = 12
			5.58 (3.78)	11.25 (3.60)
		Average	7.18 (3.22)	10.67 (2.52)
		SJS	6.53 (4.13)	10.43 (3.13)
	Sixth	Actual	*N*=17	*N*=15
			7.47 (2.53)	10.00 (1.46)
		Average	8.29 (2.39)	9.31 (1.38)
		SJS	8.91 (1.88)	9.36 (1.45)
	Eighth	Actual	*N*=16	*N*=14
			8.25 (2.89)	10.36 (1.86)
		Average	7.40 (2.47)	10.19 (1.89)
		SJS	7.61 (3.16)	10.27 (1.60)
Sex	Girls	Actual	*N*=28	*N*=26
			7.75 (3.11)	10.08 (2.12)
		Average	8.04 (2.46)	10.04 (1.99)
		SJS	8.00 (3.21)	10.07 (2.04)
	Boys	Actual	*N*=15	*N*=17
			6.41 (3.10)	11.20 (2.73)
		Average	7.09 (2.90)	10.18 (2.13)
		SJS	7.57 (3.02)	9.83 (2.24)
Order	U/D	Actual	*N*=19	*N*=19
			5.89 (3.28)	10.58 (2.27)
		Average	6.68 (2.97)	9.84 (2.32)
		SJS	6.67 (3.70)	10.14 (2.59)
	D/U	Actual	*N*=26	*N*=22
			8.23 (2.69)	10.41 (2.54)
		Average	8.40 (2.16)	10.29 (1.75)
		SJS	8.74 (2.27)	9.86 (1.62)

Note: Numbers in parentheses are standard deviation. SJS: social judgment schema model. U: Ultimatum game. D: Dictator game.

Analyses of group decision-making processes: Theory and model

To reveal the processes of group decision making, we compared observed group offers with the predictions from two different models that represent different aggregate principles. One is the social judgment scheme model (SJS; Davis, 1996), which predicts that the group will offer a weighted average of individual offers biased toward majority.[2] For instance, if individuals A and B preferred to offer 10 coins and C preferred to offer only 4 coins, SJS predicts that the group will offer 10 coins. This is because SJS assumes that each individual is not equally influential in a group and that the majority enjoys a superior position. Another model that will be examined is averaging. This assumes that each individual is equally influential regardless of her status in a group. In the above example, the prediction is that the group will offer 8 coins [(10 + 10 + 4)/3].

As was discussed, majority principle is widely observed in both group judgment and decision-making situations and the SJS model has been shown to fit empirical data better than the averaging model (Davis, 1996; Ohtsuboet al., 2002). Thus, the SJS model serves as the primary criterion for evaluating group decision-making processes.

In general, in the SJS model no initial difference between prosocial and egoistic offers in terms of social influence is assumed. If the majority principle works, groups with a majority of egoistic individual offers will opt for an egoistic group offer, whereas groups with a majority of fair offers will opt for a fair group offer. However, if egoistic arguments are much more influential than

---

[2] In SJS, a group decision (X_G) is predicted by the following equation:

$$X_G = C_1 X_1 + C_2 X_2 + \ldots + C_n X_n,$$

Where $C_1 + C_2 + \ldots + C_n = 1$ and n is the number of group members. The weight C_i is determined by the following equation:

$$C_i = \frac{\sum_{j=1}^{n} f(|X_i - X_j|)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f(|X_i - X_j|)}, \text{ where } i \neq j.$$

The function $f(x)$ is further defined by the following equation:

$$f(|X_i - X_j|) = e^{-\theta(|X_i - X_j|)}, \text{ where } i \neq j.$$

Following the experimental results by Davis (1996) and Ohtsubo and colleagues (2002), $\theta = 1.0$ was used in the following analyses.

prosocial arguments, prosocial arguments may not be able to overcome egoistic arguments even when they occupy a majority position. In the above example, group offers will deviate from the SJS prediction in the selfish direction. On the other hand, if prosocial arguments are at least equally influential as egoistic arguments, actual group offers will not deviate in an egoistic direction from SJS predictions. If prosocial arguments are more influential, actual group offers will deviate from SJS in a fairer direction.

