

Children's selective trust in promises

Margherita Isella^{1,2,*}, Patricia Kanngiesser^{2,3,*}, Michael Tomasello^{2,4}

¹Department of Philosophy, Vita-Salute San Raffaele University, Milan

²Department of Developmental and Comparative Psychology, Max Planck Institute for
Evolutionary Anthropology, Leipzig

³Faculty of Education and Psychology, Free University Berlin, Berlin

⁴Department of Psychology and Neuroscience, Duke University, Durham

*both authors contributed equally to this manuscript, correspondence should be addressed to:
margherita.isella1@gmail.com or patricia.kanngiesser@fu-berlin.de.

Funding for this research was provided to Margherita Isella by a Research Grant for Doctoral Candidates and Young Academics and Scientists (6 months), 2014 (50015537) from the German Academic Exchange Service (DAAD). We would like to thank Michele Herbstritt and Roger Mundry for support with data analysis and visualization, and Carl Bartl, Adrian Clement, Jana Jurkat, Kristina Kellermann, Benedikt Kovacs and Regina Sticker for their help in running the study and transcribing and coding the data. Thanks also go to Claudia Bianchi, Malinda Carpenter, Ivan Gonzalez-Cabrera, Neri Marsili and Barbora Siposova for helpful discussions.

This is a preprint version of the published article: Isella, M., Kanngiesser, P., & Tomasello, M. (2019). Children's selective trust in promises. *Child development*, 90(6), e868-e887.
<https://doi.org/10.1111/cdev.13105>

Abstract

There has been extensive research into the development of selective trust in testimony, but little is known about the development of selective trust in promises. The present research investigates children's (N = 264) selective trust in other's promises to help. In Study 1, six-year-olds selectively trusted speakers who had previously kept a promise. In Study 2, five-year-olds displayed selective trust for speakers who had previously kept a prosocial promise (promise to help). In Study 3, five-year-olds trusted a speaker, who kept a prosocial promise, over a helper. These data suggest that from the age of five children show selective trust in others' promises using prosociality, promise-keeping or both to inform their judgements.

Promises are speech acts by which speakers commit to performing future actions that are desirable for hearers (Searle, 1969). Some philosophers have suggested that promises are invitations to trust (Friedrich & Southwood, 2011; Scanlon, 1998), and play an important role in enabling cooperation (Bicchieri, 2002). In experimental studies with adults, rates of cooperation increased when players communicated and made promises to each other (Orbell, Van de Kragt, & Dawes, 1988; Ostrom, Walker, & Gardner, 1992; Sally, 1995). Yet, to avoid being subjected to free-riders, who use promises as “cheap talk” to gain others’ trust and cooperation, it is crucial to distinguish between genuine, trustworthy promises and untrustworthy promises. In particular, selective trust in those who keep their promises (and distrust in those who do not) is crucial for sustaining cooperative interactions.

The ability to understand commissive speech acts such as promises and the norms that regulate their use emerges gradually during childhood. From their pre-school years, children understand that promises ought to be kept: when another person breaks their promise to the child, 3- and 5-year-olds spontaneously protest by appealing to promissory obligations (‘you should keep your promises’), and when children themselves promise or agree to do a task, they persist longer than without a promise (Kanngiesser, Köymen, & Tomasello, 2017). By the age of 6, children take into account speakers’ intentions (rather than just their actions and the consequences of these actions) when evaluating commitments: they blame speakers for breaking their commitments intentionally but not for breaking them because of events beyond their control (Mant & Perner, 1988). Finally, from 9 years of age children understand that only commissive speech acts entail an obligation to perform the promised act: they judge the speaker as “naughty” for unfulfilled promises, but not for unfulfilled statements of intentions (‘I am going to go swimming’) or predictions (‘You will feel better tomorrow’; Astington, 1988; Mant & Perner, 1988). In summary,

during early and middle childhood, children develop the ability to assess *whether speakers have kept their promises*, and to what extent they can be held responsible. But do children also have the ability to assess *whether they can trust speakers to keep their promises*?

In order to make informed decisions about whether to trust speakers' promises, children need to be able to *evaluate* speakers' reliability (i.e., speakers' willingness and ability to provide true information; Sperber et al., 2010) and to *use* that evaluation when making a decision (see, Birch, Vauthier, & Bloom, 2008; Koenig & Harris, 2005). Speakers' reliability can be inferred through a number of cues, for example: speakers' motives and preferences, speakers' abilities to keep their promises, and speakers' past behavior.

Previous research has shown that children evaluate others based on their past promise-behavior: Six-, eight- and ten-year-olds were asked to rate how much they trusted a character who either kept its promise to help a third party or who broke its promise (Rotenberg, 1980). From six years of age, children assigned higher trust values to the promise-keeping character. However, the development of trust judgements before six years of age remains unclear, as the study did not test younger children. Moreover, the study did not investigate whether children would use their evaluation when deciding whether to trust future promises.

There is some evidence that younger children regulate their own behaviour in response to others' past reliability in first party situations. Specifically, Kidd, Palmeri, and Aslin (2013) found that three- to five-year-olds waited longer in a delay of gratification task if the experimenter had previously kept her word to the child. Yet, it is unclear whether children waited longer because they trusted the experimenter more or because they had previously experienced a rewarding interaction with her.

