Preschoolers refer to direct and indirect evidence in their collaborative reasoning

Bahar Köymen a,*, Solveig Jurkat b, Michael Tomasello c,d

a School of Health Sciences, University of Manchester, Manchester M13 9PL, UK
b Department of Psychology, University of Münster, 48149 Münster, Germany
c Department of Psychology & Neuroscience, Duke University, Durham, NC 27708, USA
d Department of Developmental and Comparative Psychology, Max Planck Institute of Evolutionary Anthropology, D-04103 Leipzig, Germany

Article history:
Received 7 October 2019
Revised 9 December 2019
Available online 31 January 2020

Keywords:
Reasoning
Reasoning by exclusion
Nonverbal reason-giving
Collaborative decision-making
Direct/indirect evidence
Peer interactions

Abstract
Collaborative reasoning requires partners to evaluate options and the evidence for or against each option. We investigated whether preschoolers can explain why one option is best (direct reasons) and why the other option is not (indirect reasons), looking at both problems that have a correct answer and those that require choosing the better option. In Study 1, both age groups produced direct reasons equally frequently in both problems. However, 5-year-olds produced indirect reasons more often than 3-year-olds, especially when there was a correct answer. In Study 2 with a nonverbal task with a correct answer, 3-year-olds produced indirect reasons more often than in Study 1, although 5-year-olds' indirect reasons were more efficiently stated. These results demonstrate that even 3-year-olds, and even nonverbally, can point out to a partner a fact that constitutes a reason for them to arrive at a correct joint decision.

Introduction
A key skill in collaborative problem solving is formulating reasons, and communicating them to one's partner about a suggested course of action. In fact, recent research on reasoning has emphasized...
the social dimensions of the process (Mercier & Sperber, 2011): individuals use reasons to justify to partners why they believe one thing versus another, all in pursuit of the best joint decision (Tomasello, 2019).

Most research on the development of children’s collaborative reasoning has focused on their comprehension of good versus bad reasons. For example, Koenig (2012) presented preschoolers with two informants with conflicting beliefs about the contents of a box. One supported her conclusion with a valid reason (e.g., “I looked and I saw an apple in the box”); the other did so with an invalid reason (e.g., “I like apples. I want there to be apples in the box”). Children were able to correctly infer the content of the box (see also Castelain, Bernard, & Mercier, 2018; Corriveau & Kurkul, 2014; Mercier, Bernard, & Clément, 2014).

Recently, young children’s production of reasons has been investigated, particularly in the context of collaborative problem solving. For example, 3-year-olds were able to recognize when they needed to give reasons—when their partner was ignorant (Baer & Friedman, 2018; Köymen, Mammen, & Tomasello, 2016). Köymen, Rosenbaum, and Tomasello (2014) had pairs of 3- and 5-year-old peers decide together where to place toy animals in a zoo. The 3-year-olds were skillful at providing relevant reasons. For instance, to justify where to place a penguin, they stated, “Because there is ice.” Domberg, Köymen, and Tomasello (2018) presented pairs of 5- and 7-year-old peers with a more challenging task. The task was more challenging because for the penguin, for instance, the dyads needed to choose one of two places that both had ice. Thus, simply pointing out a fact about the physical environment (where the ice is) was not sufficient; pointing out the limitations of the other option(s) was required. Thus, in this study, the 7-year-olds sometimes produced reasons that argued against one of the two options and referred to indirect evidence, so-called “indirect reasons,” leaving the other option as the best alternative by default (sometimes called reasoning by exclusion) such as “[The penguin should go here because] it is too crowded there [the other option].” The 5-year-olds relied mostly on “direct reasons” and referred to direct evidence, explaining why one option is better. They used indirect reasons much less often (see Felton & Kuhn, 2001, for adolescents).

