Journal of Experimental Child Psychology 179 (2019) 362–374

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Children use rules to coordinate in a social dilemma

Sebastian Grueneisen [a,b,c,*], Michael Tomasello [a,d]

[a] Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, 04229 Leipzig, Germany
[b] Department of Psychology, University of Michigan, Ann Arbor, MI 48109, USA
[c] Center for Adaptive Rationality, Max Planck Institute for Human Development, 14195 Berlin, Germany
[d] Department of Psychology and Neuroscience, Duke University, Durham, NC 27708, USA

ARTICLE INFO

Article history:
Received 1 May 2018
Revised 23 August 2018
Available online 20 December 2018

Keywords:
Cooperation
Coordination
Rules
Children
Social dilemma
Chicken game

ABSTRACT

Humans are frequently required to coordinate their actions in social dilemmas (e.g. when one of two drivers has to yield for the other at an intersection). This is commonly achieved by individuals following communally known rules that prescribe how people should behave. From relatively early in development, children swiftly pick up the rules of their culture and even start creating game rules among peers. Thus far, however, little is known about children's abilities create rules to regulate their own interactions in social dilemma situations in which individuals' interests are partially in conflict. Here, we repeatedly selected dyads of children (5- and 8-year-olds, N = 144) at random from a group and presented them with a chicken game – a social dilemma in which individuals have conflicting motives but coordination is required to avoid mutual failure. In game breaks, groups reconvened and had the opportunity to think of additional game rules. Eight- but not five-year-olds readily came up with and agreed upon impartial rules to guide their subsequent game behavior (but only after adult prompting). Moreover, when playing by the self-made rules, children achieved higher payoffs, had fewer conflicts, and coordinated with greater efficiency than when playing without a rule - which mimics the functional consequences of rules on a societal level. These findings suggest that by at least age 8, children are capable of using rules to independently self-regulate potential conflicts of interest with peers.

* Corresponding author at: Department of Psychology, University of Michigan, Ann Arbor, MI 48109, USA.
E-mail address: sgruenei@umich.edu (S. Grueneisen).

https://doi.org/10.1016/j.jecp.2018.11.001
0022-0965/

Cooperation in social dilemmas is often challenging because individuals have self-serving goals that conflict with their cooperative goals. In one paradigmatic case, the so-called Prisoner's Dilemma (Rapoport, Chammah, & Orwant, 1965) the overall aggregate good would be maximized if everyone cooperated but cooperation is difficult to obtain because each individual has a strong incentive to be selfish (i.e., everyone is individually better off by not cooperating). Another situation of special importance in human cooperation is the Chicken game (also called the Snowdrift game, Rapoport & Chammah, 1966), captured by the story of two cars racing toward one another on a single-lane road, with neither wanting to yield to the other. In this situation both participants prefer the other to incur the cost of cooperating (e.g., by yielding the road). But if it looks like the partner will not do this, then doing it oneself is preferable to it not getting done since yielding the road is better than having a crash (i.e., everyone is worse off if no one cooperates).

When interacting repeatedly with the same partner, people often resolve such conflicts of interest by negotiating compromises, e.g. taking turns at yielding the road (Bornstein, Budescu, & Zamir, 1997; Helbing, Schönhof, Stark, & Holyst, 2005; Grueneisen & Tomasello, 2017). These person-to-person agreements become inefficient or impossible, however, when partners constantly change, as is frequently the case among people in modern societies who engage with strangers in economic exchanges, in public interactions (such as car traffic), or in large-scale companies with many employees. Let us call these societal-level interactions.

To meet the challenges of coordinating in societal-level interactions, humans create communally known rules (e.g. whichever car faces a yield sign should yield). Rules of this type are self-reinforcing in the sense that once they are commonly known it is in each person's interest to follow them (it is in my interest to yield if everyone else thinks it is my turn to yield, see McAdams, 2015; Hoffman, Yoeli, & Navarrete, 2016). They are collectively established to solve societal-level coordination problems by anticipating recurrent conflicts and offering a mutually beneficial (in the long run) solution, thereby reducing the "transaction costs" associated with achieving coordination by eliminating deliberation and negotiation in each case (Young, 2003; Lewis, 1969; Sugden, 1986; Hovenkamp, 1995; Shapiro, 2011).

A key developmental challenge for children is to pick up the various rules of their culture in order to participate in coordinated activities, and they start doing this from early in development (Piaget, 1932). Already at preschool age, for instance, children readily infer arbitrary rules by observing adult behavior (Kenward, Karlsson & Persson, 2011; Schmidt, Butler, Heinz, & Tomasello, 2016) or via minimal perceptual and linguistic cues (Roberts, Ho, & Gelman, 2017; Orvell, Kross, & Gelman, 2018). They also show an early understanding of the generic, even objective, nature of rules as is evidenced by their tendency to protest in response to others' rule violations using so-called generic normative language (e.g., "That's not how it's done!") (Rakoczy, Warneken, & Tomasello, 2008; Schmidt, Rakoczy, & Tomasello, 2011; Josephs, Kushnir, Gräfenhain, & Rakoczy, 2016).