Analyses of group decision-making processes in the dictator game

We first compared the predicted offers of SJS and the averaging model. If fair offers occupied a majority position in most of the groups, SJS should be larger (fairer) than averaging-model predictions and vice versa. Results show that the overall SJS predictions (M = 7.84, SD = 3.11) were not statistically fairer than the predictions of the averaging model, M = 7.77, SD = 2.59; t (43) = −.29, p = 0.78, one tailed. However, this tendency was different across grades: In third grade the average predictions of the averaging model (M = 7.52, SD = 3.15) were larger than the predictions of the SJS model (M = 6.53, SD = 5.13) although this difference failed to reach statistical significance, t(10) = 1.51, p = 0.08. This result indicates that in third grade selfish offers occupy a majority position. In sixth and eighth grade the SJS predictions (sixth: M = 8.91, SD = 1.88; eighth: M = 7.62, SD = 3.16) were larger than the averaging model's predictions (sixth: M = 8.29, SD = 2.39; eighth: M = 7.39, SD = 2.47), although this difference was only significant in sixth grade, t(16) = −1.85, p = 0.04, one tailed, and not in eighth grade, t(15) = −0.68, p = 0.26. These results imply that groups in sixth and eighth grade were composed of a fair majority. A repeated ANOVA in which grade was entered into a model revealed a significant interaction effect of Grade × Model, F (2, 34) = 3.53, p = 0.04.

In the same vein, we checked the weights attached to each group member's individual offer in the SJS model. In each group the three individual offers were ranked from the most generous to the most selfish. Following from the SJS equation, the weight attached to an individual member's offer is an exponential function of the distances between a given member's preference and all other group members' preferences. The weight attached to any member decreases exponentially as an increasing function of the discrepancy of that individual's offer from the other members of the group. Following from this, the weight of the individual in the middle position (which should be most similar to any of the other individuals in the group) is always the largest. However, the question is whether the offer of the most egoistic or of the fairest child in a group receives a larger weight. Our analyses show that in third grade, the offer of the most egoistic child has a larger

weight (M = 0.33, SD = 0.17) than the offer of the most generous child (M = 0.19, SD = 0.19), whereas in sixth and eighth grade it was the other way around. The weight attached to the offer of the most generous individual (sixth: M = 0.36, SD = 0.14; eighth: M = 0.32, SD = 0.18) was on average larger than the weight of the most egoistic child (sixth: M = 0.20, SD = 0.16; eighth: M = 0.22, SD = 0.18). In third and eighth grade this difference was not statistically significant but in sixth grade it was, $t(16) = 2.21$, $p = 0.04$.

Table 2.1 shows the actual DG offers compared to the DG offers predicted by the two models for the three grades. In third and sixth grade, the actual offers were more selfish than both model predictions. As discussed above, this indicates that on average egoistic arguments were more influential in these two age groups than prosocial arguments. A paired t test showed that in third grade, the actual DG offers were statistically different from the averaging model, t (11) = −2.85, p = 0.02, but were not different from SJS, t (10) = −1.11, p = 0.30.[3] The opposite picture emerged in sixth grade. Actual DG offers were statistically different from the SJS model, t (16) = −2.59, p = 0.02, but were not different from the averaging model, t (16) = −1.67, p = 0.12. In eighth grade, actual DG offers were generally fairer than the model predictions but were not significantly different from SJS, t (15) = 1.31, p = 0.21. On the other hand, the averaging model predictions differed significantly from actual offers, t (15) = 2.40, p = 0.03. Taken together, these results imply that in third and eighth grade a majority seems to be more influential in the group decision-making process, whereas in sixth grade all three group members are equally influential in determining the group offer.