Children's trust in reliable sources of information has been thoroughly studied in the context of children's learning from others' testimony (i.e., information communicated by others via assertions; see, Heyman & Legare, 2013; Mascaro & Morin, 2014; Mills, 2013, for overviews). For instance, two-year-olds were more likely to learn new labels for *familiar* objects from an accurate speaker (i.e., who had provided accurate information about other familiar objects) than an inaccurate one; yet, they learnt labels for *novel* objects indiscriminately from both speakers (Koenig & Woodward, 2010; Krogh-Jespersen & Echols, 2012). This suggests that two-year-olds selectively override their own beliefs, but still trust others generally when learning novel information. From three to four years of age children become more selective: After receiving conflicting testimony about (novel) objects from two speakers, three-year-old children trusted the testimony of a benevolent speaker (i.e., a speaker described as nice and behaving nicely) more than that of a malevolent one (i.e., a speaker described as mean and behaving badly; Mascaro & Sperber, 2009), and the testimony of a knowledgeable speaker (i.e., a speaker declaring to have information about familiar objects) more than the testimony of an ignorant one (i.e., a speaker declaring to not have information about familiar objects; Harris et al., 2012; Koenig & Harris, 2005). Between 3 and 4 years of age, children begin to trust a previously accurate speaker over an inaccurate one (Koenig, Clement, & Harris, 2004; Corriveau, Meints, & Harris, 2009; Pasquini, Corriveau, Koenig, & Harris, 2007; Rakoczy, Warneken, & Tomasello, 2009). By four to five years of age children also become gradually better at rejecting the testimony of a speaker who has been described as or shown to be dishonest (Heyman, Sritanyaratana, Kimberly, & Vanderbilt, 2013; Mascaro & Sperber, 2009; Vanderbilt, Liu, & Heyman, 2011).

The above line of research has focused on children's social learning of facts and thus has tested children's ability to evaluate assertions. However, assertions and promises are two different

illocutionary acts: in performing them, speakers are ‘doing different things with words’ (Austin, 1962). Specifically, the two speech acts differ in four out of the seven defining features of illocutionary acts (Searle & Vandervaken, 1985; see also, Brandom, 1994; Watson, 2004, for similar comparisons of assertions and promises): (1) Assertions and promises have different purposes (illocutionary points): In asserting, speakers commit to a proposition *being* true. For instance, in asserting ‘Tom is a dancer’ the speaker commits to the truth of the proposition ‘Tom is a dancer’. In promising, speakers commit to *making* a proposition true. For instance, in promising ‘I will pick up the cake’ the speaker commits to making the proposition ‘I will pick up the cake’ true. (2) Assertions and promises have different constraints on the propositional content: assertions can concern a past, present or future fact (i.e., speakers’ actions or other types of facts). In particular, all testimony studies presented children with past or present facts (e.g., ‘There is a spoon in the box’; Mascaro & Sperber, 2009). Promises, however, always concern future actions of speakers. (3) Assertions do not need to concern events that are desirable for hearers. Promises always concern actions that are desirable from hearers’ point of view. For example, saying ‘I promise I will bring cake’ is only a promise if the hearer desires that the speaker brings cake. (4) Assertions express speakers’ beliefs: For example, the assertions ‘Tom is a dancer’ expresses the speaker’s belief that Tom is a dancer. Promises express speakers’ intentions: For example, the promise ‘I promise to bring the cake’ expresses the speaker’s intention to bring the cake. Given these differences, it is an open question whether selective trust in promises follows a different or the same developmental trajectory as selective trust in assertions.

We therefore tested whether three- to six-year-olds would selectively trust promises of reliable speakers. Specifically, children could choose between two speakers that promised to help the child. One speaker had previously kept a promise to a third party and the other speaker had

previously broken a promise to a third party. If children take speakers' reliability into account when evaluating a new promise, they should choose the speaker that had previously kept its promise. We focused on past behavior as a reliability cue because it has been frequently used in previous studies on children's selective trust in testimony (e.g., Clement et al., 2004). We first tested three-year-olds as a number of studies found selective trust to emerge at this age in binary choice paradigms (Birch et al., 2007; Harris et al., 2012; Mascaro & Sperber, 2009). Then, we successively added older age groups until we had determined the age at which children showed selective trust in promises. This allowed us to map, for the first time, the developmental trajectory of selective trust in promises.

Study 1a

In Study 1a, we tested whether 3-year-olds selectively trusted promises of reliable speakers. Children saw videos where speakers were asked to promise to a third party (i.e., a promise to brush its teeth or to clean up its toys). In one video, the speaker kept the promise and in another video, a different speaker broke the promise. The promises concerned somewhat unpleasant every-day tasks that we expected even young children to be familiar with. Moreover, the promises always involved self-directed actions to minimize a prosocial interpretation. After the video sequences, the two speakers promised children to help retrieve some puzzle pieces that were out of children's reach. If children take speakers' reliability into account when evaluating a new promise, they should choose the speaker that had previously kept its promise to the third party.

Method

Participants. Twenty-four 3-year-olds (*Mean*: 3 years 6 months, *Range*: 3 years 4 months – 3 years 8 months, 14 girls) participated in Study 1a. Two additional children took part in the

warm-up but were excluded from the final sample because they refused to complete the procedure. Children were recruited in a medium-sized German city through a database of parents who had agreed to have their children participate in developmental studies in their child's kindergarten. They were all native speakers of German and came from various socioeconomic backgrounds. The database contained around 20000 children and was only accessible to a very limited number of people for recruitment purposes. All parents were provided with information about the lab, gave informed consent, and had the possibility to contact the lab at any time as well as to withdraw their consent. Data were collected in 2015.

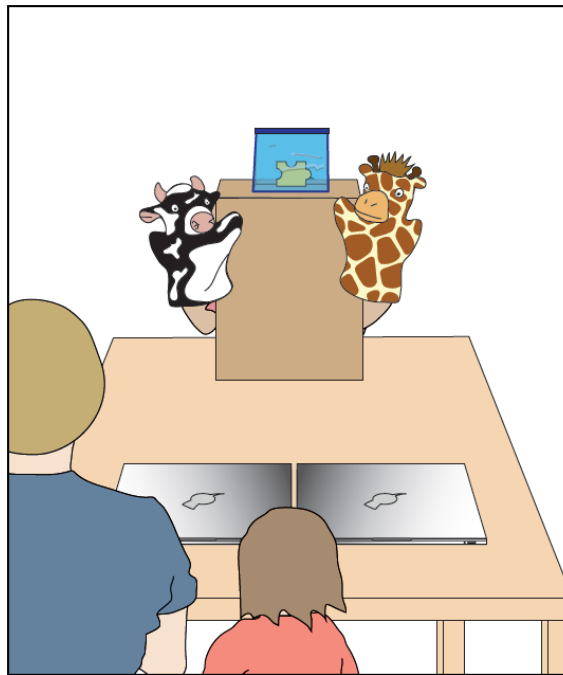


Fig. 1. Set-up of Studies 1, 2 and 3. E1 sat next to the child and E2 operated the two hand puppets (while hiding behind the box). The clear plastic bag contains the puzzle piece(s) that children needed to complete their puzzle. Videos were shown on two laptops, placed side-by-side in front of the child.