But rather than viewing these rules as inalterable facts handed down by authorities, by late in the preschool period children appear to have an appreciation of the conventionality of rules, that is, that rules exist because (and only because) people collectively agree that they do (Schmidt, Rakoczy, Mietzsch, & Tomasello, 2016; Hardecker, Schmidt, & Tomasello, 2017; Zhao & Kushnir, 2018; Gilbert, 1989). What attests to this most clearly is that by five years of age children sometimes even establish their own arbitrary game rules in peer groups, readily transmit them to novices, and – just as with rules created by adults – treat them as objective standards in the sense that they demand conformity from all group members (Nobes, 1999; Göckeritz, Schmidt, & Tomasello, 2014; Hardecker et al., 2017). Late preschoolers also acknowledge that under certain circumstances rules can be changed (Riggs & Young, 2016; Hardecker et al., 2017; Zhao & Kushnir, 2018; Turiel, 1983; Smetana, 1981), which also indicates an early recognition of their social construction.

To date, however, children's creation, application, and enforcement of rules has almost exclusively been examined in the context of arbitrary games. While this approach has been extremely fruitful in revealing children's developing conceptual understanding of rules, little is currently known about their abilities to use rules in social dilemma situations in which individuals have conflicting motives but coordination is required for cooperative success – arguably the context in which the need for such rules is most pressing (Shapiro, 2011). This is likely to be more challenging since, in contrast to arbitrary game contexts in which children might even find it inherently enjoyable to create and play with rules, a feature of rules in social dilemmas is that they inevitably constrain individuals' abilities to unrestrictedly pursue their own self-interest. Furthermore, little is known about the functional consequences that rules have on children's cooperative efforts.

In one exceptional study, triads of five-year-olds were first primed to agree on a rule to allocate some rewards (to spin a wheel of fortune). In a subsequent test, children persisted longer at using this rule if it was fair and everyone had an equal chance of winning the greatest reward than when the experimenter unexpectedly introduced a wheel that systematically favored one child, rendering the rule unfair (Grocke, Rossano, & Tomasello, 2015). This finding points to an early appreciation of the legitimacy of rules. However, children in this study did not have to coordinate their decisions or to come up with the rules themselves, nor were they aware that they would face a conflict of interest later, all of which are important aspects of generating societal-level rules in social dilemma situations.

The current study therefore had two main objectives: first, to investigate whether children can come up with, agree on, and autonomously enforce rules in a context that mimics human societal interactions in some of its key respects; and second, to test if children's rules already share some of the functional consequences that rules are purported to have on a societal level, namely, to facilitate coordination success, mitigate conflict, and reduce transaction costs (e.g., Lewis, 1969; Sugden, 1986; Hovenkamp, 1995; Shapiro, 2011).

To this end, we presented 5- and 8-year-old children with a child-friendly social dilemma requiring coordination to avoid mutual failure (Chicken Game). On each trial, two children were randomly selected from a group of four children to play while the others watched with the inevitable conflict on display for all to see. In game breaks, the group of four was reconstituted and encouraged to find ways for improving their game efforts by establishing rules. This situation thus captures some central features of societal-level interactions in which groups create rules for individuals to apply in subsequent encounters.

We reasoned that successfully devising and applying rules in this context requires a principled application of equality principles in the sense that rules apply to all group members equally and have to be adhered to even if this results in personal disadvantage on a given trial. Since the tendency to act in accordance with fairness principles even at a personal cost increases substantially over the course of middle childhood (Blake & McAuliffe, 2011; Smith, Blake, & Harris, 2013; House et al., 2013), we predicted that 8-year-olds would establish rules more successfully than 5-year-olds. Moreover, based on the finding that children more consistently apply impartial over partial procedures (Grocke et al., 2015) we predicted that children would predominantly establish impartial rules and that playing by the rules would be associated with more equitable reward divisions. Finally, in line with conceptual work on the functional consequences of rules, we predicted that playing by the rules would be associated with greater coordination success (measured by group payoffs), a decrease of verbal conflicts, and a reduction of the transactions costs (as assessed by children's negotiation efforts and their reliance on explicit verbal agreements).

Material and methods

Participants

Seventy-two 5-year-olds (mean age = 5 years, 6 months; 50% girls) and 72 eight-year-olds (mean age = 8 years, 6 months; 47% girls) participated in the study. Children were tested in groups consisting of four boys (n = 12), four girls (n = 11), or two boys and two girls (n = 13). They were from the same school (although not always from the same class) and thus mostly familiar with each other. The sample size was specified prior to data collection (while this was a costly study to run, we settled for 18 groups per age group to at least approximate the recommendations made by Simmons, Nelson, and Simonsohn (2011) of at least 20 observations per cell). Three additional groups (two of 8-, and one of 5-year-olds, all boys) stopped participating throughout the game and were therefore excluded from the analysis. Children were mostly Caucasian, from middle class backgrounds, and were recruited and tested at urban schools and daycare centers. Informed consent was obtained from children's caregivers prior to the study and children had the opportunity to terminate participation at any time.

Apparatus and design

Children were presented with two battery-powered toy trains – one blue and one yellow – each carrying three marbles. The trains operated on wooden BRIO train tracks. In their starting positions they were about two meters apart, facing each other on a straight track. Each train had two stations – one at the end of the straight track behind the opposing train and the other at the end of tracks that swerved off the straight track (Fig. 1). Players could exchange all marbles their train successfully transported to a station for stickers. Each train could be steered to either of its two stations via a switch accessible to the player assigned to that train. The switches were covered up so that players could not see their partner's switch from their starting position. A barrier separating the players prevented children from actively interfering with each other's switches. However, by leaning forward children could see one another while discussing their game choices.

Once switched on, the trains simultaneously moved towards each other. If both trains went straight they crashed against each other and all marbles were lost. If one train swerved, the train going straight retained all of its marbles. The swerving train, however, lost two marbles which fell into opaque boxes from where they could not be retrieved anymore. If both trains swerved, they both lost two marbles (Table 1).