To examine if the difference between SJS and actual offers was modified by the other factors, an ANOVA was conducted on difference scores (i.e., actual offers minus SJS predictions) with grade, gender, and task order as independent variables. First, a main effect of grade was significant, F (2, 44) = 4.93, p= 0.01. The average difference score was −0.62 (SD = 1.87) in third grade, −1.44 (SD = 2.29) in sixth grade, and 0.63 (SD = 1.93) in eighth grade. Second, an interaction effect of grade and sex was found, F (2, 44) = 3.34, p = 0.05. Girls deviated more toward an egoistic direction in third grade (M = −1.00, SD = 2.03) but tended toward a fairer direction in sixth and eighth grade (sixth: M = 0.39, SD = 1.40; eighth: M = 0.28, SD = 1.80). Boys, on the other

[3] As the purpose of the t test in this section is to investigate model fit, keeping a 5% significance level increases the risk of erroneously accepting a model (see Kerr, Stasser & Davis, 1977). In such analyses, a significance level of 20% is usually used to reject a tested model. In the sixth grade, both models should be rejected if we use this criterion. As our concern is with the direction and the amount of deviation from SJS, the comparison with the averaging model should be treated as supplementary.

hand, tended toward a fairer direction in third and eighth grade (third: $M = 0.05$, $SD = 1.56$; eighth: $M = 1.66$, $SD = 2.22$) but to a much more selfish direction in sixth grade ($M = –2.62$, $SD = 2.59$).

Group decision-making processes in dictator game and moral reasoning

In line with Haidt (2001), we argued that moral reasoning ability is important for changing others' moral intuitions and making a prosocial argument influential during a group's discussion in DG. Our results so far show that groups in third grade tended to consist of a selfish majority. Consistent with group decision-making research, this selfish majority was more influential than the generous minority, and the group offer shifted in the selfish direction. Groups in sixth and eighth grade were composed of a generous majority. In eighth grade, this majority was more influential in the group decision-making process and prosocial arguments are more influential in the group decision-making process. In sixth grade, however, the group offers deviated from the SJS prediction in an egoistic direction. This means that the selfish minority group member was relatively more influential in the group decision-making process in sixth compared to eighth grade. I would like to argue that the relatively stronger influence of the generous majority in eighth grade is connected to generous children using more highly developed moral arguments to persuade the selfish group member. However, we did not find an overall difference between the moral reasoning abilities of sixth- and eighth-grade participants. Has our hypothesis failed?

For our next analyses, we checked for every single group in sixth and eighth grade whether the SJS or the averaging model described the aggregation of individual offers into one group offer best. For each group, we calculated the absolute difference scores between the actual group offer and the group offer predicted by either of the two models. The model with the smallest difference score was treated as the best-fitting model for this group. According to this analysis, the decision-making process of eight groups in sixth grade could be best modeled by the averaging model and of three group by the SJS model. For six groups the predictions were identical. In eighth grade, the decision-making process of nine groups could be best modeled by the SJS model, of five groups by the averaging model, and two groups were taken out of the analysis because of identical predictions (see Table 2.2).

Next, we checked whether the moral reasoning ability of the generous majority was on a higher level, the same level, or a lower level than the moral reasoning ability of the selfish child. In the three groups in sixth grade in which the SJS model was identified as the best-fitting model, in no case was the moral reasoning ability of the generous majority higher than the moral reasoning ability of the selfish child. In two cases, it was the same, and in one case, the selfish participants

Table 2.2. Classification of groups according to which model predicted the group decision-making process best and comparison of moral reasoning abilities of group members in sixth and eighth grade

Sixth grade		
SJS best N = 3	Averaging best N = 8	Same Prediction N = 6
Moral reasoning ability S = G; N = 2 S > G; N = 1	Moral reasoning ability S > G; N = 5 S < G; N = 2 Missing; N = 1	
Eighth grade		
SJS best N = 9	Averaging best N = 5	Same Prediction N = 2
Moral reasoning ability S < G; N = 7 S = G; N = 1 Missing; N = 1	Moral reasoning ability S < G; N = 2 S = G; N = 1 S > G; N = 1 Missing; N = 1	

Note: S = G: moral reasoning of selfish child identical to moral reasoning of generous child; *S > G:* moral reasoning of selfish child higher than moral reasoning of generous child; *S < G:* moral reasoning of selfish child lower than moral reasoning of generous child.

revealed a higher moral reasoning ability. Five out of eight averaging groups revealed a higher moral reasoning ability of the selfish minority child compared to the generous majority, in two cases it was reversed, and in one case no information was available because of missing data. In seven of nine SJS groups in eighth grade, the generous majority had a higher moral reasoning ability than the selfish minority child, in one group it was the same, and in one group no information was available because of missing data. In two of the five averaging groups the moral reasoning ability of the generous majority was higher than the moral reasoning of the selfish minority, in one group it was the same, and in one group, the selfish participant had a higher moral reasoning ability. For one group, no information was available (see Table 2.2).

