



Children's reasoning with peers in cooperative and competitive contexts

Andreas Domberg^{1,2*} , Bahar Köymen² and Michael Tomasello^{1,3}

¹Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

²School of Health Sciences, University of Manchester, UK

³Department of Psychology and Neuroscience, Duke University, Durham, North Carolina, USA

We report two studies that demonstrate how five- and seven-year-olds adapt their production of arguments to either a cooperative or a competitive context. Two games elicited agreements from peer dyads about placing animals on either of two halves of a playing field owned by either child. Children had to produce arguments to justify these decisions. Played in a competitive context that encouraged placing animals on one's own half, children's arguments showed a bias that was the result of withholding known arguments. In a cooperative context, children produced not only more arguments, but also more 'two-sided' arguments. Also, seven-year-olds demonstrated a more frequent and strategic use of arguments that specifically refuted decisions that would favour their peers. The results suggest that cooperative contexts provide a more motivating context for children to produce arguments.

Statement of contribution

What is already known on this subject?

- Reasoning is a social skill that allows people to reach joint decisions.
- Preschoolers give reasons for their proposals in their peer conversations.
- By adolescence, children use sophisticated arguments (e.g., refutations and rebuttals).

What the present study adds?

- Cooperation offers a more motivating context for children's argument production.
- Seven-year-olds are more strategic than five-year-olds in their reasoning with peers.
- Children's reasoning with others becomes more sophisticated after preschool years.

Reasoning is a core aspect of decision-making. In their Argumentative Theory, Mercier and Sperber (2011) characterize reasoning – in the sense of giving and evaluating reasons – as a social skill, with which speakers win others over to their point of view. The theory is motivated by the concept of Epistemic Vigilance (Sperber *et al.*, 2010), which guards hearers from being deceived and requires speakers to support their claims with mutually accessible evidence to overcome this mistrust. Tomasello (2014), on the other hand, emphasizes that reasoning is also important in cooperative contexts. Because all parties

*Correspondence should be addressed to Andreas Domberg, Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, Leipzig 04103, Germany (email: domberg@posteo.de).

share an interest in reaching the best decision, who produces the 'winning' argument becomes less important.

Children's use of arguments has been mostly studied in the context of peer conflicts (Dunn & Munn, 1987; Eisenberg & Garvey, 1981; Phinney, 1986). It has been observed that two- to five-year-old children justify opposition, and use of conciliatory justifications increases with age. Recent studies investigated how children reason cooperatively with peers, finding that in justifying their proposals, three- and five-year-olds adjusted the informativeness of justifications to their mutual knowledge with their partners to reach the correct joint decisions (Köymen, Mammen, & Tomasello, 2016; Köymen, Rosenbaum, & Tomasello, 2014). However, studies on young children's use of justifications focused either entirely on conflictual/competitive or entirely on cooperative contexts without a systematic comparison of how young children's arguments change across both contexts. Such comparison would reveal more about children's ability to pursue communicative goals argumentatively.

Research with adolescents compared speakers' argumentation across contexts, however. Felton and Kuhn (2001) had adolescents discuss controversial topics in dyads of consenting or dissenting partners. Consenters were asked to defend their view, and dissenters were asked to reach agreement. Compared to adults, adolescents showed much less strategic argumentation, for example weakening a dissenter's argument. Kuhn and Udell (2003) found that only after training, adolescents produced such powerful arguments as *refutations*, refuting potentially opposing arguments, in contrast to *affirmations*, which advance the speaker's point of view. More importantly, when adolescents were instructed to reach agreement, in contrast to winning over their interlocutor, they produced more complex arguments (e.g., refutations) and less biased argumentation, showing 'two-sided reasoning' (Garcia-Mila, Gilabert, Erduran, & Felton, 2013; see Kuhn, 2015 for a review).