In each trial (i.e., each round of the game), two children were assigned to play the game and the others sat on chairs opposite the train setup. Children could freely communicate at all times. The experiment consisted of 20 trials in total but the exact number of trials was unknown to children. The order in which children played was randomized with the constraint that all children played 10 times and faced each other player at least 3 and not more than 4 times. A die – half blue and half yellow – was used to assign trains. Two small toys (another die and a 'wheel of fortune' – a rotatable horizontal wheel with an arrow symbol protruding laterally) were placed next to the apparatus as additional potential clues for rule creations. The experimenters never commented on these items.

Procedure

Introduction to the task

All children in a group were simultaneously introduced to the task by an experimenter (E1). E1 first showed how the switches worked and then asked each child to try it. Using a single train E1 then demonstrated how the marbles could be retrieved. A second experimenter (E2) started the train via remote control and the train went to its straight station. E1 collected the three marbles and exchanged them for three stickers from E2. This was repeated again with the train going to its swerving station and E1 collected only one sticker. Throughout the demonstration E1 recounted how many stickers he had won through which action. The process was then repeated with the second train so that children saw each train going once to each of its stations. The group was asked to restate the payoffs associated with going straight and with swerving and E1 provided corrective feedback in the rare case of a mistake (this was not a formal pre-test but rather an additional opportunity to state the game payoffs and to make sure children are tuned in to the task). E1 also explained that the trains would crash if both trains went straight and that in that case all marbles would be lost. While E1 prepared the first test trial with the help of a third experimenter (E3) each child received a drawing task and a sticker book into which they could put their stickers between trials.

Test trials

Before each trial, E2 and E3 left the test room and E1 announced which of the four children would play the game and which children would be observers. E1 then rolled the blue-yellow dice to determine which train the first player would steer; the other train was assigned to the second player. The default position of the switches was identical for both trains but was counterbalanced across trials. On the first three trials E1 reiterated the payoff structure using the following instruction: "Now you have the chance to win stickers. Remember, by going straight one gets three and by swerving, one. If both

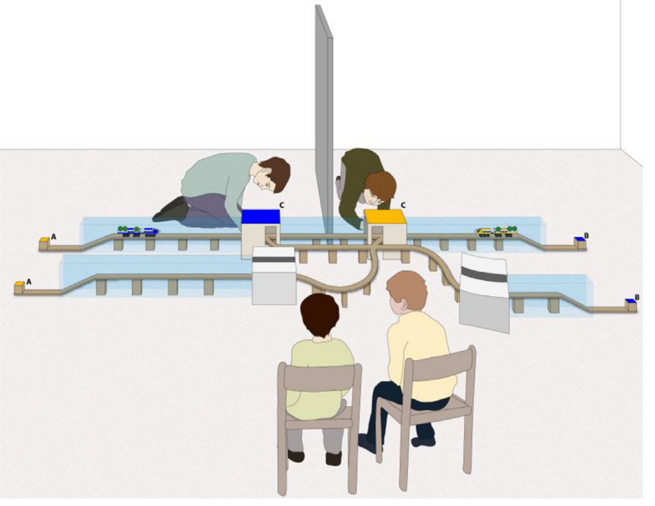

Fig. 1. Experimental setup. A = Yellow train stations; B = Blue train stations; C = Switches. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Payoff matrix of the game.[a]

		Player 2	
		Go straight	Swerve
Player 1	Go straight	0 | 0	3 | 1
	Swerve	1 | 3	1 | 1

[a] The left number in each cell corresponds to Player 1′s payoff, the right numbers to Player 2′s payoff.

trains go straight no one gets any. You guys sort it out yourselves. Shortly, I will come back and start the trains." On all subsequent trials E1 used the same instructions without restating the payoff structure. E1 then went behind a barrier, allegedly to do some work (this was done so that children felt they could interact freely and to reduce potential self-presentation motives). After 15 seconds E1 returned, started the trains via remote control, and went back behind the barrier. Once the trains started moving it took them another 15 seconds to reach the switch so that children had a total of roughly 30 seconds on each trial to negotiate their choices. When the trains had reached their stations children retrieved the marbles and exchanged them for stickers from E2 who re-entered between trials. While children put their stickers into their sticker books E1 and E3 prepared the following trial. After trials in which the trains had crashed E1 took all marbles off the trains and put them into a disposal box for children to see that these marbles would not be used again. If children complained or asked questions about how to play the game E1 always responded with "You sort it out yourselves."

Rule creation

E1 interrupted the game after 6, 10, and 14 trials to encourage children with increasingly explicit instructions to think of a game rule (if children used a rule already no further breaks were introduced). In the first break E1 said: 'You have been playing really well and collected many stickers. However, why don't you all think of a rule in order to play the game even better.' In the second break E1 said: 'You have been playing really well and collected many stickers. But sometimes you have to quarrel about who goes straight. I have an idea. Why don't you all think of a rule to decide who can go straight.'

In the third break E1 asked each child to show him his/her sticker book while the other children continued their drawing task. E1 attempted to seed a rule by asking the last child, determined randomly, to propose the rule that the blue train always goes straight. Children were then left alone again to discuss potential rules. During the discussion breaks children were left alone until a rule was agreed upon or until children stopped discussing ideas. The game then continued as before.

