Children's Norm Enforcement in Their Interactions With Peers

Bahar Köymen
Max Planck Institute for Evolutionary Anthropology

Elena Lieven
Max Planck Institute for Evolutionary Anthropology and The University of Manchester

Denis A. Engemann
Institute of Neuroscience and Medicine—Cognitive Neuroscience (INM-3) and University Hospital of Cologne

Hannes Rakoczy
University of Göttingen

Felix Warneken
Harvard University

Michael Tomasello
Max Planck Institute for Evolutionary Anthropology

This study investigates how children negotiate social norms with peers. In Study 1, 48 pairs of 3- and 5-year-olds ($N = 96$) and in Study 2, 48 pairs of 5- and 7-year-olds ($N = 96$) were presented with sorting tasks with conflicting instructions (one child by color, the other by shape) or identical instructions. Three-year-olds differed from older children: They were less selective for the contexts in which they enforced norms, and they (as well as the older children to a lesser extent) used grammatical constructions objectifying the norms ("It works like this" rather than "You must do it like this"). These results suggested that children's understanding of social norms becomes more flexible during the preschool years.

Most everyday human actions are normatively structured within social practices such that people act in ways consistent with social expectations. Through participating in daily interactions with their caregivers and peers, children learn about the norms of their social groups (Schieffelin & Ochs, 1986; Tomasello, 1999, 2009; Turiel, 1983, 1998). This study investigates children's understanding of conventional norms regarding game rules through looking at their norm enforcement in their spontaneous peer interactions.

According to Piaget (1932/1965), children's understanding of social rules begins with egocentric or objectivist stages in which they view norms as non-negotiable, unalterable facts and only later in middle childhood do they arrive at a mature understanding of social norms as a product of rational negotiation between equals. In contrast to the Piagetian view, research within the "domain theory" framework has shown that very early on at preschool age, children are able to distinguish different "domains" of social norms (Helwig & Turiel, 2003; Killen, 1991; Smetana, 2006; Turiel, 1998). Many empirical studies using interview methods have demonstrated that preschoolers differentially evaluate moral norm violations (e.g., hitting someone) and conventional norm violations (e.g., eating a snack while standing instead of sitting), and rate the latter as less serious, less punishable, more context dependent, and more authority dependent (Nucci & Nucci, 1982a, 1982b; Smetana, 1981; Smetana & Braeges, 1990; Smetana, Schlagman, & Adams, 1993; Turiel, 1983). With age, children reason more flexibly about transgressions because they are able to take into account more factors such as interpersonal concerns and group functioning (Horn, 2003; Killen & Stangor, 2001; Smetana, 2006). For instance, Killen and Stangor (2001) have shown that, as compared to younger school-aged children, older children justify exclusion more if there is a threat to group

functioning, although all age groups rate straightforward exclusion of a child from a group as unfair.

In addition, some recent experimental studies have documented that from early on, young children enforce social norms to third parties. A series of studies investigated children's interventions to a puppet committing a mistake within pretense and other rule-governed games (Rakoczy, 2008; Rakoczy, Brosche, Warneken, & Tomasello, 2009; Rakoczy, Warneken, & Tomasello, 2008; Wyman, Rakoczy, & Tomasello, 2009). The converging evidence suggests that both 2- and 3-year-olds intervene and protest against the mistakes of the puppet. Three-year-olds particularly display normative protests (e.g., teaching, critiquing) using normative language such as "No! It does not go like this!" (Rakoczy et al., 2008, p. 877), whereas 2-year-olds' protests are more descriptive and imperative such as "No! Not in this hole!" (Rakoczy et al., 2008, p. 877). While enforcing rules to third parties, 3-year-olds show some understanding about the context specificity of the rules and evaluate the same action as a norm violation in certain contexts but not in others. For instance, Wyman et al. (2009) created two pretend play contexts in which the same object, a yellow stick, was used as a toothbrush in one and as a carrot in the other. Three-year-olds normatively protested when the puppet used the contextually inappropriate pretend identity.

One limitation in both experimental and interview methods is that they rely on noninteractive measures by analyzing children's one-shot responses to norm violations of a puppet or hypothetical characters who do not respond back or challenge children's beliefs about norms. Some observational studies have analyzed how children enforce norms within their spontaneous peer conflicts, which are a significant part of peer interactions (Corsaro, 1994; Kyratzis, 2004). Studies have documented that in their peer conflicts, children refer to norms of property entitlements or possessions, turn taking, or aggression (Dunn & Munn, 1987; Hay, 1984; Hay & Ross, 1982; Ingram & Bering, 2010; Much & Shweder, 1978; Shantz, 1987; Smetana, 1984). In facing such conflicts, children are responsive to their peers' protests and attempt to resolve the conflict (Hartup, French, Laursen, Johnston, & Ogawa, 1993; Hartup, Laursen, Stewart, & Eastenson, 1988; Hay & Ross, 1982; Killen & Turiel, 1991; Shantz, 1987). Nucci and Nucci (1982a) demonstrated that school-aged children respond to peers' moral transgressions (e.g., throwing sand at someone) by explaining the effect of the act on victim's rights and welfare, such as *You got it in my eyes. It hurts like hell* (p. 1339), and to

conventional transgressions (e.g., spitting on the grass) by stating the normative rule such as *You're not supposed to spit* (p. 1339).

In the studies on children's naturalistic peer interactions, children's response types to norm violations were mostly categorized into functional categories related to the content of the message such as "injury or loss statement," but how these normative expressions were formulated was not analyzed linguistically. For instance, the way children mark agency in their [normative] formulations might reveal their perspective on an event (Berman & Slobin, 1994; Budwig, 1995; Kyratzis, 2009). When a child uses a transitive construction (an animate agent is acting on an object), *You must put the pen here*, they report the act from a speaker-subjective point of view: how things are seen by the speaker or what seems right for the speaker. When they use inanimate subjects, *The pen goes here*, they report the same act from a more factual point of view: how things are seen by everyone or what seems true for everyone (Berman & Slobin, 1994). Although both statements are normative, in the latter case, known as "middle constructions," the same action is more objectified and the role of the speaker or the causal agent is downgraded when inanimate subjects are used (Budwig, Stein, & O'Brien, 2001; Kemmer, 2003).

Although observational methods take into account the interactive nature of normative peer conflicts, they still have limitations. Investigators do not have control over the peer conflict in terms of which child knows what about the norm, the type of norm violations (or the type of normative conflict), the number of parties involved in the conflict, so forth. Most importantly, it is very difficult to draw conclusions about developmental changes in the understanding and use of norms because the contexts of peer conflicts are often not comparable at different ages. One cannot compare how long it takes children to resolve a normative conflict or what kinds of linguistic strategies they deploy in their norm enforcements, if the sampled normative conflicts are not analogous. One major gap in the literature, therefore, is the analysis of children's normative disputes with peers in highly interactive, but still experimentally controlled settings.