Analyses of group decision-making processes in the ultimatum game

A paired *t* test revealed that the overall SJS predictions (M = 9.98, SD = 2.09) were not statistically fairer than the predictions of the averaging model, M = 10.09, SD = 2.02; t (41) = 0.53, p = 0.30, one tailed. This tendency held across all grades. A repeated measures ANOVA with the factors grade and model revealed no significant grade or interaction effects.

Table 2.1 shows that the actual offers in UG were fairer than the offers predicted by both models, and this trend holds across grades. Paired *t* tests between actual offers and either SJS or averaging model predictions revealed no significant difference in third grade, SJS: t (11) = 0.88, p = 0.40; averaging: t (11) = 0.76, p = 0.46, and eighth grade, SJS: t (13) = 0.36, p = 0.72; averaging: t (13) = 0.61, p = 0.55. In sixth grade, however, the predictions of both models were significantly different from the actual UG offers, SJS: t (14) = 2.45, p = 0.03; averaging: t (14) = 2.39, p = 0.03. Thus, we cannot tell which of the group decision-making models fit the data best in either grade, because both SJS and the averaging model either simultaneously fit or did not fit.

ANOVA on the difference score (actual–SJS) with sex, grade, and task order as independent variables showed no significant main or interaction effects. The results imply that, across all conditions, prosocial arguments were much more influential than selfish arguments. This is probably because in UG, it is easier to make an egoistic individual realize that egoistic offers may be rejected by the responder groups.

Discussion

In this study, we explored the influence of spontaneous altruistic preferences and moral reasoning on the prosocial decisions of children and adolescents in two economic games. In line with Haidt (2001, 2003) we questioned the importance of moral reasoning ability as a cause of individual prosocial (or moral) behavior, but we argued that moral reasoning ability plays an important role when people with prosocial preferences try to persuade more selfish group members during a group discussion. Our results show that individual offers in both dictator and ultimatum games were indeed correlated with altruistic preference measured in the post-session questionnaire but not with level of moral reasoning. Both individual offers and altruistic preference were not significantly different between third, sixth, and eighth grades, but level of moral reasoning was higher in the older age groups. Individual offers in the ultimatum game were much fairer than offers in the

dictator game, which indicates that the participants in our study understood the strategic difference between the two games.

Although our study supports Haidt's (2001, 2003) doubts about moral reasoning as a cause of prosocial behavior and replicates previous findings that prosocial behavior is not correlated with developmental level of moral reasoning (Hart & Fegley, 1995), we believe that the story may be much more complicated. Empathy, as a cause of prosocial behavior, is thought to be a complex of several psychological mechanisms (Casebeer & Churchland, 2003; Hoffman, 2000; Preston & de Waal, 2002). For instance, cognitive perspective taking, the ability to imagine oneself in another's situation, is thought to be an important component of empathy (Preston & de Waal, 2002). At the same time, reasoning from another's perspective is thought to be a component of higher stages of moral reasoning (Keller & Edelstein, 1991; Keller & Reuss, 1984; Kohlberg, 1969). Even though moral reasoning is not a direct cause of prosocial behavior, it may be of influence as a developmental process. A growing body of research in cognitive neuroscience further indicates that the relationship between reasoning and emotion is far more complicated than previously thought (for a review of moral reasoning and emotion in this field, see Casebeer & Churchland, 2003). Their mutual relationship in development therefore needs further investigation.