Young children’s production of indirect reasons to justify their proposals, thus, has been studied only using the more difficult task in which there was no correct option—only better and worse options based on a comparison of evaluative judgments. Therefore, it is not clear whether the infrequent use of indirect reasons by young children is due to the difficulty of producing indirect reasons (e.g., holding conflicting perspectives simultaneously in mind and reasoning by exclusion) or to the kinds of tasks used (e.g., tasks without an objectively correct answer, tasks without any salient reason against one option). Tasks with an objectively correct answer are perhaps easier for young children because the children can simply refer to observable “facts” in the form of direct or indirect reasons to back up their claims. It is well documented that around 3 years of age children’s social cognition, particularly their reasoning, is dominated by the “objective perspective” (Tomasello, 2019). For instance, in false belief tasks, 3-year-olds systematically indicate that a person with a false belief would still act in accordance with reality (the “objective” truth), and they have a hard time representing two (subjective) perspectives simultaneously (Perner & Roessler, 2012; Wellman, Cross, & Watson, 2001). It is also around 3 years that children show sensitivity to objective “rights” and “wrongs” and justify their moral judgments through normative statements such as “It belongs there” (Köymen, Lieven, et al., 2014; Rakoczy, Warneken, & Tomasello, 2008).

In the current two studies, we investigated how 3- and 5-year-old children used direct and indirect reasons to convince their partners. In Study 1, we investigated whether the presence of an objectively correct answer to a task influences children’s production of these reasons spontaneously in their peer conversations. We asked 3- and 5-year-old peer dyads to place a toy animal (e.g., a polar bear) in a toy zoo because at this age children would be familiar with a zoo setup (Köymen, Rosenbaum, et al., 2014). In the correct-answer condition, there was one obviously correct option (e.g., an ice landscape) and one incorrect option (e.g., a savannah landscape). In the better-answer condition, both options were equally plausible (e.g., both options had an ice landscape, although they differed in other ways). We predicted that both age groups would mostly rely on direct reasons (e.g., “There is ice”) in both conditions (Domberg et al., 2018; Mercier, 2011). However, they might rely on indirect reasons more often in the correct-answer condition because it is easier to point out the incorrect alternative in the correct-answer condition (e.g., “That is too hot for polar bears”) than in the better-answer condition.
We predicted that 5-year-olds would produce indirect reasons more often than 3-year-olds because 5-year-olds would be better at considering the two conflicting perspectives simultaneously (e.g., favoring one perspective by explaining why the other is not good).

In Study 2, we focused on the correct-answer condition using a simpler (potentially nonverbal) task. We made the evidence for and against each option more salient and made the production of indirect reasons less linguistically demanding. We predicted that both 3- and 5-year-olds would be able to produce indirect reasons, although 5-year-olds would produce indirect reasons more efficiently than 3-year-olds.

**Study 1**

**Method**

**Participants**

72 3-year-olds (\(M = 3;8\) [years;months], range = 3;5–3;11; 36 girls) and 72 5-year-olds (\(M = 5;8\), range = 5;5–5;11; 36 girls) in 72 same-age and same-sex dyads participated in the study. The sample size was determined prior to data collection and was based on prior studies with a similar design (e.g., Köymen et al., 2016). An additional 6 dyads could not be included in the analyses because they did not complete the task (\(n = 3\) dyads) or the videos were accidentally deleted (\(n = 3\) dyads). Children were native speakers of German and had various socioeconomic backgrounds.

**Materials**

In each trial (two warm-up trials and three experimental trials), children needed to place a toy in one of the two homes, each of which had a background picture. In the first warm-up trial, a toilet needed to be placed in a bathroom or a living room. In the second warm-up trial, a refrigerator needed to be placed in a kitchen or a bedroom.

**Experimental trial with a polar bear.** In the better-answer condition, both options had ice backgrounds. Each option had two relevant objects in the foreground (fish and igloo in one option; meat and ice cave in the other). In the correct-answer condition, one option had an ice background with all four relevant objects; the other had a savannah background without any objects in the foreground. We avoided putting any irrelevant objects in the foreground of the incorrect option so as to avoid confusion and distraction in children.

**Experimental trial with a horse.** In the better-answer condition, both options had meadow backgrounds. Each option had two relevant objects in the foreground (a carrot and a saddle in one option; a brush and an apple in the other). In the correct-answer condition, one option had a meadow background with all four relevant objects; the other had a desert background without any objects in the foreground.

**Experimental trial with a bird.** In the better-answer condition, both options had forest backgrounds. Each option had two relevant objects in the foreground (a nest and seeds in one option; a birdhouse and a worm in the other). In the correct-answer condition, one option had a forest background with all four relevant objects; the other had an ocean background without any objects in the foreground.