Coding

Each trial led to one of the three possible outcomes – coordination, both swerve, or crash – leading to group payoffs of 4, 2, and 0 stickers, respectively (i.e., the payoff of both players combined, these were coded as 2, 1, and 0 for the analysis). For each trial we counted payoffs of individual children and overall group payoff. All on-task communication was transcribed from videotape and coded by the first author who was thus not blind to the study's predictions. Each utterance was coded as containing normative language (commonly viewed as a way of expressing rules and of creating pressure to conform; Rakoczy et al., 2008), protest, normative protest, a negotiations attempt, or received no code (i.e., codes were mutually exclusive). In addition, we coded for each trial whether the two players had a dispute in that trial, whether or not players had formed a joint agreement about how to play before making a choice, and whether a rule was present (see Table 2 for definitions and examples). These communicative categories were used to assess how children communicated rules and enforced compliance (normative language and normative protest) and whether rules affected conflicts (verbal disputes and protest) and the efficiency of the coordination process (reliance on explicit agreements and use of negotiation attempts).

A rule was regarded as established if a child proposed a rule and no one protested or offered counter-proposals. A rule proposal was defined as suggesting an agent-neutral strategy aimed at coordination to be applied over multiple trials (i.e. 'I go straight' was not considered a rule proposal whereas 'the blue train always goes straight' or 'the tallest player goes straight' would have been). On any given trial the rule was coded as present if it was followed, mentioned explicitly or implicitly, or if any child protested against a proposed or carried out departure from the rule by another child. A second coder who was blind to the study's predictions and who was trained based on the category descriptions as well as hypothetical examples of each category recoded 20% of all trials. There was

Table 2
Coded communicative categories.

Communicative category	Definition	Example	κ coefficient
Normative language	Use of normative vocabulary, particularly deontic modal verbs (must, have to, ought to) normative adjectives (right, wrong), or making explicit reference to an agreement	"You must swerve." "I did it wrong!" "Remember, we have agreed on this."	0.87
Protest	Expressing objection to another's choice or proposed strategy	"Hey! That's mean!" "No! Don't do it like that!"	0.84
Normative protest	Expressing objection to another's choice or proposed strategy using normative language	"Hey! You must not do that!" "You broke the rule!"	0.92
Negotiation attempt	Providing an argument for why favoring a specific strategy	"Let me go straight because I have fewer stickers."	0.72
Agreements	Children explicitly reach a joint decision of how they will choose by both stating complementary strategies or one child approving of the another's proposal	Child 1: "I go straight" Child 2: "Ok, I swerve." / Child 1: "Let's both swerve." Child 2: "Ok."	0.69
Disputes	Children state opposing strategies at least three times or nothing is said after both children state opposing strategies	Child 1: "I go straight." Child 2: "No me!" Child 1: "No me!"	0.74
Rule presence	A previously established rule is followed, mentioned explicitly or implicitly, or a child protests against a prosed or actual rule violation by another child	"Remember, the blue train must go straight." / "You broke the rule."	0.94

generally good to very good agreement between coders (Landis & Koch, 1977, see Table 2 for κ coefficients).

Analysis

To analyze children's game behavior in relation to the rules they established we ran several generalized linear mixed models (GLMM, Baayen, 2008). This analytic approach has the advantage that it allows the residuals of the response measure to be non-normally distributed (e.g., to follow binomial or poisson distributions) and that it can handle complex random effect structures (e.g., to have multiple observations from the same individual or multiple observations from individuals from the same group).

All models were fitted in R (R Core Team, 2014) using the function 'glmer' of the R-package lme4 (Bates, Maechler, Bolker & Walker, 2014) and the function 'lm' for one linear model. For each model we ran several diagnostics (see Table 1 in Supplementary material for details). These were unproblematic unless otherwise stated. When fitting GLMMs, we included all random intercepts and random slopes components to keep type 1 error rates at the nominal level of 5% (Schielzeth & Forstmeier, 2009). We always first compared a full model including all test and control predictors, random intercepts, and random slopes to a null model that did not include the test predictors but was otherwise identical. This tests whether the test predictors combined had an effect on the response variable and thereby avoids multiple testing issues (this is analogous to running an ANOVA test for multi-level factors before applying post hoc tests, see Forstmeier & Schielzeth, 2011). We only considered and report the effects of individual test predictors (using likelihood ratio tests) if the full-null model comparison indicated an effect.

Results

Preliminary analysis

To validate the incentive structure of the game and to rule out that children were motivated to let the trains crash rather than to win stickers we ran a preliminary analysis looking at whether the game outcome (crash vs. coordination) predicted players' protest after the game outcome had become apparent (a model also including the trials in which both children swerved yielded almost identical results but was highly unstable). Both 5- and 8-year-olds were substantially more likely to protest after a crash than after successful coordination (5-year-olds: $\chi^2 = 17.74$, df = 1, $p < .001$; 8-year-olds: $\chi^2 = 66.54$, df = 1, $p < .001$) indicating that children showed far more discontent when they did not win any stickers.[1]

Rule creations

In line with our prediction, 8-year-olds created rules more often than 5-year-olds (Fisher's Exact Test, $p = .001$). Over the whole game, all groups of 8-year-olds, but only 50% of 5-year-olds, established a rule and 8-year-olds did so much earlier in the game (although they hardly did so without any adult prompts, see Table 3). In fact, whereas only 3 groups of 5-year-olds (16.7%) created a rule between trial 1 and trial 14 (i.e., before a rule was seeded) 16 groups of 8-year-olds (88.8%) did so.