In the analyses of the individual offers, we further found that the order in which the games were played influenced the amount of offers. The offers in the dictator game became more selfish when it was played after the ultimatum game and vice versa. This finding might be attributed to a contrast effect; experience with the first game might have highlighted the strategic advantage of proposers' position in the dictator game. This reasoning is supported by the result that offers in the ultimatum game increased when it was played after the dictator game. Again, it could be that the contrast to the dictator game illuminated the risk of rejection of selfish offers in the dictator game. Since this effect of task order was more prevalent in boys than in girls, we might conclude that boys chose their offers more strategically than girls or were in fact more aware of the structural differences between the games.

In general, however, effects of gender on both individual and group offers in both games were relatively weak. This is in line with the economic model, which would predict that females and males behave similarly, since the principles of utility maximization apply for both genders equally. Therefore, experimental economic studies often do not include the gender variable into their analyses. Of the few empirical studies that do, no strong main effects of gender on behavior in dictator and ultimatum games are reported (see Camerer, 2003, for an overview). These findings parallel results from psychological research, which found only small differences in the prosocial behavior of adolescent and adult males and females (see Eisenberg and Fabes, 1998; Fabes and

Eisenberg, 1996). Gender differences were bigger among children than among adolescents and adults, with girls being more prosocial than boys. However, as Fabes and Eisenberg (1996) showed in a meta-analysis, gender differences are especially pronounced for self- and other-report measures but not for physiological and behavioral indices of prosocial behavior. This also mirrors the gender effects found in our study: We found only weak effects of gender for behavioral data (offers in dictator and ultimatum games), but larger main effects of gender for the self-report data in the post-questionnaire (moral reasoning and altruistic preference).

Instead of looking at the role of moral reasoning within individuals, our study intended to demonstrate that moral reasoning is likely to have an effect on individuals in group decision making. We showed such a possibility by analyzing how individual preferences are aggregated into group decisions via discussion. Given the economic irrationality of prosocial arguments (especially in the dictator game), it seems to be difficult to make them influential without either more highly developed moral reasoning or greater perspective-taking ability. This hypothesis is in line with the findings from Burleson and Fennelly (1981) and Clark and Delia (176, 1977). Thus, when children can use higher-stage moral reasoning in a context of persuasion, they should be more successful in influencing group choices.

Analyses of the group decision-making processes in the dictator game revealed a clear age difference although we did not find any age effect at the individual level. According to our hypothesis, actual group offers in the dictator game would deviate from the model predictions in an egoistic direction. However, in third grade, groups were composed of a majority of selfish children and consequently, the group offers shifted in the selfish direction. Thus, our hypothesis could not be tested in this case because of group compositions. In sixth and eighth grade, groups were composed of a generous majority and a selfish minority child. But whereas in eighth grade, SJS predicted group offers and no egoistic shift was observed, the egoistic minority child was relatively more influential in sixth grade, since the group offers were best predicted by the averaging model, in which all three group members are equally influential. These results indicate that the influence of prosocial arguments during the group discussion was different for the two age groups. However, our assumption that this difference is likely to be attributed to higher levels of moral reasoning in the older age group, which may have made prosocial arguments equally influential as egoistic arguments, is not yet fully supported by the data. The overall level of moral reasoning was not different in sixth grade compared to eighth grade. On the other hand, our more qualitative analyses revealed that in sixth grade the more selfish child (who was also more influential than the generous child) had a higher level of moral reasoning. In eighth grade, the more generous and more influential child showed a higher ability of moral reasoning than the more selfish child. Thus, only

in eighth grade did generous adolescents have the ability to persuade a more selfish other to agree to a prosocial group decision.

Concerning our argument that moral reasoning is the explanation for an age difference in group decision-making processes, some alternative explanations may be possible. We hypothesized that the development in moral reasoning is necessary for making moral arguments more persuasive and reasoned that this is the source of the age effect at the group-level analyses. It also seems to be the case that the overall ability to reason persuasively increases with age. It follows that what makes moral arguments persuasive is not the development in moral reasoning but the general cognitive reasoning ability that makes one's arguments more persuasive in wide contexts. The research by Clark and Delia and colleagues (Clark & Delia, 1976, 1977; Delia et al., 1979) does not provide decisive data on this issue. However, if it is the general ability of persuasive reasoning that made prosocial arguments more powerful in the eighth grade, it should also increase the influence of egoistic arguments. If both selfish and prosocial arguments became more persuasive in the eighth grade but the relative influence of prosocial to egoistic arguments was kept constant, we should obtain the same results at the group level in sixth and eighth grade. This is because the degree of deviations from the SJS model prediction (i.e., majority rule) is dependent on *relative* strength of social influence between prosocial and egoistic arguments. Our experiment showed, however, that the deviation from SJS in the egoistic direction disappeared in eighth grade, implying that the influence of prosocial relative to egoistic arguments increased with age. We do not deny the possibility that it is the general ability of persuasive reasoning that made prosocial arguments more influential in the eighth grade. Still, we need to explain why the development of general reasoning ability has an unequal influence on prosocial and egoistic arguments. More thorough data is necessary to deepen our understanding of how the processes that influence prosocial arguments increase with age.