**Procedure**

The procedure of this study was approved by the research ethics committee of Max Planck Institute for Evolutionary Anthropology, Department of Developmental and Comparative Psychology (project title: “Collaborative reasoning”). The study took place in quiet rooms of the nurseries. In the first warm-up trial, the experimenter (E) asked children to name the two rooms: the bathroom and the living room. E then pointed at the sealed box with the toilet and said, “You two will decide together to which room the toilet should go.” E then asked why after children chose the correct option so as to
prime children to provide reasons for their decisions. Children often said, “Because that is the bathroom.” E then repeated the procedure for the second warm-up trial with the refrigerator.

In the experimental trials, E said, “Here is a stack of three puzzles [see Fig. 1 for the setup]. Start with the top one, then the middle one, and finally the bottom one. In each of these three boxes, there is an animal. For each animal, you two need to decide together which of the two homes is better and explain why. If you find the better home, you both will get a big surprise gift!” E then left the room until children were done. The only difference between the conditions was the background pictures of the two homes for each animal. In the better-answer condition, each home had an appropriate background (see the “Materials” section). In the correct-answer condition, one home had the appropriate background; the other did not. The side of the homes (left vs. right) and the order of the target animals were counterbalanced. The whole procedure lasted about 10 to 15 min.

Coding
Children’s conversations with their peer partners were transcribed, with each line corresponding to a clause. We identified lines with reasons and coded each reason for its type:

- Direct reasons explaining why an animal should go to a home (e.g., “There is ice”)
- Indirect reasons explaining why an animal should not go to a home (e.g., “It will drown there”).

A second coder, blind to the hypotheses and conditions, recoded 22% of the transcripts (16 dyads, 4 in each age group and condition). The agreements for the identification of reasons and their type were $\kappa = .81$ and $\kappa = .80$, respectively.

Results
In the correct-answer condition, children in both age groups predominantly chose the correct home in the correct-answer condition (90% of the trials). In the better-answer condition, children’s preferences for each option were 42% vs. 58% for the polar bear, 19% vs. 81% for the bird, and 17% vs. 83% for the horse (the Appendix shows the breakdown of the choices in the two age groups). Thus, for the bird and the horse, both age groups showed a preference for one option over the other in the better-answer condition.

First, we analyzed whether the frequency of children’s reasons varied between age groups and conditions. Children produced a total of 318 reasons (4 dyads did not produce any reasons; see Fig. 2 for the mean numbers of reasons across age groups and conditions). We used a generalized linear mixed model (GLMM) with Poisson error distribution because GLMM allowed us to analyze the frequency of reasons while controlling for the total number of utterances (number of lines in the transcript) produced by each dyad with an offset term. The unit of analysis was each dyad ($N = 72$). The response variable was the number of reasons that each dyad produced. The full model included age group (3
or 5 years), condition (correct-answer or better-answer) and their interaction, gender, and the offset term of the number of utterances (log-transformed). The null model included gender and the offset term. The full model improved the fit as compared with the null model ($\chi^2 = 31.83$, $df = 3$, $p < .001$). However, the interaction between age and condition was not significant ($\chi^2 = 0.58$, $df = 1$, $p = .446$), so we dropped this interaction term. The reduced model revealed only a significant age difference ($\chi^2 = 28.89$, $df = 1$, $p < .001$) (Fig. 2); the 5-year-olds produced reasons more often than the 3-year-olds. There was no significant condition difference ($\chi^2 = 0.73$, $df = 1$, $p = .392$).

Next, we analyzed whether the frequency of direct and indirect reasons varied between age groups and conditions using GLMM with Poisson error distribution. The offset term was the total number of reasons. In the first analysis, the unit of analysis was each dyad ($N = 68$; 4 dyads that did not produce reasons were excluded from this analysis). The response variable was the number of direct reasons. The full model included age group (3 or 5 years), condition (correct-answer or better-answer) and their interaction, gender, and offset term of the total number of reasons (log-transformed). The null model included only gender and the offset term. The full model did not improve the fit as compared with the null model ($\chi^2 = 2.98$, $df = 3$, $p = .394$), suggesting that the frequency of direct reasons did not differ significantly across age groups or conditions (Fig. 3). In the second analysis, the response variable was the number of indirect reasons. The full and null models were the same as in the previous analysis. The full model improved the fit as compared with the null model ($\chi^2 = 29.35$, $df = 3$, $p < .001$). The interaction between age and condition was significant ($\chi^2 = 4.58$, $df = 1$, $p = .032$); the 5-year-olds produced...
indirect reasons more than the 3-year-olds (who produced almost none) in both conditions, but the difference was greater in the correct-answer condition.