Materials/Set-Up. We set up the task in a quiet room in children's kindergarten. The setup included a table with a game (a puzzle in one trial, a wooden necklace in the other trial; order counterbalanced across children), two laptops, a big basket, and a plastic bag with puzzle/necklace

pieces (for details, see Fig. 1). Children participated in two trials, in which they interacted with different animal puppets (played by a female experimenter – E2): a frog and a mouse in the puzzle trial, and a cow and a giraffe in the necklace trial. These hand puppets were the speakers that promised to children. We used animal hand puppets because they were child-friendly and to avoid confounds due to speaker characteristics. Hand puppets have been widely used in testimony studies (e.g., Clément, Koenig, & Harris, 2004; Einav & Robinson, 2011; Mascaro and Sperber, 2009; Rakoczy, Warneken, & Tomasello, 2009; Robinson and Nurmsoo, 2009).

Procedure. The procedure was divided in five phases: (i) a familiarization phase to introduce children to the experimenters and the set-up, (ii) a game phase to create a helping situation, (iii) a video phase to provide children with information about the speakers' past behaviour, (iv) a control question phase to assess children's understanding of the videos, and (v) a selection phase to test children's selective trust in the puppets.

Familiarization. A male experimenter (E1) briefly warmed up each child in his or her kindergarten group (e.g., by playing with a music toy car). In the testing room, E1 introduced the child to E2 and familiarized the child with the study set up. E2 sat quietly behind the basket to animate the hand puppets.

Game phase. In each trial, E1 showed the child a game. We used a puzzle in one trial and a necklace game in the other trial to keep children interested in the games. E1 then encouraged the child to finish the respective game, while E1 turned away to read. When the child had used all the available game pieces, E1 pointed out (if the child did not find out by him or herself) that one piece was missing and that this missing piece was in a transparent box on the basket (out of the child's reach). The two animal puppets appeared on either side of the transparent box. E1 told the child

that one of them could help the child retrieve the missing piece, but that they would first watch some videos.

Video phase. During the videos, the puppets rested next to the transparent box, pretending to sleep. Children watched two videos, one about each hand puppet. The use of video stimuli, and not live demonstrations, ensured that the presentation of the two speakers was closely matched and that all participants watched the same presentations (see, Corriveau & Harris, 2009; Koenig & Harris, 2005; Liu, Vanderbilt, & Heyman, 2013, for the use of video stimuli in testimony studies). E1 showed only one video per laptop, so that the videos matched the locations of the puppets (i.e., on the left laptop, E1 showed a video about the left puppet). E1 always showed the videos from left to right. Each video lasted approximately 2 minutes. In each video, a bear-character asked the respective hand-puppet (frog/mouse/cow/giraffe) to promise to do a task. Specifically, in the puzzle trial the bear asked the hand puppet (frog/mouse) to promise to brush its teeth before reading a book (*Please, first brush your teeth and then you can read your favorite book. Do you promise that you will first brush your teeth?*). In the necklace trial, the bear asked the cow/giraffe to promise to clean up its toys before eating cookies (*Please, first clean up your toys and then you can eat the cookies. Do you promise that you will first clean up?*). The hand puppet then promised to do the task (*Yes, I promise that I will first brush my teeth.*). The next part varied between videos: in one video, the respective puppet kept its promise (while saying *Now I brush my teeth.*); in the other video, the respective puppet broke its promise (e.g., immediately read the book). Both videos ended in the same way with the speaker reading/eating (*Now I read the book / Now I eat the cookies.*). See Figure 2 for screenshots and the video script.