In this study, the group decision-making process has been investigated in the tradition of a *social combination approach* (e.g. Davis, 1973, 1982). The basic idea of this line of research is to find a rule or function that translates the individual group members' preferences into one group decision. Thus, the challenge is to find a model that links group members' inputs to group output; such models are usually called social decision schemes. In the present study, two social decision schemes, a majority model and an averaging model, have been tested, and we have seen that different decision schemes are used in the three age groups studied. However, the exact processes by which group members with different allocation preferences influence and persuade others during group discussion still await rigorous investigation and I will follow up on this topic in the next chapter. I am especially interested in the kinds of strategies a generous majority in eighth grade

used that made them more effective in convincing a selfish minority than a generous majority in sixth grade. Investigating the actual group discussions will provide a useful extension to the analysis performed in this study. Moreover, social decision schemes are not formulated to predict a final decision of one single group, but rather the overall distribution of decisions across a sample of groups. As our qualitative analyses has shown, for some groups in the sixth grade, the majority model would have been a better description of the group decision-making process, although the majority of groups were best modeled by an averaging group decision-making process. The differences and similarities in these groups cannot be captured using only a social combination approach.

The analysis of the group discussions will also help to clarify the influence of moral reasoning "in action." It might not be enough that a generous majority rather passively possesses a higher moral reasoning ability, but that this skill has to be translated into convincing arguments that might have to be repeated and stressed in order to be understood and acted upon by the other group member(s). Haidt (2001) acknowledges the important influence of social influence (e.g. Sherif, 1935), conformity (e.g. Asch, 1956), and social comparison (e.g. Festinger, 1954) processes on moral intuitions and judgment in the reasoned-persuasion link and the social-persuasion link of his model. I think it necessary to identify the routes of social influence from a developmental perspective, to investigate how strongly exposure to prosocial or egoistic arguments influences a person's mind, and how long this influence might last.

Concerning the group decision-making process in the ultimatum game, our results indicate, interestingly, that prosocial arguments were much more influential than selfish arguments, and this effect was observed in every grade. We assume that this is because inherently egoistic individuals can be convinced to agree to fairer offers by making them realize the risk of rejection by responders. Again, this hypothesis can be followed up via analyzing the transcripts of the group discussions.

This study represents a first step in connecting research on prosocial behavior from economic and developmental psychological perspectives. I have tried to shed more light on the concept of "social preferences" and their developmental trajectory. From the provisional results, we can conclude that in situations of individual decision making, prosocial choices seem to be more influenced by a spontaneous preference for altruistic and fair allocations than by moral reasoning. This finding is more in line with the equity, reciprocity, and competition (ERC) and inequality-aversion models of Bolton and Ockenfels (2000) and Fehr and Schmidt (1999) than with Rabin's (1993) fairness equilibrium model. In a group situation, in which one's own social preferences have to be defended, moral reasoning and perspective taking might play more important roles. Our study therefore also demonstrates that the type of decision maker matters. Studying group as opposed to

individual behavior is also an important topic in economics, since the decision-making agents in many real-life situations are groups (e.g. committees, companies). Moreover, as we have seen in the study reported here, group behavior cannot be readily inferred from the individual behavior of the group members (see also Bornstein, Kugler, & Ziegelmeyer, 2002; Davis, 1992). Future research should therefore focus on the decision strategies of both individuals and groups.

62