Discussion

In this study, 5-year-olds produced more reasons than 3-year-olds in both conditions. Both age groups predominantly used direct reasons in both conditions. The 5-year-olds produced indirect reasons more often than the 3-year-olds, especially in the correct-answer condition. Children’s knowledge about the animals and their habitats is unlikely to explain this age difference because this age difference was observed only with indirect reasons and not with direct reasons. Nonetheless, children overall did not produce indirect reasons very frequently in their spontaneous peer interactions, in line with the findings of Domberg et al. (2018); the 3-year-olds produced almost none at all.

In the correct-answer condition one option was obviously correct, whereas in the better-answer condition either option was fine, so children’s proposals were rarely challenged, although children could still provide indirect reasons against their less favorite option (e.g., “There is too little grass there [for the horse],” “The igloo is for penguins [and not for polar bears]”). This agreement might have reduced the frequency of the reasons, especially the indirect reasons. Therefore, in Study 2 we focused more closely on the correct-answer condition and simplified the task. First, we gave children an adult partner who acted unsure and asked children to justify their decisions. Second, we also increased the saliency of the evidence for or against each option and reduced the linguistic demands of the task such that simply pointing at the evidence would be enough to justify children’s proposal. A recent study by Mascaro, Aguirre, Brun, Couderc, and Mercier (2019) used a nonverbal task and found that when 2- to 4-year-olds’ claims were challenged by an adult (e.g., “No, the dog is not there”), the children were able to correct the adult by pointing to where the dog actually was.

Study 2

In Study 2, we exposed children to a clean toy and a toy that leaves traces (e.g., has paint on its feet or wheels). Each toy was then hidden in one of two houses, and one house had traces in front of it. In the direct-evidence condition, children together with an adult partner were asked to identify the house containing the toy that left traces; in the indirect-evidence condition, they were asked to identify the house containing the clean toy. In both conditions, simply pointing at traces would be sufficient to justify a location. However, the pointing would be a direct reason in the direct-evidence condition (indicating where the sought-after toy is); it would be an indirect reason in the indirect-evidence condition (indicating where the sought-after toy is not). We predicted that both age groups would refer to the traces in both conditions, producing direct reasons in the direct-evidence condition and indirect reasons in the indirect-evidence condition. We also predicted that in the indirect-evidence condition 5-year-olds would be more efficient in their indirect reasons and would rely on less informative reasons, or simple reasons, such as pointing at the traces, whereas 3-year-olds might produce more informative or more elaborate indirect reasons (e.g., “This is the doll’s traces, so the car is in that house”) because sensitivity to over-informative utterances emerge around 5 years of age. That is, 5-year-olds, unlike 3-year-olds, start rejecting over-informative utterances or get confused by over-informative utterances (see Davies & Katsos, 2010; Morisseau, Davies, & Matthews, 2013).

Method

Participants

48 3-year-olds ($M = 3;10$, range = $3;6–4;0$; 24 girls) and 48 5-year-olds ($M = 5;10$, range = $5;8–5;10$; 24 girls) participated in the study. The sample size was determined prior to data collection based on prior studies with a similar design or similar analyses (e.g., Misch, Over, & Carpenter, 2016). An additional 20 3-year-olds (7 in the direct-evidence condition and 13 in the indirect-evidence condition) and 2 5-year-olds (1 in each condition) were excluded from analyses because they incorrectly
answered the question about the location of the toy. Children were native speakers of German and had various socioeconomic backgrounds.

**Materials**

We used a doll and a car hidden in two houses.