Our main question here is whether the differences in argumentation observed with adolescents across cooperative and competitive contexts emerge earlier in development, in five- and seven-year-old peers' conversations. We chose these age groups because around age 5, children were observed to produce and critically evaluate arguments (Corriveau & Kurkul, 2014; Köymen *et al.*, 2014, 2016; Mercier, Bernard, & Clément, 2014). In Study 1, dyads of five- and seven-year-old peers were asked to reach agreement about placing toy animals in a zoo. Each child owned one half of the zoo. The rewarding scheme either favoured finding nice homes for every animal ('finding the best solution' – cooperative condition), or having more animals on one's own side ('winning' – competitive condition). We predicted that the cooperative condition would elicit more objective two-sided argumentation than the competitive condition. We explored any differences between conditions or age groups in the overall frequency, content, and especially type of arguments, that is, whether they were affirmations or refutations, as the literature suggests that even adolescent children tend to focus on the explanations for 'why' (justifying a desired outcome), rather than 'why not' (refuting an undesired outcome; Kuhn & Udell, 2003).

In Study 2, children played a similar game, but we made sure children knew a set of arguments to test a specific interpretation of the results from Study 1, namely, whether the bias observed in the competitive condition had cognitive or strategic reasons. One child per dyad first played the game with the experimenter and learned critical arguments. Later, he/she played with a naïve peer either in the cooperative or competitive condition. Critically, in the competitive condition, by reproducing the critical arguments, the trained child would benefit the naïve child and worsen his/her own chances of winning. We

predicted that in the competitive condition, trained children would withhold these arguments more often. As in Study 1, we explored any age differences in cooperative and competitive contexts.

STUDY I

Method

Participants

Forty-eight 5-year-olds ($M = 5;8$, range = 5;5–5;11, 24 girls) and 48 seven-year-olds ($M = 7;6$, range = 7;0–7;11, 24 girls) in 48 same-age and same-sex dyads participated in the study in their kindergartens and schools. Children in dyads knew each other (although their degree of familiarity with each other varied) and were randomly assigned to the cooperative or competitive condition in equal numbers. All children were native German speakers with various socio-economic backgrounds. Participation was subject to prior parental consent, and the procedure was approved by the departmental Ethics Committee at the Max Planck Institute for Evolutionary Anthropology.

Materials

Children placed toy animals onto two backgrounds. One was used as warm-up, comprising two pictures. The other was used in the testing phase, depicting a zoo with four cages (Figure 1). The two cages in each half were the following: (1) a 'warm cage', suitable for animals in warm climates; (2) a 'cold cage', suitable for animals in cold climates. The cold and warm cages in each half differed so that children could produce different arguments. Moreover, the orientation of the cages on each half was such that each child would see their own cages upright, but their peer's upside down.

The toy animals were released from a cardboard dispenser with 17 slots. Pulling the lowest card would release the animal at the bottom slot into a transparent box with a lid (see Figure 2 for the list of animals). Finally, a 'surprise box' held stickers as rewards.

Warm-up phase

The study took place in quiet rooms of preschools in a German city. The procedure took *c.* 15 min. The experimenter (E) seated both children facing one another with the dispenser to their side. E first introduced how the dispenser works by asking children to pull out the lowest card together. When the animal was released into the transparent box, E explained the game rules: 'First we say what animal this is, then to which cage it should go and most importantly why we put it there'. In the warm-up phase, E primed children to give reasons and exemplified three argument topics: one about habitat, two about animal safety/interaction within cages.

The first animal, introducing the habitat argument, was a fish. Almost all children said that it should go to the ocean, because it lives there. If children's justifications did not involve the habitat topic, E provided the justification: 'because fish live in water'. The second item was a group of baby chicks. Again, almost all children said they were chicks, they should go to the farm because they live in farms. If they did not, E provided the justification. The third animal was a baby fox to prime the safety argument. All children said that it was a baby fox, and it should go to the farm. E primed the safety argument by saying, 'And foxes hunt chicks, that is why it also wants to go to the farm. Now the chicks get scared, right? But, it's a baby fox. It can't hurt the chicks yet, so let's place them