Rules were independently created 22 times and, in total, children came up with nine different rules (see Supplementary material, Table 2, for a list of all rules created by children). Most rules followed an impartiality principle using arbitrary criteria to decide who goes straight (e.g. "always the blue train goes straight"). Some rules aimed specifically at creating equal payoff distributions (e.g. "the player with fewer stickers goes straight"). The only rule that privileged particular players ("the player whose name appears first in the alphabet goes straight") was frequently violated and abandoned after only a few trials. Some groups added further amendments to their rules. Most noteworthy, after repeated rule violations one group of 8-year-olds agreed to punish future violators by introducing sanctions

[1] This analysis was kindly suggested by an anonymous reviewer.

Table 3
Cumulative number of groups having created a rule at different experimental stages.

Experimental stage	5-year-olds	8-year-olds
Up to break 1 (trials 1 – 6)	0 (0.0%)	1 (5.5%)
Up to break 2 (trials 1 – 10)	2 (11.1%)	10 (55.5%)
Up to break 3 (trials 1 – 14)	3 (16.7%)	16 (88.8%)
Overall (trials 1 – 20)	9 (50.0%)	18 (100%)

(i.e. violators had to give up some stickers). Hence, in accordance with our prediction, children overwhelmingly established rules that were fair.

We ran an additional exploratory analysis examining whether children's agreements and disputes early in the game affected the likelihood of rule formation (e.g., because successful agreements may forestall the need to form rules). This revealed that the number of agreements and disputes children had in the first six trials of the game did not predict the number of trials children played with a rule (multiple regression, $F(2, 33) = 1.17$, $p = .324$).[2]

Consequences of rules

Fairness of payoff divisions

As a measure of how unfairly payoffs were distributed within groups we computed unfairness scores for each dyad comprised in a group by dividing the payoff difference between children by the overall dyad payoff (using only the trials played by that particular dyad). The mean of these dyadic unfairness scores produced the unfairness score of a group. Contrary to our prediction, a linear model revealed that the combined effects of age, the proportion of trials played with a rule, and their interaction (while controlling for sex) did not significantly affect the unfairness of payoff distributions (full-null-model comparison, $F_{1,30} = 1.79$, $p = .171$). We therefore did not further inspect the effects of individual predictors.

Coordination success

In general, children of both age groups coordinated their decisions at high levels and achieved mean group payoffs significantly above chance level of 2.5 payoffs per trial (i.e., the per-trial payoff children would have been obtained had they chosen randomly between swerving and going straight; 8-year-olds: mean group payoff per trial = 3.47, SD = 0.37, $t(17) = 11.15$, $p < .001$; 5-year-olds: mean group payoff per trial = 3.16, SD = 0.48, $t(17) = 5.85$, $p < .001$, see Supplementary materials, Table 5, for mean payoffs for each trial by age group).

A GLMM examining the effect of rule presence and age on group payoffs revealed that neither the interaction between rule presence and age ($\chi^2 = 0.17$, df = 1, $p > .250$) nor age significantly affected payoffs ($\chi^2 = 0.32$, df = 1, $p > .250$; while controlling for sex, trial number, the random effects of the group, dyad, and subject, and the random slopes of trial number and rules presence nested within the group). However, there was a significant main effect of rule presence indicating that – in line with our prediction – groups achieved higher payoffs when playing with than when playing without a rule ($\chi^2 = 4.27$, df = 1, $p = .039$).

Conflict

Two GLMMs controlling for age, sex, trial number, the random effect of the group, and the random slopes of trial number and rule presence nested within groups revealed that when playing with a rule children had fewer verbal disputes ($\chi^2 = 15.25$, df = 1, $p < .001$) and protested less against their partners ($\chi^2 = 19.68$, df = 1, $p < .001$) than when playing without a rule.

[2] This analysis was kindly suggested by an anonymous reviewer.

Negotiation/efficiency

Two GLMMs controlling for age, sex, trial number, the random effect of the group, and the random slopes of trial number and rule presence nested within groups revealed that when playing with a rule children made fewer negotiation attempts ($\chi^2 = 8.17$, df = 1, $p = .004$) and formed fewer explicit agreements before choosing ($\chi^2 = 5.88$, df = 1, $p = .015$) than when playing without a rule indicating that rules were related to a decrease in bargaining about how to play. By contrast, children used significantly more normative language when playing with a rule ($\chi^2 = 17.88$, df = 1, $p < .001$) suggesting that children highlighted the presence of rules and encouraged compliance. Normative protest was not significantly affected by rule presence ($\chi^2 = 3.05$, df = 1, $p < .081$).

Finally, we examined if forming a joint agreement of how to play in a given trial had an effect on coordination success (i.e., group payoffs) and whether this effect was mediated by the presence of a rule. A GLMM revealed a significant interaction between agreements and rule presence ($\chi^2 = 6.88$, df = 1, $p = .009$; while controlling for age, sex, trial number, the main effect of rule presence and the random effect of the group). Posthoc tests revealed that in the absence of a rule group payoffs significantly increased when children had previously formed an explicit agreement about how to play compared to when no agreement had been formed ($\chi^2 = 20.67$, df = 1, $p < .001$). In contrast, when playing with a rule group payoffs were high regardless of whether children had explicitly agreed on a joint strategy prior to choosing ($\chi^2 = 0.15$, df = 1, $p = .696$) indicating, once more, that rules removed the need for explicit negotiation (Fig. 2). For detailed model descriptions and model outputs including estimates and standard errors, see Supplementary material, Tables 1, 3, and 4.