In this study, therefore, we experimentally created two contexts in which children had to interactively negotiate the rules of a game in dyadic peer conversations. In these two contexts, children of different ages (3, 5, and 7 years) interacted with one another in the same context and had equal footing such that they all learned the same game with the same rules from the same adult sources. We elicited

peer conflicts to observe possible developmental changes in how normative disputes are resolved and what linguistic constructions (e.g., marking agency) children deploy. We used a methodology similar to that of Hartup et al. (1993) and taught children conflicting rules of a sorting game comparable to the Dimensional Change Cart Sort used by Zelazo, Frye, and Rapus (1996). In Study 1, pairs of 3- and 5-year-olds were presented with simple sorting games. In Study 2, pairs of 5- and 7-year-olds were presented with complex sorting games. In both studies, in the incompatible context, one child was taught to sort the items by shape, and the other by color, such that the placement of the first item by a child would be a rule violation for his or her partner. In the other context, the compatible context, both children were taught to sort the items by both color and shape such that the placement of the first item by a child would not necessarily be a rule violation, but the appropriateness of that action would depend on which game the children decide to play and would require children to take the perspective of their peer partners. We first analyzed whether children distinguished these two contexts through looking at the presence of normative conflicts. We hypothesized that normative conflicts would take place in the incompatible context more than the compatible context. The difference would be greater in the older age groups because they would be better at discriminating the contexts where an action is wrong from those contexts where an action is not necessarily wrong. Next, we analyzed how easily the children agreed on a rule by looking at how long it took them to agree, and predicted that they would agree in the compatible context sooner than incompatible context. Older children would agree on a rule sooner than younger children because they would be more flexible in seeing alternative rules as acceptable. And finally, we explored how children marked agency as a discursive strategy in their normative expressions to see whether they objectified the rules or they used more person-oriented statements.

These age groups were selected because the literature shows that by 3 years of age children are able to follow and detect violations of game rules (Kalish, Weissman, & Bernstein, 2000; Rakoczy et al., 2008; Zelazo et al., 1996), by 5 years of age they are able to take the perspective of their interactive partners (Wellman, Cross, & Watson, 2001), and in middle childhood (ages 7 and 8 years) children show some further flexibility in their reasoning on moral norms and social conventions (Conry-Murray & Turiel, 2012; Horn, Killen, & Stangor, 1999; Killen & Stangor, 2001; Theimer, Killen, & Stangor, 2001).

Study 1

Method

Participants

Forty-eight 3-year-olds ($M = 3;9$, range $= 3;5–3;11$, 24 boys and 24 girls) and forty-eight 5-year-olds ($M = 5;8$, range $= 5;5–5;11$, 24 boys and 24 girls) participated in Study 1. The 96 children were grouped into 48 dyads (24 dyads of 3-year-olds and 24 dyads of 5-year-olds) matched by gender. Within the 3-year-olds, 11 dyads participated in the compatible context, and 13 dyads in the incompatible context. Within the 5-year-olds, 13 dyads participated in the compatible context, and 11 dyads in the incompatible context. The dyads were formed on the basis of the recommendations of their teacher and were composed of children who were friends. All children were monolingual German speakers, except for the two 3- and two 5-year-old children who were bilinguals. The second languages of these bilingual children were Arabic, Russian, English, and French.

Materials

There were two sorting games to control for the possible difficulty or preference for one of the sorting dimensions (color and shape).

Board game. Children were instructed to sort the 14 magnets that came out of a dispenser according to animal shape (a hedgehog or a bird) and/or color (yellow or green). There were four types of magnets: a yellow hedgehog, a green hedgehog, a yellow bird, and a green bird. Children were asked to place these tokens on a 56.5×34 cm magnetic board on which there were two rows. On the head of one row there was a picture of a yellow hedgehog, and on the other row there was a picture of a green bird, signaling one level of each of the two sorting dimensions.

Tube game. Children were instructed to sort the 14 wooden tokens that came out of a dispenser according to geometric shape (ball or cube) and/or color (blue or red). There were four types of wooden tokens: a blue ball, a red ball, a blue cube, and a red cube. Children were asked to place these tokens into two 6×29.7 cm plexiglass tubes on a platform. In front of one tube there was a red cube, and in front of the other tube there was a blue ball,

signaling one level of each of the two sorting dimensions.

Procedure

The study took place in a quiet room of preschools in a midsize German city. Each dyad played both games and the order of the games was counterbalanced. There were two sessions for each game: individual teaching phase and testing phase. The entire session lasted around 45 min and all sessions were videotaped.

In the individual teaching phase of the *compatible context*, both sets of rules (the shape rule and the color rule) were presented to each child and the presentation order of the rules were counterbalanced. We will describe the procedure in which the board game was played first and the color rule was introduced before the shape rule. The first child, Child A, went through the individual teaching phase, while the other child, Child B, was outside with the second experimenter. The experimenter told Child A, "We will play two games. First, we will play the color game" and instructed Child A to sort the 10 tokens by color within the board game ("All the green ones go here and the yellow ones go here. The green one never goes here, because they all go there and the yellow ones go there"). The first 2 tokens were demonstrated by the experimenter, the next 2 tokens were sorted by the experimenter and the child together, and the rest of the 6 tokens were sorted by the child alone. If the child made a mistake, the experimenter reminded him or her of the rule. Children did not have problems understanding the rules. Then the experimenter introduced the second set of rules within the board game to Child A, "We just played the color game. Now we will play the animal game," and repeated the same instructions for sorting the same tokens by animal shape. After Child A left the room, Child B entered. The same instructions were delivered but the order of the rules was reversed such that Child B sorted the magnets by animal first and by color next. In other words, the last rule each child learned was different. This constituted the first teaching phase. In the first testing phase, both children were in the room together. The experimenter told them, "Now you will play together" and left the room. The children played together and sorted the 14 tokens. Then the same procedure was repeated for the tube game. This time, Child B went through the teaching phase first and was taught to sort the wooden tokens by color first and by shape next. Then, Child A was taught to sort the wooden

tokens by shape first and by color next. Eventually, if a child learned the color rule first and shape rule next in the first game, he or she learned the shape rule first and the color rule next in the second game and vice versa.

In the *incompatible context*, the individual teaching phases were identical to the compatible context, except that only one set of rules was presented to each child. We will again describe the procedure in which the board game was played first. The experimenter told Child A, with Child B out of the room, "We will play the board game" and instructed the child to sort the tokens by color. Then, with Child A out of the room, the experimenter told Child B, "We will play the board game" as well; however, the experimenter instructed the child to sort the tokens by shape. In the first testing phase, the experimenter told both children, "Now you will play the board game together" and left the room. The children sorted the 14 tokens that came out of the dispenser. Then the same procedure was repeated for the tube game. This time, Child B went through the teaching phase first and was instructed to sort the tokens by color, and then Child A was taught to sort the tokens by shape. In other words, if a child learned the color rule in the first game, he or she learned the shape rule in the second game and vice versa. In the second testing phase, the experimenter told the children "Now you will play the tube game together." The children sorted the 14 tokens that came out of the dispenser.

The children were not told anything about what instructions their peer received. The order of 10 tokens in the teaching phase and 14 tokens in the testing phase were the same across dyads. One 5-year-old in the compatible condition refused to play the tube game so that session was dropped from the analyses. Thus, the data set has 95 sessions from 48 dyads.

Coding

All the sessions were transcribed using the transcription conventions of CLAN (Macwhinney, 2000). Each line in a transcript depicted an utterance with a maximum of one clause. "Utterance" was defined as a clause (complex sentences were divided into several lines), a conversational turn, or a group of words separated from one another by a pause of 2 s. Exact repetitions within a conversational turn were excluded from the analyses. We first identified the point at which the children reached agreement, and then extracted the on-task protest utterances until agreement. Finally, we

coded each on-task protest utterance for normativity, agency, and transitivity. A second coder recoded 22% of the transcripts for all the variables. The agreement between the two coders is reported at the end of the respective sections: (a) agreement on a rule, (b) extraction of on-task utterances, (c) extraction of protest utterances, and (d) normativity, agency, and transitivity.