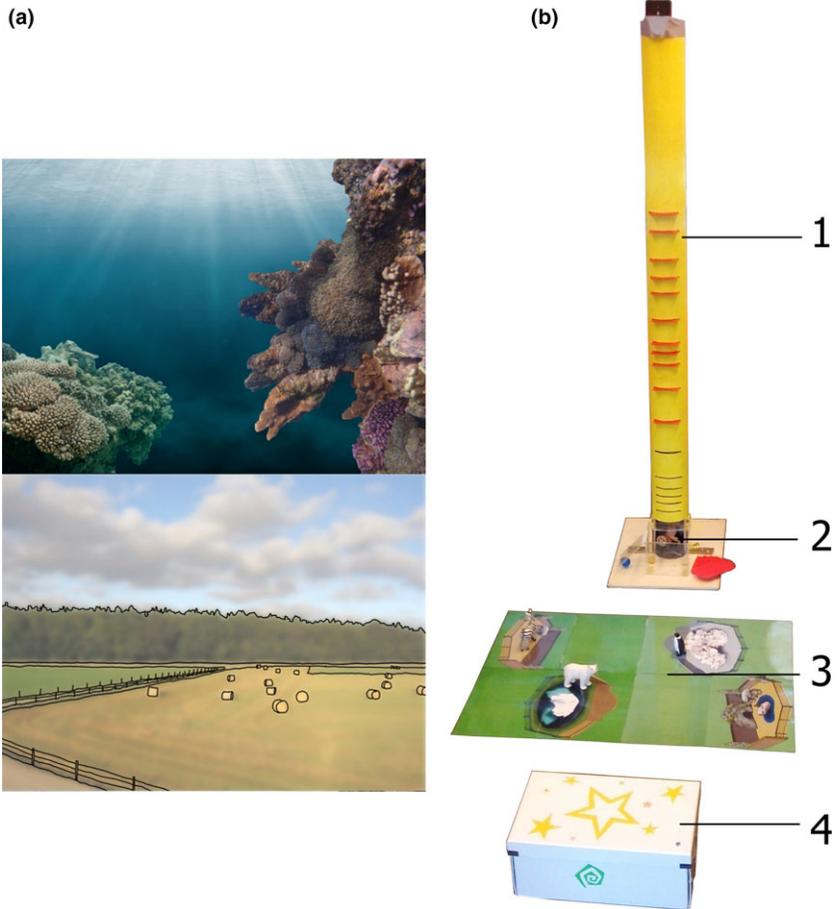


Figure 1. (a) Background images for the warm-up phase: an underwater scene (above) and a farm scene (below); (b) Experimental set-up during the testing phase: (1) dispenser with (2) plexiglass box with lid, (3) zoo background, and (4) surprise box. [Colour figure can be viewed at wileyonlinelibrary.com]

together'. The safety arguments alerted children to attend to what animals they place together in cages.

Experimental phase

After the warm-up, E introduced the zoo, placing the zoo background with four cages (see Figure 1) and said,

'Our zoo has four cages, they have sand like this [point: warm cages], or an iceberg like this [point: cold cages]. This is your half, [Child A], and this is yours, [Child B]. Now you play the game like before. First, say what animal it is; second, where it should go; and then why it goes there. All animals go into cages and not outside, because otherwise they scare the visitors'.

Children's behaviour indicated no difficulties in understanding the meaning of the materials.



Figure 2. The toy animals used in the warm-up and testing phases in the order of appearance. This order was constant for all dyads. Items 5 and 8 were pairs of usually conflicting animals and meant to elicit discussion. [Colour figure can be viewed at wileyonlinelibrary.com]

The final part of the instructions differed across conditions. In the cooperative condition, E said, 'If you find a nice home for each animal, you will both get rewards from this box [presents the gift box], but remember: they have to be nice homes. So, find a nice home for each animal, and you will both get the rewards from this box'. In the competitive condition, E said, 'If you, [Child A], have more animals on your side, only you, only [Child A], can get the reward from the box. And if you, [Child B] have more animals on your side, only you, only [Child B] can get the reward'. From an individual perspective, both children were motivated to win in both conditions. In the competitive condition, 'winning' meant 'winning individually' (i.e., *one child* gets a reward), whereas in the cooperative condition, it meant 'winning together' (i.e., *each child* gets a reward). Therefore, in case of winning (and losing), the individual pay-off was the same in both conditions.

After the instructions, E left the room, and re-entered after the fourth animal was released to remind children of the game rules, 'Remember you always have to say why each animal should go there'. After that, E did not enter until the last animal was placed, unless an animal was stuck in the dispenser, children placed an animal outside the cages, or got distracted.

Coding

Children's conversations during the experimental phase were transcribed verbatim, each line containing a clause with a single proposition. The transcript was divided into segments pertaining to the discussion of each of the 14 items.