**Procedure**

The procedure of this study was approved by the research ethics committee of Max Planck Institute for Evolutionary Anthropology, Department of Developmental and Comparative Psychology (project title: “Collaborative reasoning”). The study took place in our laboratory. The child and the two experimenters (E1 and E2) first played with a puppet for familiarization. E1 pointed out that there were two houses in the room and then introduced the two toys: the doll and the car. The first toy (counterbalanced which toy) was introduced as the “clean toy.” E1 moved the clean toy onto a piece of paper and said, “All clean!” E1 then introduced the toy that leaves traces: “Look, this is the doll/car that loves paint [E1 dipped the feet/wheels in paint].” She moved the toy onto a piece of paper and pointed out, “Look at the traces!” E2 suggested playing a game and said, “I will hide the doll in one house and the car in the other house. You two will find out in which house X is.” In the direct-evidence condition, X was the toy that left traces. In the indirect-evidence condition, X was the clean toy. After E2 hid the toys, one house had traces leading to it (Fig. 4). Therefore, in the direct-evidence condition the traces indicated where the sought-after toy was, whereas in the indirect-evidence condition the traces indicated where the sought-after toy was not.

E1 and the child sat side-by-side facing the two houses behind a barrier. E1 said, “In which house is the X?” and waited for the child to point at a house. E1 repeated this question until the child picked a house. E1 then asked three why questions: “How do you know it is there?”; “How did you figure out that it is there?”; and “It is there because . . .?” If the child produced a reason after a question, E1 did not ask more questions. If the child answered the question about the location of the toy incorrectly and never revised his or her answer during why questions, the child was excluded from the analyses. The toy that left traces and the location of traces (left vs. right) were counterbalanced.

**Coding**

We coded each child’s reason with the following categories:

1. **Simple reasons:** Pointing at traces with or without minimal verbal explanations with the following four subcategories:
   - *Direct reasons* with reference to the following:
     - *traces* (e.g., pointing, “Because of the traces”)
     - *clean path* (e.g., “Because it is clean”)

![Fig. 4. Houses from the perspective of the child and Experimenter 1.](image)
• **Indirect reasons** with reference to the following:
  – traces (e.g., pointing, “Because of the traces”)
  – clean path (e.g., “Because it is clean”)

2. **Elaborate reasons:** Identifying the toy or giving extra information with the following two subcategories:
   • **Direct reasons** (e.g., “There are the doll’s traces [so there is the doll]”)
   • **Indirect reasons** (e.g., “There are the doll’s traces [so there is the car]”)

3. **Irrelevant/no reasons** (e.g., “I heard it”).

It should be noted that same reasons were coded as direct or indirect depending on the proposal they justified. Some children spontaneously justified their choices before the *why* questions. E1 asked these children the *why* questions again, and we included the more elaborate reason of the two in the analyses. A second coder recoded 25% of the types of children’s reasons (24 children, 6 in each age group and condition). The agreement was $\kappa = 1.00$.

**Results**

In response to the question “In which house is the X?”, 84 of 96 children (87.5%) gave a correct response, and 12 children (12.5%) initially gave an incorrect response but revised their answer after being asked why. In addition, 79 of 96 children (82%) provided a relevant reason for their proposal.

First, we compared the two age groups in each condition in terms of the number of children who produced a relevant reason. In the direct-evidence condition, there was no significant difference between the numbers of 3- and 5-year-olds who produced a relevant reason, $\chi^2(1, N = 48) = 0.19, p = .663$. In the indirect-evidence condition, the number of 5-year-olds who produced a relevant reason was significantly higher than that of 3-year-olds, $\chi^2(1, N = 48) = 4.25, p = .039$ (Fig. 5).

Next, we focused on the indirect reasons and whether the number of children who produced indirect reasons varied depending on age and condition. In each age group, the number of children who used an indirect reason was significantly higher in the indirect-evidence condition than in the direct-evidence condition [3-year-olds: $\chi^2(1, N = 35) = 13.17, p < .001$; 5-year-olds: $\chi^2(1, N = 44) = 21.15, p < .001$] (Fig. 6). There was no significant difference between the numbers of 3- and 5-year-olds who produced an indirect reason in the indirect-evidence condition, $\chi^2(1, N = 37) = 0.58, p = .446$. That is, both age groups predominantly relied on direct reasons in the direct-evidence condition and on indirect reasons in the indirect-evidence condition.