Agreement on a rule. The first coder marked the resolution in the transcript. Agreement was operationalized as any kind of verbal agreement (e.g., *Ja Ente zu Ente* "Yes ducks with ducks") and/or two consecutive compatible moves in the game (nonverbal agreement). In about 10% of the sessions, children did not reach agreement or they agreed to disagree, each playing by their own rule or moving to off-task games. In these cases, the agreement was marked at the point when children stopped protesting to one another. If children never reached agreement and continued to negotiate the rules until they were out of tokens to sort, then the agreement was marked at the end of the session. For the reliability on resolution, a second coder marked the episodes for when the children reached agreement. We first determined the positions in the sequences of utterances at which the two coders indicated that the children reached agreement. In the next step, we calculated the absolute difference between these positions per episode and then averaged these values across episodes. We used this value as the test statistic in the following permutation test that we ran to test its significance (Adams & Anthony, 1996; Manly, 1997). For this, we randomly allocated the decision about agreement of one of the two coders along the sequence of utterances, separately for each episode, and then determined the test statistic again. We permuted the data 1,000 times (including the original data as one permutation) and determined the p value as the proportion of permutations revealing a test statistic at least as small as that of the original data such that the test statistic, which was .90, can be treated as equivalent to the κ value (test statistic: .90, expected = 14.79, $p < .01$).

Extraction of on-task utterances. All the utterances until agreement were first coded for being on task. On-task utterances were defined as all utterances related to the rules of the game. Due to our linguistic analyses, two types of on-task utterances were differentiated: clause-level (*Hier kommen alle Igel hin* "All hedgehogs go here") and non-clause level on-task utterances (*Murmeln!* "Marbles!"). Off-task utterances were defined as utterances that were not about the game or its rules (*Das ist eine Kamera* "That is a camera"). The reliability on this was done with the five levels: clause-level on-task utterances, non-clause-level on-task utterances, off-task utterances, exact repetitions, and unintelligible utterances. The agreement was κ = .84.

Extraction of protest utterances. Next the protest utterances were identified within the on-task utterances. Protest utterances were operationalized as any kind of disagreements with the prior statement or the prior action about the rule of the game until the children reach agreement. Protests could be in the form of highlighting the mistake of the prior move, *Nein, das kommt nicht da hin* "No it doesn't go there" or asserting the alternative rule, *Grün gehört dahin* "Green belongs there." The agreement was κ = .79.

Normativity, agency, and transitivity. Each clause-level protest utterance was coded for the following three dimensions: (a) normativity, (b) agency, and (c) transitivity.

Normativity: The coding of the normativity in this study was mostly based on the coding scheme of Rakoczy et al. (2008), which defined normative protests as utterances that involve protest, critique, and teaching with the use of normative vocabulary. More specifically, an utterance was coded as normative if:

1. it stated a general rule about the game with the verbs *kommen* "come," *gehen* "go," and *gehören* "belong" in present tense such as *Nein, Grün kommt auf Grün* "No, green comes with green" or *Das gehört zu Gelb* "That belongs to yellow";
2. it had deontic modal verbs *müssen* "must" or "have to," *sollen* "ought to," *dürfen* "allowed to" such as *Das muss hier hin* "It must (go) in here";
3. it had some normative adjectives or adverbs such as *Du hast das falsch gemacht* "You did that wrong," *Das stimmt nicht* "That is not correct," *Das ist richtig* "That is right";
4. it was a formulaic normative expression such as *Es geht nicht so* "It doesn't go like that" and *Das machen wir so* "We do it like that."

All other protest utterances were coded as non-normative and included utterances that had:

1. imperatives such as *Gib mir den Vogel* "Give me the bird";
2. nondeontic modals *möchten* "would like to," *wollen* "want to," such as *Ich will Vogel machen* "I want to do the birds";

3. descriptive protests such as *Ich habe das anders gespielt* "I played it differently" and *Es ist aber grün* "But it is green."

We then categorized each session for the presence of normative conflict. If each child in a dyad displayed at least one normative protest, that episode was considered as an episode with normative conflict and all others were coded as non-normative. The agreement on normativity of on each clause-level protest utterance was κ = .94.

Agency: The subject of each protest utterance was coded as inanimate if it referred to an object or an action and animate if it referred to a person (including the generic person *Man* "One"). The agreement on agency was κ = 1.00.

Transitivity: There were five categories. The utterances were coded as:

1. *intransitive*, if there was only one participant (the grammatical subject), such as *Ich gehe* "I go" or one implied participant *Komm* "(You) Come," as well as the formulaic expression *Es geht nicht so* "It doesn't go like that";
2. *middle*, if the utterances were intransitives with inanimate subjects and highly kinetic verbs ("come," "go") that involve some motion and direction. These included the verbs *kommen* "come," *gehen* "go" (when they co-occurred with noun phrases within nonformulaic utterances), *gehören* "belong," and modals with locative such as *Grün kommt hier hin* "Green comes here" and *Es muss hier hin* "It must (go) in here";
3. *transitive*, if there were two participants (the grammatical subject and the grammatical object) such as *Ich brauche die Rote* "I need the red," *Doch, so hat der gesagt* "Yes, he did say that," or two implied participants *Frag* "(You) ask (that/him)";
4. *sein* "to be," if there were copular constructions with the verb *sein* "to be" such as *Das ist ein Igel* "That is a hedgehog";
5. *other*, if the utterances were incomplete or ambiguous in terms of whether they were transitive or not such as *Du kannst* "You can."

The agreement on transitivity was κ = .91.

Results

There were four analyses: (a) presence of normative conflict in each session, (b) number of utter-ances it took children to reach agreement, (c) marking of agency within normative protests, and (d) the co-occurrence patterns of agency, normativity, and transitivity. All the statistical analyses were run using R (version 2.15.1; R Development Core Team, 2012). There were no gender differences in any of the measures so gender was not included in the models for the sake of parsimony.

Presence of Normative Conflict

We first tested whether an episode had a normative conflict, using generalized linear mixed model (GLMM; Baayen, 2008) fit by the Laplace approximation. The analyses were run using the function lmer of the statistics package lme4 (Bates, Maechler, & Bolker, 2012). To test the significance of the full model, we compared its fit with that of a null model using a likelihood ratio test. The models were fit with binomial error structure and logit link function. The response variable was a binary measure, session with a normative conflict versus session with a non-normative conflict (or no conflict at all). The full model included all the predictors of interest: age group (3 vs. 5), context (compatible vs. incompatible), and their interaction, as well as the control predictors: the game (tube game vs. board game), the order of these games (1 vs. 2), and their interaction. The random factor of dyad ($N = 48$) was also added to the model. The null model comprised only the control predictors (game, task order, and their interaction) and random factor of dyad.

The full model improved the fit as compared to the null model ($\chi^2 = 29.59$, $df = 3$, $p < .001$). The significant interaction between age group and context suggested that both age groups had more normative conflicts in the incompatible context than in the compatible conflicts, but this difference was greater for the 5-year-olds ($B = 2.79$, $SE = 1.36$, $z = 2.06$, $p = .04$; see the top panel in Figure 1). Namely, 3-year-olds had normative conflicts in the compatible context more than 5-year-olds did.