We first identified children's game-related utterances addressed to one another. Within the game-related utterances, we identified arguments, defined as justifications for why an animal should or should not go into a cage. We excluded from the analyses those arguments that repeated or rephrased a previous argument on the same item. For each argument, we coded three properties: its content, favoured side, and type.

First, we coded arguments into five content categories:

1. *habitat*: arguments about the suitability of the cage ('It [:the polar bear] needs a cave to hide');

2. *family*: arguments construing adult and young animals of a species as kin who go together ('Because its mother is there');
3. *safety*: arguments about animal safety ('He [: the lion] would eat the zebra');
4. *species*: arguments about placing the same species of animals together ('There are already some [: penguins] there');
5. *game affordances*: arguments about the number of animals ('It's too crowded there!'). This last category was not primed in the warm-up and children spontaneously referred to this in their justifications.

Next, we coded each argument in terms of which cage it favoured: the speaker's cages ('own' arguments), the peer's cages ('other' arguments), or unclear/both (e.g., 'Let's put it on an iceberg' – which both children had).

Finally, we coded arguments in terms of their type: *affirmations* explained why an item should go to a cage (e.g., 'Because it can swim there') and *refutations* explained why it should *not* go to *another* cage (e.g., 'Otherwise the snake [: in his peer's cage] would bite it').

A second rater coded 25% of transcripts (12 dyads: three dyads per age group and condition). Agreement was $\kappa = .63$ for game-related utterances, $\kappa = .81$ for extracting arguments, $\kappa = .83$ for argument content, $\kappa = .95$ for the targeted cage, and $\kappa = .90$ for argument type.

Results

Forty-eight dyads produced a total of 7,177 utterances (94% of these utterances (6,760) were game-related). These included a total of 1,241 arguments. Within these arguments, 1,169 (94%, $M_{\text{per dyad}}: 24.4$, $SD: 18.3$) were (1) not ambiguous in terms of their goal (targeted cage), and (2) not repetitions.

There were four sets of analyses. In the first three, the unit of analysis was the dyad. We did not distinguish between child speakers within dyads, because observations were not independent. First, we analysed whether the frequency of arguments among game-related utterances differed across age groups and conditions with a two-way ANOVA with age (five- vs. seven-year-olds) and condition (cooperative vs. competitive) as between-subject factors, and gender as control factor. The response variable was the proportion of arguments to total game-related utterances. This analysis revealed two significant main effects of age and condition (Figure 3). Seven-year-olds produced more arguments than five-year-olds, $F(1, 43) = 13.71$, $p < .001$, $\eta_p^2 = .24$, and both age groups produced more arguments in the cooperative condition than in the competitive condition, $F(1, 43) = 6.31$, $p = .016$, $\eta_p^2 = .13$. There were no gender differences.

Second, we ran a two-way ANOVA with the same factors and the response variable of number of topics that each dyad based their arguments on, averaged across items. Two five-year-old dyads in the competitive condition were excluded from this analysis, because they did not produce any arguments. This analysis revealed a trend towards more diverse arguments in the cooperative condition than in the competitive condition, $F(1, 41) = 2.87$, $p = .10$, $\eta_p^2 = .13$, and significantly more diverse arguments in seven-year-olds' conversations than in five-year-olds', $F(1, 41) = 13.71$, $p = .001$, $\eta_p^2 = .24$.

Third, we investigated how often children favoured their own cages. We ran a two-way ANOVA with the same factors and the response variable of proportion of 'own' arguments to the total number of arguments produced per dyad. Again, the two 5-year-old dyads that did not produce any arguments were excluded. There was only a significant main effect of

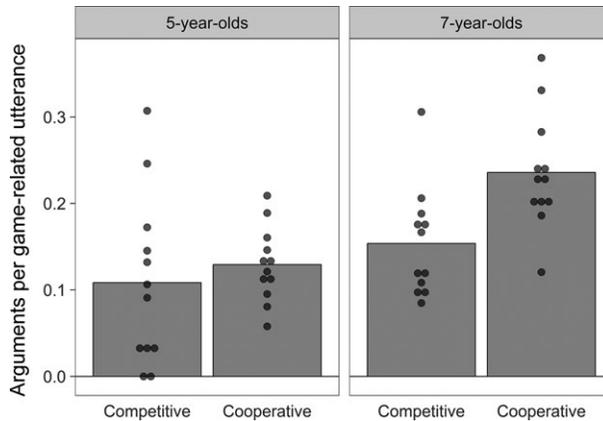


Figure 3. The proportion of arguments in overall utterances that were related to the game. Bars represent mean values per group, and dots show the outcomes for each dyad. Means (SD) for five-year-olds: competitive: .11 (.10); cooperative: .13 (.04). For seven-year-olds: competitive: .15 (.06); cooperative: .24 (.07).