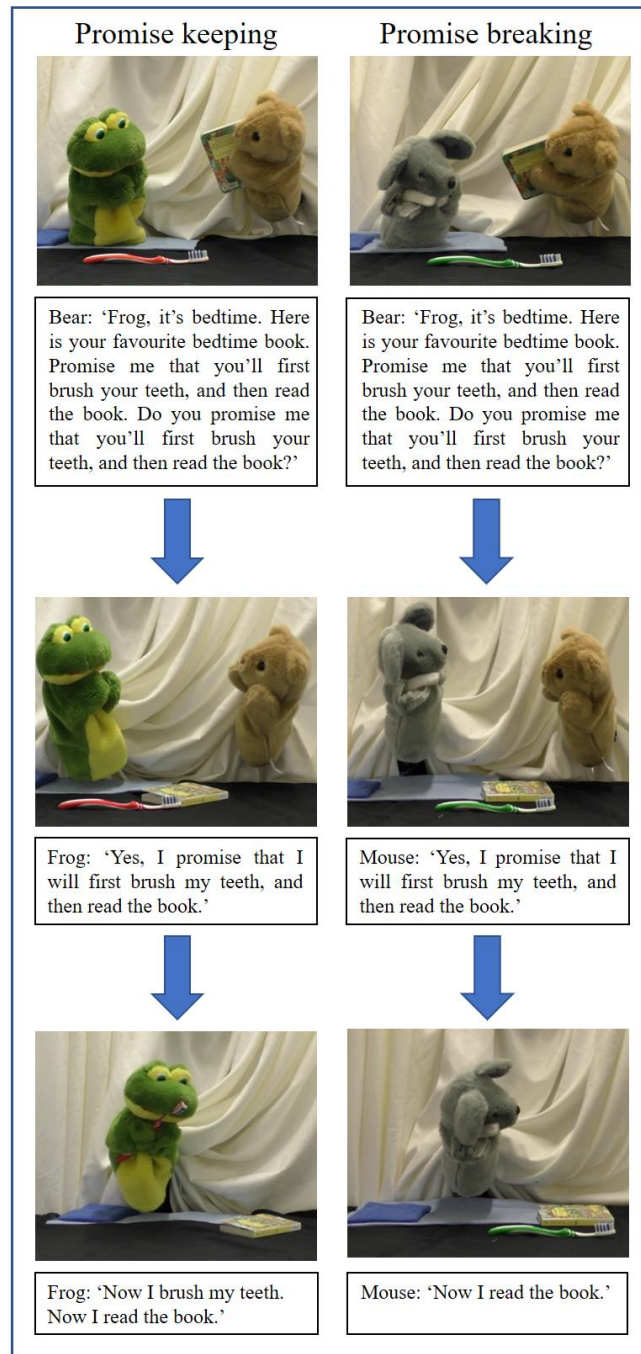


Fig. 2. Screen shots of the video stimuli and the video script used in Study 1 (see the online supplementary material for the videoclips). In both videos, the target puppets (frog/mouse) promised a third-party (Bear) to perform a task. Then one target puppet kept the promise ('promise-keeping' video), while the other broke it ('promise-breaking' video).

Control questions phase. After each video, E1 asked up to three control questions: (1) *What happened in the videos?* (2) *What did the [name of respective puppet, e.g. frog] say to the bear?* (3) *What did the [name of respective puppet, e.g. frog] do?* If children remembered both the speech act and the promised action after the first or second question, E1 did not ask any further control questions (to avoid eliciting the same answers repeatedly). If children failed to answer questions 2 or 3, E1 repeated the question once. We later coded children's answers to all of the control questions (see, Data Coding and Analyses, for details). The aim of the control questions was twofold: (1) to check whether children remembered the key events in the videos, i.e. the promise speech act and the promised action, and (2) to test whether children with better memory of the events in the videos were more likely to choose the reliable speaker. This allowed us to assess whether a failure to choose the reliable speaker would be due to a failure to comprehend the key events in the videos.

Selection phase. After the video phase, both hand-puppets promised to help the child (*Oh, I can give you the box, I promise that I will help you and give you the box.*). E1 then asked the child which hand-puppet she or he would like to receive help from (*Would you like the frog or the mouse to help you?*). However, in order to avoid giving children feedback on the hand-puppet's actual behavior between trials, E1 retrieved the missing pieces from the transparent box (*You know what? I can do it.*).

Across children, we counterbalanced the order of trials (i.e., puzzle or necklace) and the speakers' position in each trial (i.e., left side vs. right side). We also counterbalanced across children which speaker kept or broke its promise to prevent potential confounds of animal type and promise behaviour. Moreover, we counterbalanced the side of the promise-keeping speaker

for each child across the two trials. Each session lasted approximately 15 minutes. We videotaped the session from the moment the child entered the testing room.

Data coding and analyses. We coded which speaker children chose as a measure of their selective trust. For each of the two trials, we assigned a score of 1 if the child chose the reliable speaker and a score of 0 if the child chose the unreliable speaker. The scores were summed across trials in an overall *choice score* (range = 0-2). We tested choice scores against chance (chance = 1) using the *exactRankTests* package for non-parametric tests in R (Hothorn & Hornik, 2013; R Core Team, 2016).

In addition, we transcribed all of children's answers to the control questions and coded whether they recalled the promise speech act and the promised action:

a) *Speech act recollection.* Children recalled the content of the promise speech act that the target speaker uttered to the third party. For example, 'Erst Zähneputzen.' (*First teeth brushing.*), 'Okay, ich räum meine Spielsachen auf.' (*OK, I clean up my toys.*).

b) *Action recollection.* Children recalled whether the speaker did or did not do the promised action. For example, 'Der hat Zähne geputzt und das Buch gelesen.' (*He brushed his teeth and read the book.*), 'Der hat nicht Zähne putz.' (*He did not brush his teeth.*).

Children received a set of control questions after each video (2 videos per trial). For each set of control questions, we coded whether children mentioned the relevant speech act and/or action. Note that we only scored the two coding categories once per set of questions. We then summed children's recollection for the two videos per trial (range = 0-2 per category). We tested for each trial whether children's choice score (reliable speaker: yes/no) correlated with their recollection scores using point-biserial correlations.

All statistical tests were two-tailed. We established the significance threshold at .05.

25% of children's answers were coded by a second coder. Agreement between coders on choice scores was perfect (choice: Cohen's $\kappa = 1$) and on recollection scores was perfect to very good (speech act: Cohen's weighted $\kappa = .93$; action: Cohen's weighted $\kappa = 1$).

Results and discussion

Contrary to what we predicted, three-year-olds chose the reliable speaker at levels that did not differ significantly from chance ($M = 1.08$, $SD = 0.78$, $V = 60$, $p = .791$, Exact Wilcoxon-test; see Fig. 1). Eight children chose the reliable speaker in both trials, ten children in one trial, and six children never chose the reliable speaker ($\chi^2(2) = 1$, $p = .607$). To test whether E1's retrieval of the missing pieces after trial 1 affected children's choices, we analysed the first trial separately: 16 of 24 children chose the reliable speaker in trial 1 ($p = .152$; binomial test, two-sided).

When answering the control questions, 3-year-olds scored on average 0.54 (of 2; $SD = 0.74$) for recalling the speech act, and 1.46 (of 2; $SD = 0.74$) for recalling the action (for further details, see Supplementary Table 1). There was no correlation between children's choice score and their recollection scores: neither for the speech act category (trial 1: $r_{pb} = .110$, $p = .610$; trial 2: $r_{pb} = -.077$, $p = .722$), nor the action category (trial 1: $r_{pb} = -.184$, $p = .389$; trial 2: $r_{pb} = .244$, $p = .251$). Three-year-olds' choices of the reliable speaker were thus not related to their memory of the events in the videos.

In Study 1a, we failed to find evidence that 3-year-olds selectively trusted promises of reliable speakers. Three-year-olds' may have performed poorly because the procedure was too elaborate and lengthy (e.g., the critical choice questions were always preceded by a series of control questions). We thus decided to shorten and simplify the procedure in Study 1b.