Number of Utterances Until Agreement

To test how long it took the children to agree on a rule in each session, we used a repeated measures analysis of variance (ANOVA) with one within-subjects factor (game with two levels: board and tube) and three between-subjects factors: (a) age group with two levels: 3 years versus 5 years; (b) context with two levels: compatible versus incompatible; and (c) order of game with two levels: board game first versus tube game first. The response variable was the
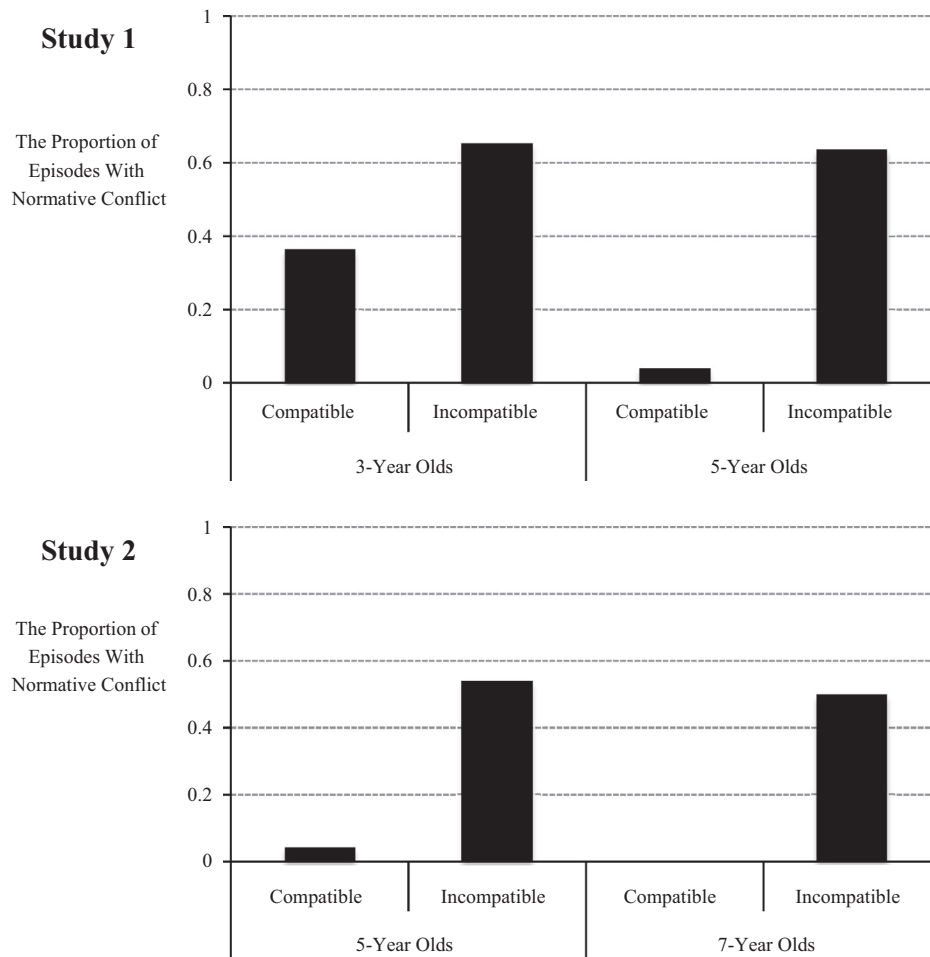
Figure 1. Proportion of episodes with normative conflict in Study 1 and Study 2.

total number of on-task utterances (both at the clause level and nonclause level and regardless of whether they were protest utterance or not) by both children in the dyad until they reached agreement. Because there were some episodes in which children had non-verbal agreements, and the number of utterances until agreement was zero, we log transformed the response variable after adding the constant 1. Through this log transformation, the assumptions of normal distribution and homogeneity of residuals were fulfilled by visual inspection of a qq-plot and residuals plotted against fitted values.

The repeated measures ANOVA was carried out with 47 dyads because one dyad (in the 5-year-old group in the compatible context) did not complete the tube game. The significant main effect of context suggested that it took both 3- and 5-year-olds significantly longer to reach agreement in the incompatible context than in the compatible context, $F(1, 42) = 52.51$, $p < .001$, $\eta_P^2 = .56$; see the left panel

in Figure 2. The main effect of age group showed a trend suggesting that it took 3-year-olds somewhat longer to reach agreement than 5-year-olds, $F(1, 42) = 3.23$, $p = .08$, $\eta_P^2 = .07$. The interaction between the context and age group was not significant.

Marking of Agency Within Normative Protests

We compared the co-occurrence patterns of normative protests with animate versus inanimate subjects across the age groups and the two contexts, using GLMM fit by the Laplace approximation. The response variable was the proportion of normative protests to all protests (normative and nonnormative). To test the significance of the full model, we compared its fit with that of a null model using a likelihood ratio test. The models were fit with binomial error structure and logit link function. The full model included all the predictors of interest, that is, age group (3 vs. 5), context
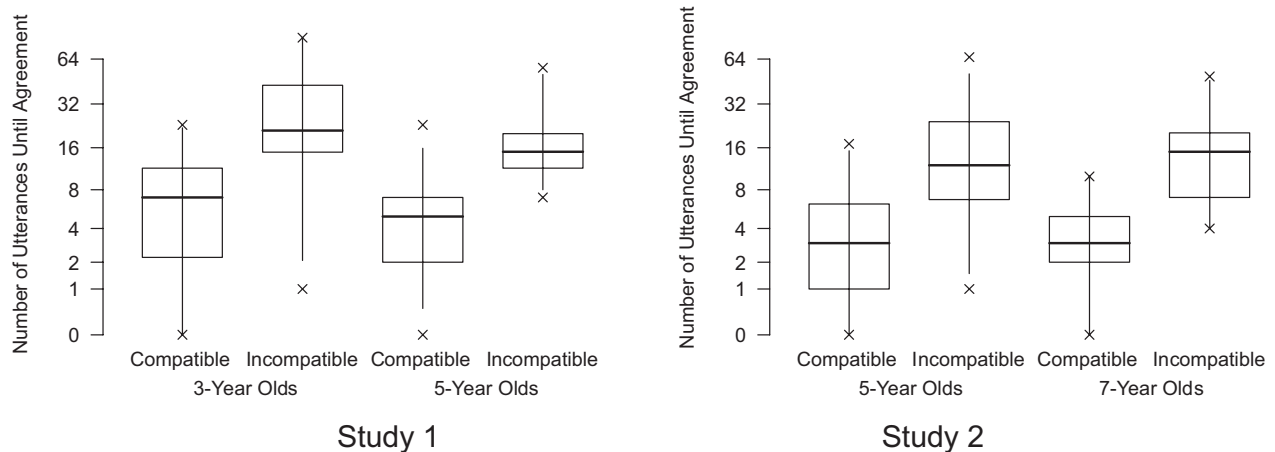
Figure 2. The number of utterances until agreement in Study 1 and Study 2. The graph displays the log-transformed values, and the numbers on the *y*-axis show the actual numbers in Study 1 and Study 2. The points indicated by "x" refer to the minimum and maximum values. The boxes indicate the values that fall into the range from the first to the third quartile. The horizontal line within each box is the median. The vertical lines below and above the boxes show the values that go from the 2.5th to the 97.5th percentile.

(compatible vs. incompatible), agency (animate vs. inanimate), and the interaction between agency and age group, as well as the control predictors of the game (tube game vs. board game), the order of these games (1 vs. 2), and their interaction. The random factors in the models were dyad and games nested in dyads (see the explanation below). There were many sessions without protests, especially in the compatible context with 5-year-olds so there were not enough data points in the compatible context to make comparisons of interaction terms involving the context (e.g., age group and context; agency and context). Thus, we did not include interaction terms involving the context in the full model. The null model comprised only the control predictors (game, the game order, and their interaction) and the two random factors. In each model we weighted the response by the number of utterances per dyad.

Thirty-three dyads of 48 showed protests. As there were two sessions for each dyad (one for each game), not all 33 dyads showed protests in both sessions and there were 58 sessions with protests. Thus, the random factors in the models were dyad ($N = 33$) and games nested in dyads ($N = 58$).