condition: Both age groups favoured their own cages significantly more often in the competitive condition than in the cooperative condition, $F(1, 41) = 12.46$, $p = .001$, $\eta_p^2 = .23$. There were no age or gender differences.

Finally, we investigated to what extent children used the two types of arguments (affirmations and refutations) in their ‘own’ versus ‘other’ arguments, across age groups and conditions, using a generalized linear mixed model (GLMM) with binomial error distribution. The unit of analysis was individual arguments, and the response variable was the binary measure of whether an argument favoured the speaker’s cages (‘own’ arguments) or their peer’s (‘other’ arguments). The full model included the predictors: age group (five- vs. seven-year-olds), condition (cooperative vs. competitive), argument type (affirmation vs. refutation), and their three-way interaction, the control factor of gender, and the random factor of dyad (as we had repeated observations per dyad). To test the significance of the full model, we compared its fit to that of a null model, which only included gender and the random factor of dyad. The full model improved the fit ($\chi^2 = 24.74$, $df = 7$, $p < .001$). Compared to a reduced model without the three-way interaction, the full model still improved the fit, suggesting a significant three-way interaction ($\chi^2 = 5.30$, $df = 1$, $p = .021$; Figure 4).

To better understand the three-way interaction, we ran *post hoc* analyses on the subsets that correspond to combinations of age groups and argument types:

1. Using affirmations, five-year-olds favoured their own side significantly more in the competitive than in the cooperative condition ($\chi^2 = 7.12$, $df = 1$, $p = .008$).
2. Using refutations, five-year-olds showed no condition difference as to which side they favoured ($\chi^2 = 1.04$, $df = 1$, $p = .307$).
3. Using affirmations, seven-year-olds favoured their own side significantly more in the competitive than in the cooperative condition ($\chi^2 = 4.83$, $df = 1$, $p = .028$).
4. Using refutations, seven-year-olds favoured their own side significantly more in the competitive than in the cooperative condition ($\chi^2 = 16.13$, $df = 1$, $p < .001$).

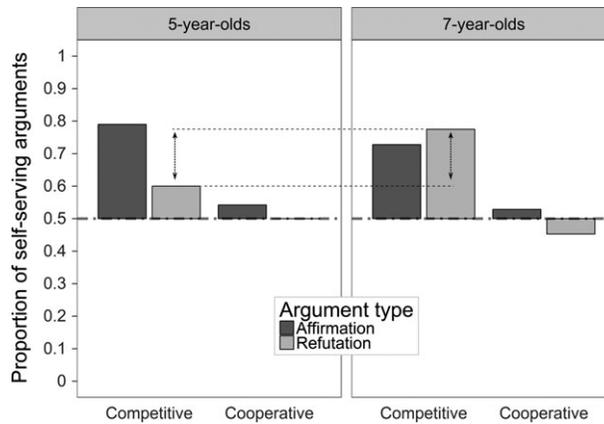


Figure 4. The proportion of 'own' arguments in overall arguments, distinguished by argument type. The difference marked in the proportions of refutations in the competitive condition refers to seven-year-olds' better grasp of this argument type as compared to five-year-olds.

Discussion

Our results suggest that children in both age groups produced systematically different samples of arguments across the two contexts. In the cooperative condition, they produced not only more arguments, but also, the set of arguments that they produced was balanced in the sense of favouring the speaker's and the peer's side equally often. In the competitive condition, children favoured their own side, but did so differently across ages. Five-year-olds mostly relied on affirmations, explaining to their peer why a toy animal should go to their own side. In contrast, seven-year-olds relied not only on affirmations but also on refutations, explaining to their peer why a toy animal should not go to their peer's side. Moreover, seven-year-olds overall produced more arguments and covered more topics in their arguments than did five-year-olds.