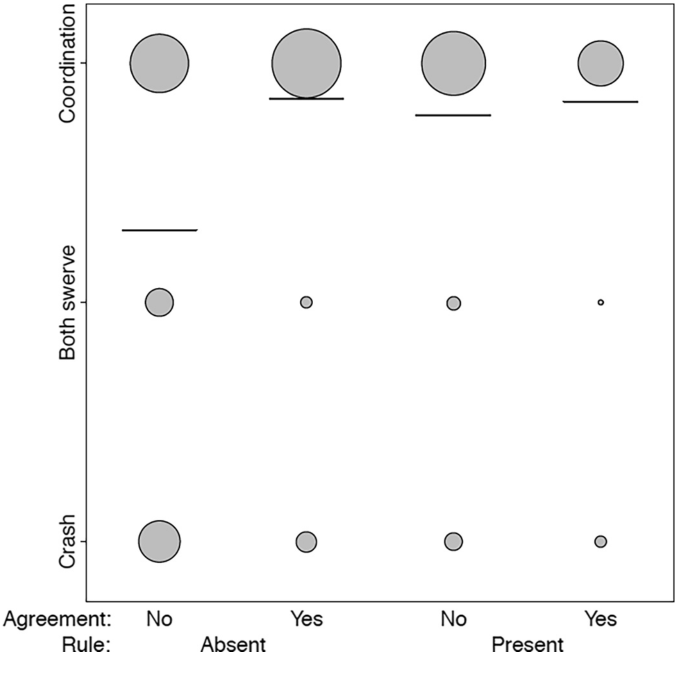

Fig. 2. Effects of agreements and rule presence on game outcomes. Circle size represents the frequency with which the specific outcomes displayed on the y-axis were achieved. Black lines display model estimates. When rules were absent, children were significantly more likely to achieve successful coordination if they had formed an explicit verbal agreement prior to choosing than when no agreement had been formed. When a rule was present, children coordinated at high levels regardless of whether they had previously agreed on a joint strategy.

Discussion

Using rules to regulate behavior is an important aspect of meeting a central challenge of human social functioning: achieving coordination in social dilemmas. While previous research has shown that preschoolers sometimes create arbitrary game rules in groups (Göckeritz et al., 2014; Nobes, 1999) the current experiment demonstrates that in response to adult prompts eight-year-old children are able to create rules with peers as a device for solving coordination problems among themselves even when their immediate interests are directly opposed.

Theoretically, the current experiment attempted to capture the way rules are often first collectively established and subsequently adhered to in societal interactions among individuals sharing mutual expectations for rule compliance. This is not to say that all rules are generated this way (e.g. conventional rules often spontaneously arise over recurring interactions without being purposefully designed; Hume, 1739; Young, 2003; Centola & Baronchelli, 2015). However, people – at least sometimes – jointly and deliberately create rules to anticipate disputes and prescribe how people ought to behave in specific situations (e.g. formal laws, official decrees, or organizational regulations, Shapiro, 2011; McAdams, 2015). A key aspect of such rules is that they apply to all group members equally and, while constraining individual self-interest, facilitate coordination one the whole.

Since 5-year-olds have previously been shown to create rules both spontaneously (Göckeritz et al., 2014) and in response to an experimenter's prompts (e.g., Hardecker et al., 2017) it is unlikely that 5-year-olds struggled with the current task due to a general inability to create rules. Instead, what differed in the current experimental setup from previous ones is that rules had to be formed and applied in the context of a social dilemma in which adequate rules inevitably limited individual children's self-interest. The greater success of 8-year-olds may therefore be related to older children's tendency to apply impartiality and fairness norms in a more principled manner even when this involves incurring a personal cost (Blake & McAuliffe, 2011; Smith et al., 2013).

Indeed, the great majority of rules that children established used impartial criteria to decide who can go straight (while a small minority of rules was specifically aimed at creating equal payoff divisions). Previous research has shown that 5-year-olds persist longer at using impartial over partial game rules suggesting that an early appreciation of the legitimacy of certain rules is already present in younger children (Grocke et al., 2015). Six-to-eight-year-olds have also been shown to prefer impartial procedures over partial ones when distributing resources between recipients (Shaw & Olson, 2014). The current study extends these findings by showing that by age 8 children can also collectively devise, agree on, and apply impartial rules, even when their own material payoffs are directly affected. By comparison, 5-year-olds appeared to have difficulties with exactly with this latter aspect when attempting to create rules: their discussions about potential rule candidates were often limited to children repeatedly claiming that they would go straight suggesting that they struggled with establishing rules that limited the pursuit of their individual goals.

Contrary to our prediction, playing with a rule did not lead to more equally distributed payoffs overall. The most likely reason for this is that in groups using arbitrary criteria to decide who goes straight the time horizon was too short for game success to average out across group members. A larger sample would also have been desirable to address this question. However, this finding also hints to the possibility that, similar to adults (Bolton, Brandts, & Ockenfels, 2005; Kimbrough, Sheremeta, & Shield, 2014; Starmans, Sheskin, & Bloom, 2017), what children consider to be fair is not necessarily an equal distribution of resources but rather a level playing field that does not systematically favor particular individuals (as also indicated by the high number of impartial rule creations).