The full model improved the fit as compared to the null model ($\chi^2 = 2{,}558.4$, $df = 4$, $p < .001$); the significant interaction between agency and age group suggested that both 3- and 5-year-olds used inanimate subjects more than animate subjects in their normative protests. This difference was greater for 3-year-olds who used inanimate subjects

in their normative protests almost exclusively, whereas 5-year-olds were relatively more flexible in their choices in the agency of the subjects in their protests ($B = -0.93$, $SE = 0.23$, $z = -4.13$, $p < .001$).

Agency and Transitivity

To see which type of constructions these animate and inanimate subjects co-occur in, all the 644 protest utterances were coded for their transitivity, which had five levels: utterances with copular (constructions with *sein* "be"), transitives, middles, intransitives, and other.

In their normative protests both 3- and 5-year-olds used predominantly middle constructions with inanimate subjects, such as *Nein, die rote Murmel kommt hier rein* "No the red marble goes in here," 80% and 66% of the time, respectively (see the top panel of Table 1). Three-year-olds almost exclusively used middles in their normative protests, whereas 5-year-olds seemed a bit more discursively flexible and used other constructions in their normative protests such as intransitives with inanimate subjects (*Das stimmt nicht* "That is not correct"; 14%) and transitives with animate subjects (*Du musst die blauen rein* "You have to put the blue ones in"; 13%). In their non-normative protests both 3- and 5-year-olds used mostly transitive constructions with animate subjects such as *Gib mir den Vogel* "Give me the bird" and some copular constructions with inanimate subjects such as *Das sind doch keine Vögel* "These are not birds."

Table 1

The Distribution of Protest Utterances Across Agency and Construction Types

| | 3-year-olds | | | | 5-year-olds | | | |
| | Normative | | Non-normative | | Normative | | Non-normative | |
	Animate	Inanimate	Animate	Inanimate	Animate	Inanimate	Animate	Inanimate
Study 1								
Sein "be"	0% (0)	3% (7)	2% (4)	31% (76)	0% (0)	7% (6)	2% (2)	32% (28)
Intransitive	0% (0)	6% (13)	5% (13)	1% (2)	0% (0)	14% (12)	3% (3)	0% (0)
Middle	0% (0)	80% (181)	0% (0)	0% (0)	0% (0)	66% (56)	0% (0)	0% (0)
Transitive	9% (20)	0% (0)	56% (137)	0% (0)	13% (11)	0% (0)	59% (52)	2% (2)
Other	0% (0)	3% (6)	2% (6)	2% (4)	0% (0)	0% (0)	1% (1)	0% (0)
Total	9% (20)	91% (207)	66% (160)	34% (84)	13% (11)	87% (84)	66% (58)	34% (30)

| | 5-year-olds | | | | 7-year-olds | | | |
| | Normative | | Non-normative | | Normative | | Non-normative | |
	Animate	Inanimate	Animate	Inanimate	Animate	Inanimate	Animate	Inanimate
Study 2								
Sein "be"	0% (0)	3% (3)	7% (8)	20% (22)	0% (0)	11% (12)	3% (4)	13% (19)
Intransitive	0% (0)	9% (9)	4% (4)	4% (4)	1% (1)	6% (6)	6% (9)	2% (3)
Middle	0% (0)	61% (59)	0% (0)	0% (0)	0% (0)	57% (60)	0% (0)	0% (0)
Transitive	23% (22)	0% (0)	65% (72)	0% (0)	21% (22)	1% (1)	75% (112)	0% (0)
Other	1% (1)	2% (2)	0% (0)	0% (0)	2% (2)	2% (2)	1% (2)	0% (0)
Total	24% (23)	76% (73)	76% (84)	24% (26)	24% (25)	76% (81)	85% (127)	15% (22)

Note. The percentages are calculated within normative and within non-normative protests for each age group. The numbers in parentheses show the raw frequencies.

Discussion

The results suggested that 3-year-olds were less selective for the contexts in which they enforced norms than 5-year-olds did. That is, 3-year-olds had normative conflicts also in the compatible context. Second, when facing conflict about the rules of the game, 3-year-olds insisted on their rule longer and were more reluctant to consider the other rule as a plausible alternative, as compared to 5-year-olds. Finally, in their normative protests, 3-year-olds used more limited discursive strategies, almost exclusively inanimate middle constructions, objectifying the norms. We will discuss these findings in detail in General Discussion.

To have a more complete developmental story, we conducted a second study in which we analyzed the negotiations of 5- and 7-year-old children because in middle childhood children show some further flexibility in their reasoning on moral norms and social conventions (Conry-Murray & Turiel, 2012; Horn et al., 1999; Killen & Stangor, 2001; Themier, Killen & Stangor, 2001). In Study 2, we presented 5- and 7-year-olds with the same sorting games. Although task difficulty was not an issue in Study 1, to make the tasks more interesting for 5- and 7-year-olds, each game had a more complicated version, in which there were four levels to each of the two sorting dimensions.

Study 2

Method

Participants

Forty-eight 5-year-olds (*M* = 5;8, range = 5;5–5;11, 24 boys and 24 girls), who did not participate in Study 1, and forty-eight 7-year-olds (*M* = 7;2, range = 6;10–7;11, 24 boys and 24 girls) participated in Study 2. The 96 children were grouped into 48 dyads (24 dyads of 5-year-olds and 24 dyads of 7-year-olds) matched by gender. Half of the dyads participated in the compatible context, and the other half in the incompatible context. The dyads were formed on the basis of the recommendations

of their teacher and were composed of children who were friends. All children were monolingual German speakers, except for the two 7-year-olds who were bilinguals (the second language of one child was Russian, and the second language for the other child was not specified by his parent).

Materials

Each game had a more complicated version.

Board game. There were 4 animal shapes (a hedgehog, a bird, a bunny, and a frog), 4 colors (yellow, green, orange, and brown), and, thus, 16 different types of magnets combining each animal shape with each color. Children were asked to place these tokens on a 56.5 × 34 cm magnetic board on which there were four rows. On the head of each row, there was a picture of a yellow hedgehog, a green bird, an orange bunny, or a brown frog, signaling one level of each of the two sorting dimensions.

Tube game. There were 4 wooden shapes (a ball, a cube, a triangle, and a star), 4 colors (blue, red, white, and pink), and, thus, 16 different types of wooden tokens combining each shape with each color. Children were asked to place these tokens into four 6 × 29.7 cm plexiglass tubes on a platform. In front of each tube there was a red cube, a blue ball, a white triangle, or a pink star, signaling one level of each of the two sorting dimensions.

In both games, children sorted 12 tokens in the teaching phase and 16 tokens in the testing phase. The order of 12 tokens in the teaching phase and 16 tokens in the testing phase were the same across dyads.

Procedure

The procedure, the counterbalancing, and the coding were exactly the same as Study 1. The kappa values reported in Study 1 for the interrater reliability on the coding apply to Study 2 (half of the reliability files were from Study 2). The instructions were slightly different. In the teaching phases of both games, the experimenter told each child, for instance, for the color rule in the board game, "All the yellows must go in this row; all the greens must go in this row; all the browns must go in this row; and all the oranges must go in this row" and delivered the instructions the same way for the shape rule in both games. The use of the modal verb *müssen* "must" in the instructions could have been interpreted as encouraging normative language. However, the frequency of the modal *müssen* "must" within on-task utterances did not change between Study 1 and Study 2. The models (fit with poisson error structure) included the control and random predictors, as well as age as a covariate. Adding the study type (1 vs. 2) to the model did not improve the fit, $\chi^2 = 1.64$, $df = 1$, $p = .19$.