Finally, we compared whether the number of children who produced elaborate reasons varied depending on age and type of reason (direct or indirect reason). Because children almost never used indirect reasons in the direct-evidence condition (Fig. 6), we did not look at whether children’s use of
elaborate reasons varied across conditions. There was no significant difference between the numbers of 3- and 5-year-olds who produced an elaborate direct reason, $\chi^2(1, N = 52) = 1.69, p = .193$; both age groups mostly used simple direct reasons (Fig. 7). However, the number of 3-year-olds who produced an elaborate indirect reason was significantly higher than that of 5-year-olds, $\chi^2(1, N = 27) = 4.69, p = .030$.

**Discussion**

In this simplified task, children were able to produce indirect reasons more frequently than they did in Study 1. In the direct-evidence condition, 3- and 5-year-olds were very similar; both age groups provided simple reasons at equal rates, mostly by pointing at the traces. In the indirect-evidence condition, however, 3-year-olds had a harder time with the task, as evidenced by the high rates of drop-outs; two thirds of the 3-year-old dropouts were in the indirect-evidence condition (see also Mascaro et al., 2019, who also found reduced accuracy without adult scaffolding in a similar task). Moreover, 3-year-olds provided a (direct or indirect) reason in the indirect-evidence condition less often than 5-year-olds. The pattern of 3-year-olds might be explained by the difficulty of cognitively representing two conflicting perspectives (see Wellman et al., 2001).

Although there was no significant age difference in the frequency of indirect reasons, 3-year-olds produced more elaborate indirect reasons than 5-year-olds. This finding that children were over-informative stands in contrast to studies showing that preschoolers can produce statements that are often optimally informative (Köymen, Rosenbaum, et al., 2014; Köymen et al., 2016) or under-
informative (Matthews, Lieven, & Tomasello, 2007). Nonetheless, the more elaborate indirect reasons produced by the 3-year-olds in our study can be explained by their difficulty in referring to indirect evidence. As evidenced by both studies, producing indirect reasons required less effort for 5-year-olds, who could judge the minimal amount of information necessary to communicate their reasons.

Overall, these findings are in line with those by Mascaro et al. (2019), who found that nonverbal tasks facilitate young children’s reason giving. But children’s reason giving in their study was limited to direct reasons; children actually showed the adult where a toy was (turned around a box) to support their claim. Our findings go beyond this by demonstrating that children are able to communicate indirect reasons, that is, not by showing or pointing to a toy’s location but rather by indicating the evidence (the traces) for believing the toy to be in one location rather than in another.

General discussion

In two studies, we found that when referring to direct evidence, 3- and 5-year-old children reason in very similar ways with their partners. However, by 5 years of age, and only in simplified situations at 3 years, when children are able to cognitively represent two conflicting perspectives, they are able to refer to indirect evidence (indirect reasons) for their proposals by explaining why an alternative is incorrect. Particularly simplified nonverbal tasks, in which the evidence for or against an option was salient, facilitated the production of indirect reasons by preschool children. In contexts of collaborative problem solving where partners justify their beliefs to one another, especially when the linguistic demands are minimized, young children display sophisticated reasoning skills (e.g., reasoning by exclusion) that nevertheless develop significantly across the preschool years.

Acknowledgments

We thank Elena Rossi, Claudia Salomo, Anne Tomm, Sinja Thielicke, and Henning Francik for their help with recruitment and data collection; Franziska Zenner for her help with coding; and all the day-care centers and children for their friendly cooperation.

Appendix

Percentages (and numbers) of dyads that chose each option

<table>
<thead>
<tr>
<th>Correct-answer condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Polar bear</td>
</tr>
<tr>
<td>Horse</td>
</tr>
<tr>
<td>Bird</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Better-answer condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Polar bear</td>
</tr>
<tr>
<td>Horse</td>
</tr>
<tr>
<td>Bird</td>
</tr>
</tbody>
</table>

Note. The percentage of dyads that chose each option (the number of dyads). In the better-answer condition, for the bird, Option 1 had the birdhouse and the earthworm; Option 2 had the nest and the seeds. For the horse, Option 1 had the brush and the apple; Option 2 had the saddle and the carrot. For the polar bear, Option 1 had the fish and the igloo; Option 2 had the cave and the meat.
References