Since in the competitive condition, children's arguments were fewer and favoured their own side, one possible explanation is that even though they noticed the evidence supporting arguments for their peer's side, they ignored it. An alternative explanation is that they restricted the search for evidence to their own side and paid no attention to the evidence on their peer's side. Considering that seven-year-olds in both conditions produced about an equal number of refutations in favour of their own side – arguments that are based on evidence on their peer's side – they seemed to have considered the evidence on their peer's side, but possibly abstained from producing the arguments that would use it. Five-year-olds, in contrast, produced fewer refutations and showed no condition difference with respect to refutations favouring either side. In their case, it is less clear whether they were aware of the evidence on their peer's side.

Therefore, in Study 2, we controlled for children's knowledge of specific arguments. One child, the trained child, played the game with the experimenter before playing with his/her naïve peer and learned a set of critical arguments that were unusual and easy to remember. These arguments were affirmations, more easily used by five-year-olds. In the testing phase, these critical arguments would support placing animals on the naïve peer's side, benefitting the trained child in the cooperative condition, but worsening his or her

chances of winning in the competitive condition. Moreover, forming any arguments apart from the trained ones was made difficult for the trained children, since their cage offered only unusual objects (e.g., piano). We predicted that seven-year-olds would reproduce the critical arguments in the cooperative condition, but withhold them more often in the competitive condition. The reason is that their behaviour in Study 1, especially their production of self-serving refutations, suggested that they had considered both sides of the zoo in search for arguments, but seemed to withhold the ‘other’-serving arguments in the competitive context. We explored whether five-year-olds distinguish between conditions when having learned explicitly about their peer’s side encouraged them to attend to it.

STUDY 2

Method

Participants

Seventy-two five-year-olds ($M = 5;5$, range = 5;0–5;11, 36 girls) and 72 seven-year-olds ($M = 7;5$, range = 7;0–7;11, 36 girls) in 72 same-age and same-sex dyads, who did not participate in Study 1, participated in the study. Children in dyads knew each other (although their degree of familiarity with each other varied) and were randomly assigned to the cooperative or competitive condition in equal numbers. All children were native German speakers with various socio-economic backgrounds. This study fulfilled the same ethics and participation standards as Study 1.

Materials

The same dispenser, warm-up pictures, and surprise box were used as in Study 1. The animals in the warm-up were the fish and baby chicks. In the training and testing phases, the playing ground was a picture of a zoo (Figure 5). Each cage contained three objects, each of which had a circle by it, indicating where an animal should be placed. The ‘rich’ cage contained an alarm clock, a brush, and a handkerchief. The other, ‘poor’ cage contained a toy lamp, a piano, and a refrigerator.



Figure 5. The set-up and materials of Study 2. The main difference: only one cage on each side with three items in it, and a target circle (yellow) that embodies the rule ‘one animal per goal’. The animals above were used in warm-up, those below in testing, the two rightmost in each row being fillers. [Colour figure can be viewed at wileyonlinelibrary.com]

There were five animals to place (Figure 5): three critical ones, and two fillers. In the training phase, the three critical animals were a baby bear, a baby lion, and a baby elephant; the filler animals were a zebra and a penguin. In the testing phase, the three critical animals were a bear, a lion, and an elephant; the filler animals were a chimp and a snake. We used baby and adult versions of the critical animals and different filler animals across the training and testing phases because our pilot data suggested that when the exact same animals were used in both phases, children treated the trained arguments normatively and did not act strategically.

Warm-up

The warm-up phase was identical to that of Study 1, but comprised only two items, a fish and a pair of baby chicks. The baby fox was not included because we did not want to prime children about the relations of animals placed in each cage. Another difference to the Study 1 warm-up was that when asking children for an argument, E1 accepted (or suggested if necessary) any arguments about what the animal could do at each location. For example, E1 said, 'If I say 'the fish goes into the water because it can breathe there', is that a good reason?' This was to prime children to critically evaluate the arguments.

Training phase

For the training phase, one child, the naïve child (N), left the room, playing a game outside with another experimenter. E1 sat by the poor cage and seated the trained child (T) next to the rich cage, saying:

This is your cage, and this is mine. We will play the game just like before. We say first the name of the animal, where it should go, and why the animal should go there. That's the most important part. The animals don't go outside the cages. Why? Because they would run away. The one who has more animals in their own cage at the end wins.