Results

The same four analyses were carried out: (a) presence of normative conflict in each session, (b) number of utterances it took children to reach agreement, (c) marking of agency within normative protests, and (d) the co-occurrence patterns of agency, normativity, and transitivity.

Presence of Normative Conflict

We ran a GLMM model in which the response variable was a binary measure, session with a normative conflict versus session with a non-normative conflict (or no conflict at all). The full model included all the predictors of interest: age group (5 vs. 7), context (compatible vs. incompatible), and their interaction, as well as the control predictors: the game (tube game vs. board game), the order of these games (1 vs. 2), and their interaction. The random factor of dyad ($N = 48$) was also added to the model. The null model comprised only the control predictors (game, task order, and their interaction) and random factor of dyad.

The full model improved the fit as compared to the null model ($\chi^2 = 32.89$, $df = 3$, $p < .001$). However, the interaction between age group and context was not significant ($B = 16.21$, $SE = 3,932.43$, $z = 0.004$, $p = .99$), so this interaction was removed from the model to get interpretable tests of the main effects. The only significant main effect of context suggested that both 5- and 7-year-olds had normative conflicts more in the incompatible context than compatible context ($B = 4.34$, $SE = 1.19$, $z = 3.62$, $p < .001$; see the bottom panel in Figure 1).

Number of Utterances Until Agreement

We used a repeated measures ANOVA with one within-subjects factor (game with two levels: board and tube) and three between–subjects factors: (a) age group with two levels: 5 versus 7, (b) context with two levels: compatible versus incompatible, and (c) order of game with two levels: board game first versus tube game first. The response variable was the total number of on-task utterances (both at the clause level and nonclause level and regardless of whether they were protest utterance or not) by both children in the dyad until they reached

agreement. We log transformed the response variable. Through this log transformation, the assumptions of normal distribution and homogeneity of residuals were fulfilled by visual inspection of a qq-plot and residuals plotted against fitted values.

The repeated measures ANOVA revealed only a significant effect of context, suggesting that it took both 5- and 7-year-olds significantly longer to reach agreement in the incompatible context than in the compatible context, $F(1, 43) = 49.28$, $p < .001$, $\eta_p^2 = .53$ (see the right panel in Figure 2).

Marking of Agency Within Normative Protests

To compare the co-occurrence patterns of normative protests with animate versus inanimate subjects across the age groups and the two contexts, we ran GLMM in which the response variable was the proportion of normative protests to all protests (normative and non-normative). The full model included all the predictors of interest, that is, age group (5 vs. 7), context (compatible vs. incompatible), agency (animate vs. inanimate), and the interaction between agency and age group, as well as the control predictors of the game (tube game vs. board game), the order of these games (1 vs. 2), and their interaction. The random factors in the models were dyad and games nested in dyads. There were many sessions without any protest, especially in the compatible context so there were not enough data points in the compatible context to make comparisons of interaction terms involving the context. Thus, we did not include interaction terms involving context in the full model. The null model comprised only the control predictors (game, the game order, and their interaction) and the two random factors. In each model we weighted the response by the number of utterances per dyad.

The full model improved the fit as compared to the null model ($\chi^2 = 1,570.0$, $df = 4$, $p < .001$); however, the interaction between agency and age group was not significant ($B = 0.12$, $SE = 0.19$, $z = 0.61$, $p = .54$), so this interaction was dropped from the model to get the interpretable results of the main effects. The only significant main effect of agency suggested that in their normative protests, both 5- and 7-year-olds used inanimate subjects more than they used animate subjects ($B = 3.08$, $SE = 0.09$, $z = 33.78$, $p < .001$).

Agency and Transitivity

To see which type of constructions these animate and inanimate subjects co-occur in, all 461 protest utterances were coded for their transitivity, which had five levels: utterances with copular (constructions with *sein* "be"), transitives, middles, intransitives, and other. Five- and 7-year-olds did not differ from one another and showed the same general trend as the 5-year-olds in Study 1. Namely, they predominantly used inanimate middles (61% for 5-year-olds, 57% for 7-year-olds) and some animate transitives (23% for 5-year-olds, 21% for 7-year-olds) in their normative protests (see the bottom panel of Table 1). In their non-normative protests they used mostly transitive constructions with animate subjects and some copular constructions with inanimate subjects.

Discussion

The results of Study 2 suggested that 5-year-olds and 7-year-olds did not differ in any of the measures and the patterns were similar to the 5-year-olds in Study 1. This could be a ceiling effect due to our tasks, which did not require abstract level of reasoning. In dealing with more abstract concepts like stereotypes, children's sociomoral reasoning shows progression after preschool years (Conry-Murray & Turiel, 2012; Horn et al., 1999; Killen & Stangor, 2001; Theimer, Killen, & Stangor, 2001). For instance, in evaluating two candidates, "a boy and a girl equally good at ballet," for a ballet club, school-aged children mostly sided with the candidate who goes against the stereotype based on the reasons like "Boys don't get a chance to take ballet" (Killen & Stangor, 2001, p. 183), as opposed to preschoolers who sided with the gender-appropriate candidate in comparable contexts (Theimer et al., 2001).

General Discussion

The findings suggested that the way 3-year-olds negotiated game rules with their peers differed from that of the older children. In line with the quantitative results, Example 1 shows a segment of a dialogue between two 3-year-olds boys, Pete and Andre, in the incompatible context in Study 1, playing the board game. Pete was instructed to sort the items by animal and Andre by color. Pete states the rule as he places the hedgehog on the animal row in line 23 (Lines marked with carets indicate the normative protests).

Example 1:

	23	Pete:	*Und der Igel kommt hier.*
			And this hedgehog comes here.
>	24	Andre:	*Nein, der kommt hierhin.*
			No, it comes here.
	25	Pete:	*Gar nich(t), kuck.*
			Not at all, look.
>	26		*Da komm(en) Vögel.*
			Birds go there.
	27		*Ja.*
			Yes.
	28	[Andre removes the magnet off the board.]	
	29	Andre:	*Aber der hat gesagt,*
			But he said that,
>	30		*Grüne komm(en) zu x hierhin.*
			Green ones go here.
	31	Pete:	*Nein, kuck.*
			No, look.
>	32		*Grün muss zu Grün.*
			Green has to (go) with green.
>	33		*Das muss so sein.*
			It has to be like this.
	34	Andre:	*Wah.*
			Wah.
>	35	Pete:	*Vogel kommt nur hierhin.*
			Bird goes only here.
>	36	Andre:	*Wä, nein, das kommt zu …*
			Wa, no, it goes to …
	37		*Nein, kuck hier.*
			No, look, here.
>	38		*Das kommt zu hierhin.*
			It goes here.
	39	[Andre removes the magnet off the board.]	
	40	Pete:	*Nein.*
			No.
	41	Andre:	*Doch, kuck.*
			Yes, look.
>	42	Pete:	*Gar nich(t), Vogel kommt hierhin.*
			Not at all, bird goes here.
	43		*Siehst du?*
			You see?
>	44		*Igel kommt hierhin, Andre.*
			Hedgehog goes here, Andre.

As Example 1 shows, these two 3-year-olds exclusively used normative protests with inanimate subjects, mostly middles, such as *Nein, der kommt hierhin* "No it comes here" (line 24), *Da komm(en) Vögel* "Birds go there" (line 26), *Grün muss zu Grün* "Green has to (go) with green" (line 32). They both insist on their version of rule over and over and do not acknowledge that the alternative rule is acceptable. In fact, they continue to argue over the rule for 66 more lines and do not eventually agree on a rule. This pattern was also evident in the quantitative analyses. As compared to older children, it took 3-year-olds longer to reach agreement and 80% of their protest utterances consisted of the same type of inanimate middle constructions, which were repeatedly used in their negotiations.