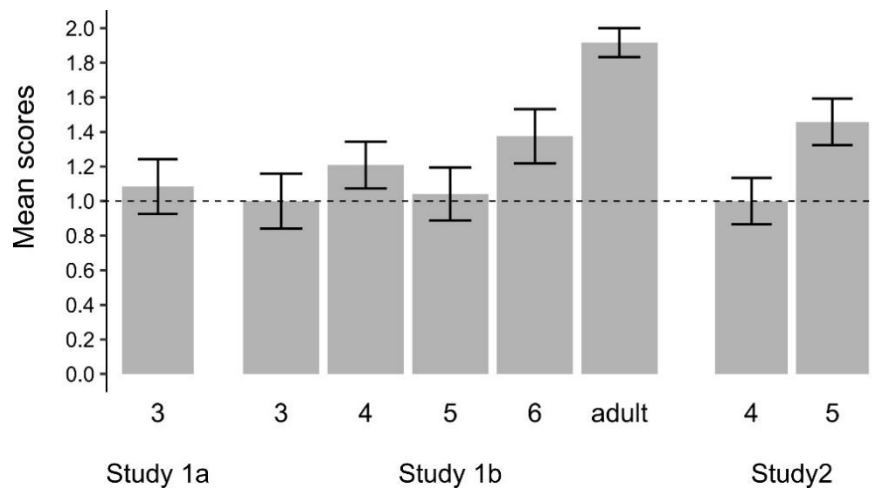


Fig. 3. Mean choice scores (\pm SE) of participants as a function of age in Studies 1a, 1b and 2. Chance (1) is indicated by the dotted line.

Study 1b

In Study 1b, we asked the choice question immediately after children had watched the videos. We again tested three-year-olds and successively added older age groups until we had determined the age at which children would selectively trust the reliable speaker. In order to ensure that the video stimuli provided meaningful information about speakers' reliability and resulted in the predicted pattern of behaviour, we also tested a sample of adults.

Method

Participants. Twenty-four 3-year-olds (M : 3 years 6 months, $Range$: 3 years 4 months – 3 years 8 months, 12 girls), twenty-four 4-year-olds (M : 4 years 6 months, $Range$: 4 years 4 months – 4 years 8 months, 11 girls), twenty-four 5-year-olds (M : 5 years 6 months, $Range$: 5 years 4 months – 5 years 8 months, 12 girls) and twenty-four 6-year-olds (M : 6 years 6 months, $Range$: 6 years 4 months – 6 years 8 months, 10 girls) participated in Study 1b. Eight additional children

were excluded from the final sample: four 3-year-olds, one 4-year-old, two 5-year-olds due to refusal to complete the procedure, and one 6-year-old due to distraction (peers entering the room). Participants were recruited and tested in the same manner as in Study 1a, except that the majority (N = 20) of 6-year-olds were tested at the after-school club of their elementary school.

In addition, twelve adults (above the age of 18, 6 women) took part in this study. They were members of the research institute in a medium-sized German city (members of the psychology department did not participate). They were all German native speakers.

Materials and procedure. The procedure was similar to Study 1a with the following modifications:

In the game phase, we used the same puzzle game across the two trials to simplify the procedure. To adjust the difficulty of the puzzle game to different age groups, we used different puzzles: the same puzzle as in Study 1a for 3-year-olds, another one for 4- and 5-year-olds, and a third one for 6-year-olds. Three-year-olds were missing two puzzle pieces in total, while older children were missing four puzzle pieces. One (two) piece(s) became available in the first and in the second trial, respectively.

In the selection phase, we simplified the procedure as follows: E1 did not retrieve the puzzle piece(s) between trials, but instead suggested to continue searching for the other missing piece(s). While the child searched, E2 secretly placed a bag with new piece(s) on the box and drew the child's attention to it (*Look, over there, the second piece is in that bag!*). At the very end of the procedure, the puppets the child had chosen handed over the bags.

We further simplified the control question phase by only asking control questions about one of the videos (i.e., the video of the speaker the child had chosen) at the end of trial 2. Finally,

since some six-year-old children spontaneously justified their choices during piloting, we asked 6-year-olds an additional question after the control questions: *Why did you choose the [name of the speaker the child chose] and not the [name of the speaker the child did not choose]?*

Adult participants were tested with the same procedure, except that we used a more difficult puzzle where about a third of the pieces were missing.

Data coding and analyses. We coded which speaker children chose as a measure of their selective trust, using the same method as in Study 1a. We analyzed the data in R using the `exactRankTests` package for non-parametric tests. We tested (1) for each age group whether children chose the reliable speaker significantly more often than chance using Wilcoxon signed rank test, and (2) whether age had an effect on children's choice scores using Kruskal-Wallis test.

As in Study 1a, we also transcribed children's answers to the control questions and coded for their recollection of the *speech act* and *action*. Since we asked each child only one set of control questions, we assigned one score (0 or 1) per category. For each age group, we tested for correlations between recall and choice scores in trial 2 (because we only asked the control questions in trial 2) using phi-correlations for two dichotomous variables.

We coded 6-year-olds' justification of their choice using the following categories:

a) *Speech act justification.* Children justified their choice by referring to the speaker's promise to the third party. For example, 'Weil der Frosch zum Bär versprochen hat hoch und heilig und dann hat er es nicht gemacht.' (*Because the frog promised the bear solemnly and then did not do it.*), 'Weil der Frosch gelogen hat und die Maus nicht.' (*Because the frog lied and the mouse did not.*).

b) *Action justification.* Children justified their choice by referring to the fact that the reliable speaker performed the promised action, and/or that the unreliable speaker did not. For example, 'Weil die Giraffe die Spielsachen eingeräumt hat.' (*Because the giraffe cleaned up the toys.*), 'Weil sie es richtig gemacht hat.' (*Because she did right.*).

c) *Socio-evaluative justification.* Children justified their choice by evaluating the speaker's behavior or appearance. We included appearance as previous studies have shown that children and adults (Bascandziev & Harris, 2014; 2016; Willis & Todorov, 2006; Todorov, Pakrashi, & Oosterhof, 2009) use it as a cue to reliability. For example, 'Weil der Frosch unerzogen ist.' (*Because the frog is not well behaved.*), 'Weil die Giraffe schöner ist.' (*Because the giraffe is prettier.*).