The order of the five animals was pseudo-randomized, without the three critical animals appearing in a row. The order was the same in the training and testing phases.

For each animal, E1 first asked T in which cage one might put the animal and then announced having a great idea, pointing out: (1) a property of the animal; (2) an item in a cage; (3) what the animal can do with that item. For the three critical animals, these argument components were as follows:

- Bear: (1) bears hibernate; (2) the alarm clock; (3) it will know when winter is over.
- Lion: (1) lion has fuzzy fur; (2) the brush; (3) it can brush its fur.
- Elephant: (1) elephants have long trunks; (2) the handkerchief; (3) it can sneeze its long trunk.

For the two filler items, unless T gave other arguments, E1 proposed the locations in his own (poor) cage: the piano for the zebra because of its black-and-white fur matching the color of the piano keys, and the refrigerator for the penguin because of its preference for cold climate.

This training phase was the same and competitive in both conditions. After the last item, E1 asserted T's victory and asked T to leave the room and join the others. E1 refilled the dispenser with the five animals for the testing phase, that is adult versions of the

critical animals and new fillers. Both children were called back in. Importantly, T was asked to sit by the poor cage and N by the rich cage.

Testing phase

E1 told both children that the game worked just as before, and said:

T already knows the zoo, this is our zoo. This is N's cage, and this is T's cage now. Let's play the game just like before. After we see the animal, we say: its name, where it should go, and most importantly, why it should go there.

Then, E1 gave the same final instructions according to the condition, as in Study 1 (see above). E1 left the room and entered only if the dispenser failed or when the children were done.

Coding

Conversations were transcribed verbatim. We coded that a critical argument was reproduced if T (1) mentioned the animal property (e.g., 'It has this fuzzy fur'), or (2) mentioned the activity (e.g., 'It hibernates'), or (3) acted out the activity, (e.g., brushing the animal's fur, making a sneezing sound). If T produced the critical argument for an incorrect animal, this was still coded as reproducing the argument because the animal would end up on N's side. Of 72 dyads with three trials each, eleven naïve children came up with one trained argument spontaneously, one of them with two. A second rater coded the transcripts of 16 dyads for the reproduction of the critical arguments. The agreement was $\kappa = .72$ (87.5%).

Results

We ran a two-way ANOVA with age and condition as between-subject factors and gender as control factor. The response variable was the number of trained arguments that were reproduced by T (0–3). The analysis revealed two significant main effects of age group and condition. Seven-year-olds reproduced the critical arguments more often than five-year-olds, $F(1, 67) = 11.73, p = .001, \eta_p^2 = .15$. Children reproduced the critical arguments more often in the cooperative condition than in the competitive condition, $F(1, 67) = 4.44, p = .039, \eta_p^2 = .06$. However, looking at each age group separately, the condition effect seems to be driven by seven-year-olds (Figure 6): Seven-year-olds produced significantly more trained arguments in the cooperative condition than in the competitive condition ($F(1, 33) = 7.41, p = .010, \eta_p^2 = .18$), but five-year-olds showed no condition difference ($F(1, 33) = .25, p = .619$).

Discussion

The results of Study 2 suggest that children produced less of the trained arguments, or more often withheld them, in the competitive condition than in the cooperative condition. However, this holds rather for seven-year-olds than for five-year-olds. A possible explanation for the lack of a significant interaction between condition and age group is that most seven-year-olds in the competitive condition withheld only one trained

the competitive condition, seven-year-olds withheld these arguments. Although five-year-olds produced more self-serving arguments in Study 1, this was based entirely on their use of affirmations. When producing affirmations was harder, as in Study 2, because no trained arguments would help them justify putting animals next to unusual items, they were not as strategic as seven-year-olds. Therefore, compared to seven-year-olds, five-year-olds relied more on simpler arguments, namely affirmations, which justified a desired proposal/outcome (why an item should be placed somewhere). Seven-year-olds, in contrast, additionally relied on refutations, which are cognitively more complex than affirmations because refutations disprove a possible undesired proposal. This shows that seven-year-olds were paying attention to not only what evidence benefits their goals (affirmations) but also what evidence might conflict with their peers' goals (refutations).