In contrast to Example 1, the following example shows how 5-year-old boys arrive at a resolution relatively more quickly. In Example 2, two 5-year-olds, Roger and Rafael are playing the tube game (Study 1). Rafael was instructed to sort the items by shape and Roger by color. Rafael places the first token in the tube according to its shape.

Example 2:

	1	Roger:	*Rafael, du musst die blauen rein.*
			Rafael, you have to put the blue ones in.
		[14 lines with non-normative protests skipped.]	
>	16	Rafael:	*Aber wir machen (e)s so.*
			But we'll do it like that.
	17		*Wie ich …*
			Like I …
	18		*Hey Roger.*
			Hey Roger.
>	19		*Du hast Recht.*
			You're right.
>	20		*X zuerst machen wir (e)s so.*
			X we do it like this first.
	21		*Wie du es willst.*
			Like you want.

In line 1, Roger objects to Rafael's placement of the token with a normative protest with an animate subject, *Rafael, du musst die blauen rein* "Rafael, you have to put the blue ones in." After this point, they continue to dispute over the rule, with non-normative protests such as *Aber wir haben es anders gespielt* "But we played it differently" for 14 lines. In line 16, Rafael insists on sorting the items by shape and uses a formulaic normative expression *Aber wir machen (e)s so* "But we do it like that." Then in line 19, Rafael implies that Roger's way can be correct as well (*Du hast recht* "You are right"). In lines 20–21, he agrees to sort the items by color and offers a solution in which they would sort Roger's way first (*Zuerst machen wir es so, wie du es willst* "First we do it like that like you want"), and his way next. Thus, 5-year-olds were quicker to adapt to alternative set of rules.

The findings reported in this study are, to our knowledge, the first systematic analyses of how children enforce norms in their spontaneous peer conflicts and how this changes developmentally. Through controlling the context (sorting game), the kind of conventional norms (rules for sorting the items), and the source of the norms (the same adult experimenter), we have documented some

developmental changes, after age 3, in terms of how selective children are with their norm enforcements in different contexts (3-year-olds are not as selective), how long it takes them to resolve a peer dispute (it takes 3-year-olds longer), and what kind of linguistic constructions they use to do so (all age groups use "objective" normative language).

In their normative interventions, children reminded their peers of the "correct" rule of the game. In this study we looked at the particular linguistic constructions used by children of different ages. All age groups, especially 3-year-olds, preferred inanimate subjects in their normative protests. The use of inanimate subjects in expressions like *This goes here*, *This is not right* has a more objective flavor and the speaker emphasizes that it is not something he wants, thinks, or believes, but it is rather a fact. The use of animate subjects *You must put this here*, *You did that wrong* holds the speaker accountable for that assessment and it does not completely free the utterance from speakers' personal preferences and idiosyncratic beliefs about the rules. Thus, 3-year-olds almost exclusively took the objective and factual view of the event and portrayed the rules as "unalterable facts." This was also reflected in the relatively greater number of utterances it took them to reach agreement, their longer persistence on one rule, and their reluctance to see the alternative rule as possible as in Example 1 above.

One explanation of why all age groups used inanimate subjects in their normative protests could be due to the instructions delivered by the adult experimenter, who used inanimate subjects (*Alle grünen kommen hier hin* "All the green ones go here"). Children could be repeating how the adult had presented the rules. However, despite receiving the same instructions, 3-year-olds still appealed to normative expressions with inanimate subjects more than older children did.

Another finding was that 3-year-olds did not distinguish the two contexts as well as the older children and enforced norms in the compatible context where there were technically no norm violations. In the compatible context, the children were presented with both rules in each game and, due to our counterbalancing, the last rule each child learned was different. The appropriateness of an action was contingent upon the joint decision about by which rule the children would play the game. As there were two possible rules, children had to take the perspective of their peer and monitor to which rule their partner was orienting. The limited perspective-taking skills of 3-year-olds might explain why they enforced norms in the compatible context (Perner,

Brandl, & Garnham, 2003; Perner, Stummer, Sprung, & Doherty, 2002). That is, when one child placed the first token without verbally marking which rule he or she had in mind, the second child had to inhibit the rule he or she had in mind, and adjust to his or her partner's rule. In these cases, some 3-year-olds did not adjust to the rules of their peers, but put forward their own rule and treated the other's actions as norm violations. The fact that they could not inhibit the rule they had in mind could also result from the limited executive functioning abilities of 3-year-olds (Carlson & Moses, 2001; Frye, Zelazo, & Palfai, 1995). Older children were better at adjusting to the rule of their peer. In fact, older children often disambiguated the rules from the beginning by asking their peer to which set of rules to commit, such as *Machen wir Farbenspiel oder Tierespiel?* "Do we play the color game or the animal game?" or by explicitly signaling to which set of rules they are orienting, such as *Zuerst das Formenspiel* "First the shape game."

Overall, the findings are informative about children's understanding of conventional norms regarding game rules. The results corroborate other studies on the fact that by 3 years of age, children are aware of such norms and responded to violations of them by referring to the rule (Nucci & Nucci, 1982a, 1982b; Nucci & Turiel, 1978; Rakoczy, 2008; Rakoczy et al., 2009; Smetana, 1981; Turiel, 1983; Wyman et al., 2009). However, this initial normative understanding seems to be less discriminate and less flexible. At age 3, children seem to be pretty dogmatic about game rules and view these as somewhat less alterable. Throughout the preschool years, normative understanding seems to get more fine tuned (Kalish, 2005; Rakoczy, 2008; Smetana, 2006; Turiel, 2002). Children become more flexible in evaluating the right way and the wrong way to play the game, and are more able to divorce the current interactive context (a peer wanting to play the tube game and sort the items by shape) from the broader general context (the rule of the tube game is to sort the items by color). They are quicker to adjust to possible alternative rules, even if this contradicts their own knowledge. So what complements children's normative understanding during these years seems to be the appreciation of the arbitrariness and the negotiability of conventional norms within rule-governed games, as envisaged by Piaget (1932/1965).

In summary, when facing a conflict or ambiguity about game rules, children bring normative order to their peer interactions. How flexibly they arrive at a resolution and the linguistic strategies they

deploy in their negotiations are informative about their normative understanding. Our results suggested that although they know much about norms even in the preschool years, young children's actual use of social norms to structure their interactions with others, especially peers, continues to develop and becomes more flexible into the school years.

References

Adams, D. C., & Anthony, C. D. (1996). Using randomization techniques to analyse behavioural data. *Animal Behaviour*, *51*, 733–738. doi:10.1006/anbe.1996.0077

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Bates, D., Maechler, M., & Bolker, B. (2012). *lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0.* Retrieved from http://CRAN.R-project.org/package=lme4

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Erlbaum.

Budwig, N. (1995). *A developmental-functionalist approach to child language*. Mahwah, NJ: Erlbaum.