Statements such as 'Weil die Maus (...) hat sich nicht an die Regel gehalten.' (*Because the mouse did not follow the rule.*) that could refer both to the speech act ('rule'), and the action ('did not follow') were scored in both categories. Next, we calculated the percentage of children who gave a valid justification (i.e., a justification that belonged to one or more of the three categories). For each coding category, we calculated the percentage of children who uttered justifications in that category (note that utterances could be scored in more than one category).

All coding was done by the same coder as in Study 1a. A second coder scored 25% of children's answers (an equal number of children per age group and gender) and 50% of adults' choices. Agreement between coders was perfect on choice (Cohen's $\kappa = 1$), good on recollection scores (speech act: $\kappa = .83$; action: $\kappa = .80$), and good on six-year-olds' justifications ($\kappa = .83$).

Results and discussion

Adults chose the reliable speaker above chance levels ($M = 1.92$, $SD = 0.29$; $V = 66$, $p < .001$), indicating that the video stimuli produced the predicted pattern of behaviour (in adults). Three-, 4-, and 5-year-olds chose the reliable speaker at levels that did not differ significantly from chance (3-year-olds: $M = 1.00$, $SD = 0.78$, $V = 52.5$, $p > .999$; 4-year-olds: $M = 1.21$, $SD = 0.66$; $V = 48$, $p = .227$; 5-year-olds: $M = 1.04$, $SD = 0.75$; $V = 49$, $p > .999$; see Fig. 1). By 6 years of age, children chose the reliable speaker significantly above chance ($M = 1.38$, $SD = 0.77$, $V = 117$, $p = .049$). When testing for effects of age on children's choice scores, we found no significant difference, $\chi^2(3) = 3.95$, $p = .267$, likely because the majority of the sample (i.e., 3- to 5-year-olds) behaved very similarly.

Next, we investigated 6-year-olds' justifications for their choices: 82% of children who chose the reliable speaker gave a valid response, while only 29% of children who chose the unreliable speaker did so. Breaking valid responses down into categories, we found that children who chose the reliable speaker referred to the speech act (29%), the action (71%), or social-evaluation (29%). Children who preferred the unreliable speaker only referred to social evaluation (see Table 1).

Table 1. Percentage of valid justifications (by children's choice score) per category in Study 1b and 2 (justifications could be scored in multiple categories).

Study	Age (years)	Speaker chosen	Valid responses N (%)	Speech act N (% of valid)	Action N (% of valid)	Social evaluation N (% of valid)
1b	6	reliable	14 (82%)	4 (29%)	10 (71%)	4 (29%)
1b	6	unreliable	2 (29%)	0 (0%)	0 (0%)	2 (100%)
2	4	reliable	2 (18%)	0 (0%)	1 (50%)	1 (50%)
2	4	unreliable	2 (15%)	0 (0%)	0 (0%)	2 (100%)
2	5	reliable	8 (50%)	2 (25%)	5 (63%)	3 (38%)
2	5	unreliable	2 (29%)	0 (0%)	1 (50%)	1 (50%)

With respect to children's recollection of the videos, we found that older children had higher scores on the control questions (see Supplementary Table 2). Across ages, children recalled the action more often (78%) than the speech act (53%); however, younger children remembered the promise less frequently than older children (three-year-olds: 25%, four-year-olds: 54%, five-year-olds: 58%, six-year-olds: 83%).

We correlated children's choice scores and their recollection scores. For the speech act category, we found a significant negative correlation for three-year-olds ($\phi = -0.435$, $p = .033$; phi correlations) and a significant positive correlation for six-year-olds ($\phi = 0.451$, $p = .027$); there was no significant correlation for the other two age groups (four-year-olds: $\phi = 0.175$, $p = .391$; five-year-olds: $\phi < .001$, $p > .999$). There were no significant correlations for the action category (3-year-olds: $\phi = 0.251$, $p = .219$; four-year-olds: $\phi = 0.262$, $p = .199$; five-year-olds: $\phi < .001$, $p > .999$; six-year-olds: $\phi = -0.134$, $p = .512$). Three-year-olds, who chose the promise-keeping puppet, did not remember the speech act (and vice versa). Six-year-olds, who chose the promise-keeping puppet, were also better at remembering the speech act. The speech act may have been more salient for six-year-olds than for younger children, resulting in higher choice and recollection scores.

Overall, these findings suggest that by the age of six, children selectively trust promises of reliable speakers. The method we adopted has been widely used in previous studies, that have found an emergence of selective trust in testimony between 3- and 4-years of age (see for instance, Clement et al., 2004). Why did children younger than six years of age fail to select the reliable speaker in our study? One explanation may be that the promises in the videos (i.e., promises to brush teeth or to clean up) differed from the promises uttered to the child (i.e., promises to help

retrieve puzzle pieces). Younger children may have found it difficult to make a transfer between the different promise-contexts when choosing a speaker. In Study 2, we therefore used videos that depicted promises to help.

Study 2

In Study 2, we presented preschoolers with videos in which speakers promised a third party to help with a task (cleaning or watering plants). We used simple promise situations that we expected children to be familiar with. The speakers' promises to a third party in the videos were thus similar to the speakers' promises to the child (first-party situation). We expected that this would lead to an earlier emergence of selective trust in promises of reliable speakers. We therefore tested four- and five-year-olds in this study. We did not include three-year-olds as we did not expect them to show selective trust even with the modified video stimuli.