The second main objective of this research was to examine if children can use rules to facilitate cooperation in a social dilemma. The results revealed that playing with a rule was related to some of the central functional properties that rules are generally purported to have on the societal level, namely, to facilitate coordination, mitigate conflicts, and reduce transaction costs (Lewis, 1969; Hovenkamp, 1995; Shapiro, 2011). First, rules were related to a significant increase of group payoffs suggesting that rule presence boosted coordination success. Second, when playing with a rule, children protested less against their partner's actions or statements and engaged in fewer verbal disputes.

They did, however, use more normative language to express existing rules – possibly to ensure common knowledge of their existence and to generate pressure to comply.

Finally, children's communication patterns and their relation to children's game choices indicate that playing with a rule made the coordination process more efficient: when no rule was in place children heavily relied on explicit verbal agreements of how to play to ensure successful coordination. When playing with a rule, by contrast, children generally made fewer negotiation attempts and formed fewer explicit agreements but nevertheless managed to coordinate at consistently high levels. Indeed, children sometimes did not discuss their game choices at all once a rule was present with players simply executing the move that the rule prescribed. Rules thus seemed to function as tacit mutually known agreements that removed the need to repeatedly negotiate strategies in any given round (which corresponds to the notion that such rules are self-reinforcing, Hoffman et al., 2016).

One limitation of the current study is that children were encouraged by an experimenter to think of rules and therefore children's rule creations were not entirely independent. This raises the potential objection that the observed age differences may not exclusively reflect age differences in children's abilities to form and sustain rules but also differences in the ability to comprehend and follow adult instructions to follow rules. However, this can also be seen as part of the phenomenon observed here: eight-year-olds' increased experience with rules in daily life, and as a consequence, their increased ability to create and regulate their own behavior with rules even in conflicts of interest may be precisely what gave them an advantage in interpreting and putting into action the experimenter's encouragement to think of rules. Moreover, a better understanding of the experimenter's prompts is unlikely to explain the kinds of rules children agreed upon or the functional consequences that rules had on children's communication and decision patterns. Whether or not children established rules during game breaks, which rules they agreed upon, or whether or not rules were subsequently adhered to, violated or abandoned was entirely for the group to decide. To increase their sense of independence children were always left alone when discussing rules and during trials, and the experimenters never commented on children's rules or game choices. Hence, while children did not have to come up with the idea of creating rules initially, the process of agreeing on rules as well as their subsequent application and enforcement was autonomous.

In conclusion, the current experiment shows that by eight years of age children are capable of devising group-wide rules prescribing how to behave in a social dilemma. Subsequent application of these rules facilitated coordination success, mitigated conflicts, and reduced bargaining costs which corresponds to some of the key functions that rules are thought to have in society. These findings attest to children's growing ability to regulate their behavior among peers in spite of having conflicting motives.

Acknowledgements

We thank Isabelle De Gaillande-Mustoe, Roman Stengelin, Yasmin Aydin, and Julia Gellermann for helping with participant recruitment and data collection. We thank Rebecca Koomen, Susanne Hardecker, and Jan Engelmann for helpful comments on an earlier draft of the manuscript, Bahar Koymen for helpful discussions, and Roger Mundry for helping with the statistical analysis. Finally, we thank all children, parents, and daycare centers for their friendly cooperation.

Funding

This work was supported by the Max Planck Society.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2018.11.001.

References

Baayen, R. H. (2008). *Analyzing linguistic data*. Cambridge: Cambridge University Press.

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7, http://CRAN.R-project.org/package=lme4.

Blake, P. K., & McAuliffe, K. (2011). ‘‘I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition, 120*, 215–224. https://doi.org/10.1016/j.cognition.2011.04.006.

Bolton, G. E., Brandts, J., & Ockenfels, A. (2005). Fair procedures: Evidence from games involving lotteries. *The Economic Journal, 114*, 1054–1076. https://doi.org/10.1111/j.1468-0297.2005.01032.x.

Bornstein, G., Budescu, D. V., & Zamir, S. (1997). Cooperation in intergroup, n-person, and two-person games of chicken. *Journal of Conflict Resolution, 41*, 384–406. https://doi.org/10.1177/0022002797041003003.

Centola, D., & Baronchelli, A. (2015). The spontaneous emergence of conventions: An experimental study of cultural evolution. *PNAS, 112*, 1989–1994. https://doi.org/10.1073/pnas.1418838112.

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology, 65*, 47–55. https://doi.org/10.1007/s00265-010-1038-5.

Gilbert, M. (1989). *On social facts*. New York: Routledge.

Göckeritz, S., Schmidt, M. F. H., & Tomasello, M. (2014). Young children's creation and transmission of social norms. *Cognitive Development, 30*, 81–95. https://doi.org/10.1016/j.cogdev.2014.01.003.

Grocke, P., Rossano, F., & Tomasello, M. (2015). Procedural justice in children: Preschoolers accept unequal resource distributions if the procedure provides equal opportunities. *Journal of Experimental Child Psychology, 140*, 197–210. https://doi.org/10.1016/j.jecp.2015.07.008.

Grueneisen, S., & Tomasello, M. (2017). Children coordinate in a social dilemma by taking turns and along dominance asymmetries. *Developmental Psychology, 53*, 265–273. https://doi.org/10.1037/dev0000236.

Hardecker, S., Schmidt, M. F., & Tomasello, M. (2017). Children's developing understanding of the conventionality of rules. *Journal of Cognition and Development, 18*, 163–188.

Helbing, D., Schönhof, M., Stark, H. U., & Holyst, J. A. (2005). How individuals learn to take turns: Emergence of alternating cooperation in a congestion game and the prisoners dilemma. *Advances in Complex Systems, 8*, 87–116. https://doi.org/10.1142/S0219525905000361.