Budwig, N., Stein, S., & O'Brien, C. (2001). Non-agent subjects in early child language: A crosslinguistic comparison. In K. Nelson, A. Aksu-Koç, & C. Johnson (Eds.), *Children's language: Vol. 11. Interactional contributions to language development* (pp. 49–67). Mahwah, NJ: Erlbaum.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*, 1032–1053. doi:10.1016/j.jecp.2004.01.002

Conry-Murray, C., & Turiel, E. (2012). Jimmy's baby doll and Jenny's truck: Young children's reasoning about gender norms. *Child Development*, *83*, 146–158. doi:10.1111/j.1467-8624.2011.01696.x

Corsaro, W. A. (1994). Discussion, debate, and friendship processes: Peer discourse in U.S. and Italian nursery schools. *Sociology of Education*, *67*, 1–26. doi:10.2307/2112747

Dunn, J., & Munn, P. (1987). Development of justification in disputes with mother and sibling. *Developmental Psychology*, *23*, 791–798. doi:10.1037/0012-1649.23.6.791

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, *10*, 483–527. doi:10.1016/0885-2014(95)90024-1

Hartup, W. W., French, D. C., Laursen, B., Johnston, M. K., & Ogawa, J. R. (1993). Conflict and friendship relations in middle childhood: Behavior in a closed-field situation. *Child Development*, *64*, 445–454. doi:10.1111/j.1467-8624.1993.tb02920.x

Hartup, W. W., Laursen, B., Stewart, M. I., & Eastenson, A. (1988). Conflict and the friendship relations of young children. *Child Development*, *59*, 1590–1600. doi:10.2307/1130673

Hay, D. F. (1984). Social conflict in early childhood. In G. Whitehurst (Ed.), *Annals of child development* (Vol. 1, pp. 1–44). Greenwich, CT: Jai Press.

Hay, D. F., & Ross, H. S. (1982). The social nature of early conflict. *Child Development*, *53*, 105–113. doi:10.2307/1129642

Helwig, C. C., & Turiel, E. (2003). Children's social and moral reasoning. In P. K. Smith & C. H. Hart (Eds.), *Blackwell handbook of childhood social development* (pp. 475–490). Oxford, UK: Blackwell

Horn, S. S. (2003). Adolescents' reasoning about exclusion from social groups. *Developmental Psychology*, *39*, 71–84. doi:10.1037/0012-1649.39.1.71

Horn, S., Killen, M., & Stangor, C. (1999). The influence of group stereotypes on adolescents' moral reasoning. *Journal of Early Adolescence*, *19*, 98–113. doi:10.1177/0272431699019001005

Ingram, G. P. D., & Bering, J. M. (2010). Children's tattling: The reporting of everyday norm violations in preschool settings. *Child Development*, *81*, 945–957. doi:10.1111/j.1467-8624.2010.01444.x

Kalish, C. (2005). Becoming status conscious: Children's appreciation of social reality. *Philosophical Explorations*, *8*, 245–263. doi:10.1080/07418820500219359

Kalish, C., Weissman, M., & Bernstein, D. (2000). Taking decisions seriously: Young children's understanding of conventional truth. *Child Development*, *71*, 1289–1308. doi:10.1111/1467-8624.00229

Kemmer, S. (2003). Human cognition and the elaboration of events. In M. Tomasello (Ed.), *The new psychology of language* (Vol. 2, pp. 89–118). Mahwah, NJ: Erlbaum

Killen, M. (1991). Social and moral development in early childhood. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development* (pp. 115–138). Hillsdale, NJ: Erlbaum.

Killen, M., & Stangor, C. (2001). Children's social reasoning about inclusion and exclusion in gender and race peer group contexts. *Child Development*, *72*, 174–186. doi:10.1111/1467-8624.00272

Killen, M., & Turiel, E. (1991). Conflict resolution in preschool social interactions. *Early Education & Development*, *2*, 240–255. doi:10.1207/s15566935eed0203_6

Kyratzis, A. (2004). Talk and interaction among children and the co-construction of peer groups and peer culture. *Annual Review of Anthropology*, *33*, 625–649. doi:10.1146/annurev.anthro.33.070203.144008

Kyratzis, A. (2009). "He take one of my tools!" vs. "I'm building": Transitivity and the grammar of accusing, commanding, and perspective-sharing in toddlers' peer disputes. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura, & Ş. Özçalõşkan (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 41–67). New York, NY: Psychology Press.

Macwhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Erlbaum.

Manly, B. F. J. (1997). *Randomization, bootstrap and Monte Carlo methods in biology*. New York, NY: Chapman & Hall.

Much, N., & Shweder, R. A. (1978). Speaking of rules: The analysis of culture in breach. In W. Damon (Ed.), *New directions for child development: Vol. 1. Moral development* (pp. 19–40). San Francisco, CA: Jossey-Bass.

Nucci, L. P., & Nucci, M. S. (1982a). Children's responses to moral and social conventional transgressions in free-play settings. *Child Development*, 53, 1337–1342. doi:10.2307/1129024

Nucci, L. P., & Nucci, M. S. (1982b). Children's social interactions in the context of moral and conventional transgressions. *Child Development*, 53, 403–412. doi:10.2307/1128983

Nucci, L. P., & Turiel, E. (1978). Social interactions and the development of social concepts in preschool children. *Child Development*, 49, 400–407. doi:10.2307/1128704

Perner, J., Brandl, J., & Garnham, A. (2003). What is a perspective problem? Developmental issues in understanding belief and dual identity. *Facta Philosophica, 5*, 355–378.

Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). Theory of mind finds its Piagetian perspective: Why alternative naming comes with understanding belief. *Cognitive Development*, 17, 1451–1472. doi:10.1016/S0885-2014(02)00127-2

Piaget, J. (1965). *The moral judgment of the child*. London, UK: Free Press. (Original work published 1932)

R Development Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0. Retrieved from http://www.R-project.org

Rakoczy, H. (2008). Taking fiction seriously: Young children understand the normative structure of joint pretence games. *Developmental Psychology*, 44, 1195–1201. doi:10.1037/0012-1649.44.4.1195

Rakoczy, H., Brosche, N., Warneken, F., & Tomasello, M. (2009). Young children's understanding of the context-relativity of normative rules in conventional games. *British Journal of Developmental Psychology*, 27, 445–456. doi:10.1348/026151008X337752

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. doi:10.1037/0012-1649.44.3.875

Schieffelin, B. B., & Ochs, E. (1986). Language socialization. *Annual Review of Anthropology*, 15, 163–191. doi:10.1146/annurev.an.15.100186.001115

Shantz, C. U. (1987). Conflicts between children. *Child Development*, 58, 283–305. doi:10.2307/1130507

Smetana, J. G. (1981). Preschool children's conceptions of moral and social rules. *Child Development*, 52, 1333–1336. doi:10.2307/1129527

Smetana, J. G. (1984). Toddlers' social interactions regarding moral and conventional transgressions. *Child Development*, 55, 1767–1776. doi:10.2307/1129924

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

Smetana, J., & Braeges, J. L. (1990). The development of toddlers' moral and conventional judgments. *Merrill Palmer Quarterly*, 36, 329–346.

Smetana, J. G., Schlagman, N., & Adams, P. W. (1993). Preschool children's judgments about hypothetical and actual transgressions. *Child Development*, 64, 202–214. doi:10.1111/j.1467-8624.1993.tb02904.x

Theimer, C. E., Killen, M., & Stangor, C. (2001). Young children's evaluations of exclusion in gender-stereotypic peer contexts. *Developmental Psychology*, 37, 18–27. doi:10.1037/0012-1649.37.1.18

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

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

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, UK: Cambridge University Press.

Turiel, E. (1998). Moral development. In: N. Eisenberg & W. Damon (Eds.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (pp. 863–932). New York, NY: Wiley.

Turiel, E. (2002). *The culture of morality: Social development, context, and conflict*. New York, NY: Cambridge University Press.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72, 655–684. doi:10.1111/1467-8624.00304

Wyman, E., Rakoczy, H., & Tomasello, M. (2009). Normativity and context in young children's pretend play. *Cognitive Development*, 24, 146–155. doi:10.1016/j.cogdev.2009.01.003

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, 11, 37–63. doi:10.1016/S0885-2014(96)90027-1