Method

Participants. Twenty-four 4-year-olds (*M*: 4 years 6 months, *Range*: 4 years 4 months – 4 years 8 months, 12 girls) and twenty-four 5-year-olds (*M*: 5 years 6 months, *Range*: 5 years 4 months – 5 years 7 months, 12 girls) participated in the study. Seven additional children were excluded from the final sample due to experimental error (one 5-year-old) or because they refused to complete the procedure ($N = 6$; four 4-year-olds, two 5-year-olds). Participants were recruited in the same manner as in the previous studies.

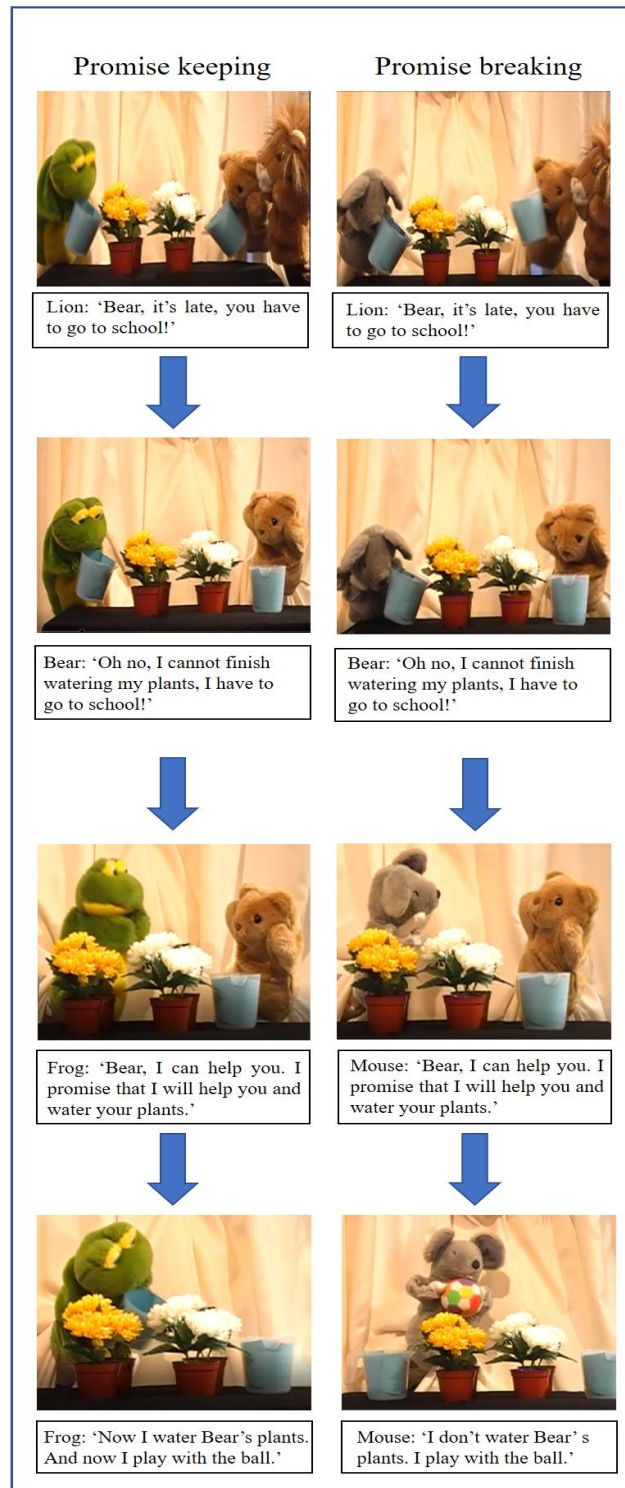


Fig. 4. Screen shots of video stimuli and script used in Study 2 (see the online supplementary material for the videoclips). Sequences were similar to Study 1. The target puppets (frog/mouse) promised to help a third-party (Bear). Then one target puppet kept the promise ('promise-keeping' video), while the other broke it ('promise-breaking' video).

Materials and procedure. Study 2 was identical to Study 1b except that we changed some features of the videos to make the promise situations in the videos more similar to the promise situations children experienced themselves. To achieve this (1) speakers promised another character to help accomplish a task, (2) speakers promised spontaneously (instead of being asked to promise), and (3) speakers always announced whether they were performing the action or not (thus highlighting whether the promise was kept or broken; see Figure 4, for details). Specifically, in trial 1 the speaker (frog/mouse) and a bear were watering plants when a second character (lion) reminded the bear that it was getting late and that it (the bear) had to go to school. The bear then stated that it would not be able to finish the task and the speaker (frog/mouse) spontaneously offered to help (*Bear, I can water your plants. I promise that I will water your plants.*). In trial 2, the speaker and the bear were cleaning up toys when the lion reminded the bear that it was getting late and that he (the bear) needed to catch the bus. Again, the speaker (cow/giraffe) promised to help (*Bear, I can clean up your toys. I promise that I will clean up your toys.*). Once the Bear left the scene, in one video the speaker kept its promise and watered the plants or cleaned up the toys (while saying: *Now I water the Bear's plants. / Now I clean up the Bear's toys.*). In the other video, the speaker broke its promise and did not water the plants or clean up the toys (while saying: *I do not water the Bear's plants. / I do not clean up the Bear's toys.*). Both videos ended with the speaker exiting the scene while playing with a ball (*Now I play with the ball.*). Furthermore, we asked all children to justify their choices.

Data coding and analyses. Data coding and analyses were similar to Study 1b, except that we used Mann-Whitney U-test to test for age effects (as we only compared two age groups). Since

the stories in the videos differed from Study 1, answers to the control questions were coded as follows:

a) *Speech act recollection.* Children recalled the speaker's promise to the third party. For instance, 'Ich gieße deine Blumen. Ich versprech es dir.' (*I water your plants. I promise.*), 'Dass sie hilft seine Blumen zu gießen.' (*That she would help to water his plants.*).

b) *Action recollection.* Children recalled whether the speaker did or did not perform the promised action. For example, 'Die hat die Blumen nicht gegossen.' (*She did not water the plants.*).