Hoffman, M., Yoeli, E., & Navarrete, C. (2016). Game theory and morality. In T. Shackelford (Ed.), The evolution of morality.

House, B., Silk, J. B., Henrich, J., Barrett, C., Scelza, B., Boyette, A., et al (2013). The ontogeny of prosocial behavior across diverse societies. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 14586–14591. https://doi.org/10.1073/pnas.1221217110.

Hovenkamp, H. (1995). Law and economics in the United States: A brief historical survey. *Cambridge Journal of Economics, 19*, 331–352.

Hume, D. ((1739).). *A treatise of human nature: being an attempt to introduce the experimental method of reasoning into moral subjects*. London: John Noon.

Josephs, M., Kushnir, T., Gräfenhain, M., & Rakoczy, H. (2016). Children protest moral and conventional violations more when they believe actions are freely chosen. *Journal of Experimental Child Psychology, 141*, 247–255. https://doi.org/10.1016/j.jecp.2015.08.002.

Kenward, B., Karlsson, M., & Persson, J. (2011). Over-imitation is better explained by norm learning than by distorted causal learning. *Proceedings of the Royal Society B, 278*, 1239–1246. https://doi.org/10.1098/rspb.2010.1399.

Kimbrough, E. O., Sheremeta, R. M., & Shield, T. W. (2014). When parity promotes peace: Resolving conflict between asymmetric agents. *Journal of Economic Behavior & Organization, 99*, 96–108. https://doi.org/10.1016/j.jebo.2013.12.021.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.

Lewis, D. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.

McAdams, R. (2015). *The expressive powers of law: Theories and limits*. Cambridge, MA: Harvard University Press.

Nobes, G. (1999). Children's understanding of rules they invent themselves. *Journal of Moral Education, 28*, 215–232. https://doi.org/10.1080/030572499103232.

Orvell, A., Kross, A. E., & Gelman, S. A. (2018). That's how "you" do it: Generic *you* expresses norms during early childhood. *Journal of Experimental Child Psychology, 165*, 183–195. https://doi.org/10.1016/j.jecp.2017.04.015.

Piaget, J. (1932). *The moral judgement of the child*. London: Routledge & Kegan Paul.

Rapoport, A., & Chammah, A. M. (1966). The game of chicken. *American Behavioral Scientist, 10*, 10–28.

Rapoport, A., Chammah, A. M., & Orwant, C. J. (1965). *Prisoner's dilemma: A study in conflict and cooperation*. Ann Arbor, MI: University of Michigan Press.

Core Team, R. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology, 44*, 875–881. https://doi.org/10.1037/0012-1649.44.3.875.

Riggs, A. E., & Young, A. G. (2016). Developmental changes in children's normative reasoning across learning contexts and collaborative roles. *Developmental Psychology, 52*, 1236–1246.

Roberts, S. O., Ho, A. K., & Gelman, S. A. (2017). Group presence, category labels, and generic statements influence children to treat descriptive group regularities as prescriptive. *Journal of Experimental Child Psychology, 158*, 19–31.

Schielzeth, H., & Forstmeier, W. (2009). Conclusions beyond support: Overconfident estimates in mixed models. *Behavioral Ecology, 20*, 416–420. https://doi.org/10.1093/beheco/arn145.

Schmidt, M. F. H., Butler, L. P., Heinz, J., & Tomasello, M. (2016). Young children see a single action and infer a social norm: Promiscuous normativity in 3-year-olds. *Psychological Science, 27*, 1360–1370.

Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2011). Young children attribute normativity to novel actions without pedagogy or normative language. *Developmental Science, 14*, 530–539. https://doi.org/10.1111/j.1467-7687.2010.01000.x.

Schmidt, M. F. H., Rakoczy, H., Mietzsch, T., & Tomasello, M. (2016). Young children understand the role of agreement in establishing arbitrary norms – but unanimity is key. *Child Development, 87*, 612–626. https://doi.org/10.1111/cdev.12510.

Shapiro, S. (2011). *Legality*. Cambridge, MA: Harvard University Press.
Shaw, A., & Olson, K. (2014). Fairness as partiality aversion: The development of procedural justice. *Journal of Experimental Child Psychology, 119*, 40–53. https://doi.org/10.1016/j.jecp.2013.10.007.
Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*, 1359–1366. https://doi.org/10.1177/0956797611417632.
Smetana, J. (1981). Preschool children's conceptions of moral and social rules. *Child Development, 52*, 1333–1336.
Smith, C., Blake, P. R., & Harris, P. L. (2013). I should but I won't: Why young children endorse norms of fair sharing but do not follow them. *PloS One, 8*. e59510.
Starmans, C., Sheskin, M., & Bloom, P. (2017). Why people prefer unequal societies. *Nature Human Behaviour, 1*, 0082. https://doi.org/10.1038/s41562-017-0082.
Sugden, R. (1986). *The economics of rights, cooperation and welfare*. Oxford: Basil Blackwell.
Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, UK: Cambridge University Press.
Young, H. P. (2003). The power of norms. In P. Hammerstein (Ed.), *Genetic and cultural evolution of cooperation*. Cambridge MA: MIT Press.
Zhao, X., & Kushnir, T. (2018). Young children consider individual authority and collective agreement when deciding who can change rules. *Journal of Experimental Child Psychology, 165*, 101–116. https://doi.org/10.1016/j.jecp.2017.04.004.