Children's use of what we termed refutations shows an interesting contrast with what Felton and Kuhn (2001) and Kuhn and Udell (2003) observed with school-age children. In these studies, children did not frequently make use of advanced argumentative strategies such as counterargument and rebuttal, which anticipate and invalidate argumentative moves of their peer. Children rather used such strategies to a greater extent only after explicit training/instructions to produce refutations (Kuhn & Udell, 2003). In Study 1, children of both age groups, five-year-olds to a lesser extent, used refutations. They explained to their peers why an animal should not go to their peer's side, thereby weakening their peer's position. Therefore, our studies demonstrate how young children are able to use complex arguments such as refutations at an earlier age than previously shown in the literature.

One limitation of our study is whether the decisions were equally 'joint' across conditions. In the strongest sense, joint decisions encompass those cases in which both participants explicitly acknowledge the decision. In the weakest sense, they also encompass cases in which an animal was placed somewhere, without attempts to undo this. In our data, children often reached joint decisions after deliberation but without verbally marking their consent to the decision in both conditions. Thus, we opted to use the weak definition because limiting the joint decisions to verbal acknowledgements would not represent children's interactions accurately.

To conclude, our results suggest that by age 7, children are able to produce sophisticated arguments such as refutations, and adapt them strategically to different discursive goals in their peer interactions. When the context is intuitive and gives children more opportunities to produce arguments, as in Study 1, even five-year-olds can produce refutations. As both studies point out, cooperative situations, in which children have joint goals, provide a more motivating context for them to produce arguments and to discuss more aspects of the question at hand.

References

- Corriveau, K. H., & Kurkul, K. E. (2014). "Why does rain fall?": Children prefer to learn from an informant who uses noncircular explanations. *Child Development*, *85*, 1827–1835. <https://doi.org/10.1111/cdev.12240>
- Dunn, J., & Munn, P. (1987). Development of justification in disputes with mother and sibling. *Developmental Psychology*, *23*, 791–798. <https://doi.org/10.1037/0012-1649.23.6.791>
- Eisenberg, A., & Garvey, C. (1981). Children's use of verbal strategies in resolving conflicts. *Discourse Processes*, *4*, 149–170. <https://doi.org/10.1080/01638538109544512>
- Felton, M., & Kuhn, D. (2001). The development of argumentative discourse skill. *Discourse Processes*, *32*, 135–153. <https://doi.org/10.1080/0163853X.2001.9651595>

- Garcia-Mila, M., Gilabert, S., Erduran, S., & Felton, M. (2013). The effect of argumentative task goal on the quality of argumentative discourse. *Science Education, 97*, 497–523. <https://doi.org/10.1002/sce.21057>
- Köymen, B., Mammen, M., & Tomasello, M. (2016). Preschoolers use common ground in their justificatory reasoning with peers. *Developmental Psychology, 52*, 423–429. <https://doi.org/10.1037/dev0000089>
- Köymen, B., Rosenbaum, L., & Tomasello, M. (2014). Reasoning during joint decision-making by preschool peers. *Cognitive Development, 32*, 74–85. <https://doi.org/10.1016/j.cogdev.2014.09.001>
- Kuhn, D. (2015). Thinking together and alone. *Educational Researcher, 44*, 46–53. <https://doi.org/10.3102/0013189X15569530>
- Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development, 74*, 1245–1260. <https://doi.org/10.1111/1467-8624.00605>
- Mercier, H., Bernard, S., & Clément, F. (2014). Early sensitivity to arguments: How preschoolers weight circular arguments. *Journal of Experimental Child Psychology, 125*, 102–109. <https://doi.org/10.1016/j.jecp.2013.11.011>
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences, 34*, 57–74. <https://doi.org/10.1017/S0140525X10000968>
- Phinney, J. S. (1986). The structure of 5-year-olds' verbal quarrels with peers and siblings. *The Journal of Genetic Psychology, 147*, 47–60. <https://doi.org/10.1080/00221325.1986.9914479>
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language, 25*, 359–393. <https://doi.org/10.1111/j.1468-0017.2010.01394.x>
- Tomasello, M. (2014). *A natural history of human thinking*. Cambridge, MA: Harvard University Press.

Received 13 February 2017; revised version received 30 August 2017