Answers to the justification question were coded as follows:

a) *Speech act justification.* Children justified their choice by referring to the speaker's promise to the third party. For instance, 'Na weil die Maus etwas versprochen hat und nicht gehalten hat.' (*Because the mouse has promised something and has not kept it.*).

b) *Action justification.* Children justified their choice by referring to the fact that the reliable speaker performed the promised action, and/or the unreliable speaker did not. For example, 'Weil die Kuh hat nicht aufgeräumt.' (*Because the cow did not clean up.*).

c) *Socio-evaluative justification.* Children justified their choice by evaluating the speakers' behavior or appearance. For example, 'Der Frosch ist ganz gemein.' (*The frog is quite mean.*), 'Weil die Giraffe so nett war.' (*Because the giraffe was so nice.*).

Justification data from one child is missing because the experimenter forgot to ask the justification question.

All data were coded by the same coder as in the previous studies. A second coder scored 25% of children's answers (an equal number of children per age group and gender). Agreement between coders on choice was perfect (Cohen's $\kappa = 1$), on recollection scores was good (speech

act: $\kappa = .824$; action: $\kappa = 1$), and on justifications was perfect to average (speech act: Cohen's $\kappa = 1$; action: Cohen's $\kappa = 1$; socio-evaluation: Cohen's $\kappa = .63$).

Results and discussion

Four-year-olds chose the reliable speaker at chance levels ($M = 1.00$, $SD = 0.66$; $V = 27.5$, $p > .999$), whereas 5-year-old children chose the reliable speaker significantly above chance ($M = 1.46$, $SD = 0.66$; $V = 104$, $p = .007$; see Fig. 1). Five-year-olds chose the reliable speaker significantly more often than 4-year-olds ($W = 183.5$, $p = .021$).

Five-year-olds gave overall more valid justifications for their choices (44%) than 4-year-olds (17%; see Table 1). Collapsing across age groups, children who chose the reliable speaker justified their responses by referring to the speech act (20% of valid responses), the action (60%), and/or social-evaluation (40%). Children who preferred the unreliable speaker justified their choices by referring to the action (25% of valid responses) and/or social evaluation (75%).

With regards to the control questions, more children recalled the actions (88%) than the speech acts (71%), and five-year-olds were better at recalling the promise speech acts (75%) than four-year-olds (67%; see Supplementary Table 2 for further details). There were no significant correlations between choice scores in trial 2 and recollection scores for the speech act category (four-year-olds: $\phi = -0.237$, $p = .246$; five-year-olds: $\phi = 0.053$, $p = .795$; phi-correlations). There was a significant positive correlation between choice scores in trial 2 and the action category for four-year-olds ($\phi = 0.411$, $p = .044$), but not for five-year-olds ($\phi = -0.193$, $p = .344$). Four-year-olds, who chose the promise-keeper, remembered the actions of that puppet more often – maybe because the action was particularly salient for them.

These results suggest that children selectively trust promises by reliable speakers from age five. This is one year earlier than in Study 1b. There are two possible explanations for this earlier emergence. First, we changed the video plots so that they resembled the situation that children experienced themselves (i.e., speakers spontaneously promised to help). This similarity may have helped children to make a transfer from one situation to the other and to better predict the speakers' behaviour. However, we cannot conclusively tell whether five-year-olds in Study 2 relied on a simple behaviour-matching strategy (i.e., expect someone to keep a promise in the same context) or engaged in trait-reasoning (i.e., expect someone to be a promise-keeper irrespective of context). While past research questioned whether preschoolers are able to reason about others' traits (e.g., Ruble & Dweck, 1995, for an overview), recent studies, using simplified procedures, have revealed that preschoolers selectively choose informants based on trait reasoning (e.g., Hermes, Behne, & Rakoczy, 2015; see, Harris, Koenig, Corriveau, & Jaswal, 2018, for an overview). It is thus, in principle, possible that children in our study engaged in trait-reasoning.

Second, in contrast to study 1 the speakers' behaviour in the videos directly helped or harmed a third party. Previous work has shown that children often prefer individuals that help over those that harm (see for instance, Hamlin & Wynn, 2011; Vaish, Carpenter, & Tomasello, 2010), and trust the testimony of benevolent speakers more than that of malevolent ones (Mascaro & Sperber, 2009). Moreover, children judge broken promises in the moral domain (i.e., broken promises that cause harm to recipients) as more severely than promises in the socio-conventional or personal domain (Hussar & Horvath, 2013). It is thus possible that 5-year-olds chose reliable speakers because of their prosocial actions. To investigate whether children used consistency or prosociality as a cue for speaker reliability, we conducted a third study.

Study 3

In this study, we gave children a choice between a speaker that promised to help another character (*promisor*) and a speaker that did not promise (i.e., uttered a neutral sentence; *neutral-speaker*). In two conditions (between-subject design), we varied whether both speakers (i) helped a third party (*prosocial condition*) or (ii) failed to help a third party (*antisocial condition*; see Fig. 4). We made the following predictions: (i) if children's selective trust is based on the speakers' prosocial or antisocial actions, children will show no preference for any speaker in both conditions (because the two speakers perform the same action - helping/not helping); (ii) if children's selective trust is based on the speakers' consistent or inconsistent behaviour, then children will prefer the promisor in the prosocial condition (because it kept its promise), and avoid the promisor in the antisocial condition (because it broke its promise). We tested 5-year-olds because Study 2 showed that children selectively trusted the reliable speaker at this age. We also included 7-year-olds to test whether there would be an age effect.

Method

Participants. Forty-eight 5-year-olds (*M*: 5 years 6 months, *Range*: 5 years 4 months – 5 years 8 months, 23 girls) and forty-eight 7-year-olds (*M*: 7 years 6 months, *Range*: 7 years 4 months – 7 years 8 months, 25 girls) participated in the study. Half of the 5-year-olds (*N* = 24; 11 girls) and 7-year-olds (*N* = 24; 12 girls), respectively, participated in the prosocial condition, and the other half participated in the antisocial condition. Four additional 5-year-old children took part in the warm up but were excluded from the final sample because they refused to complete the procedure. Participants were recruited in the same manner as in the other studies. Data were collected in 2016.

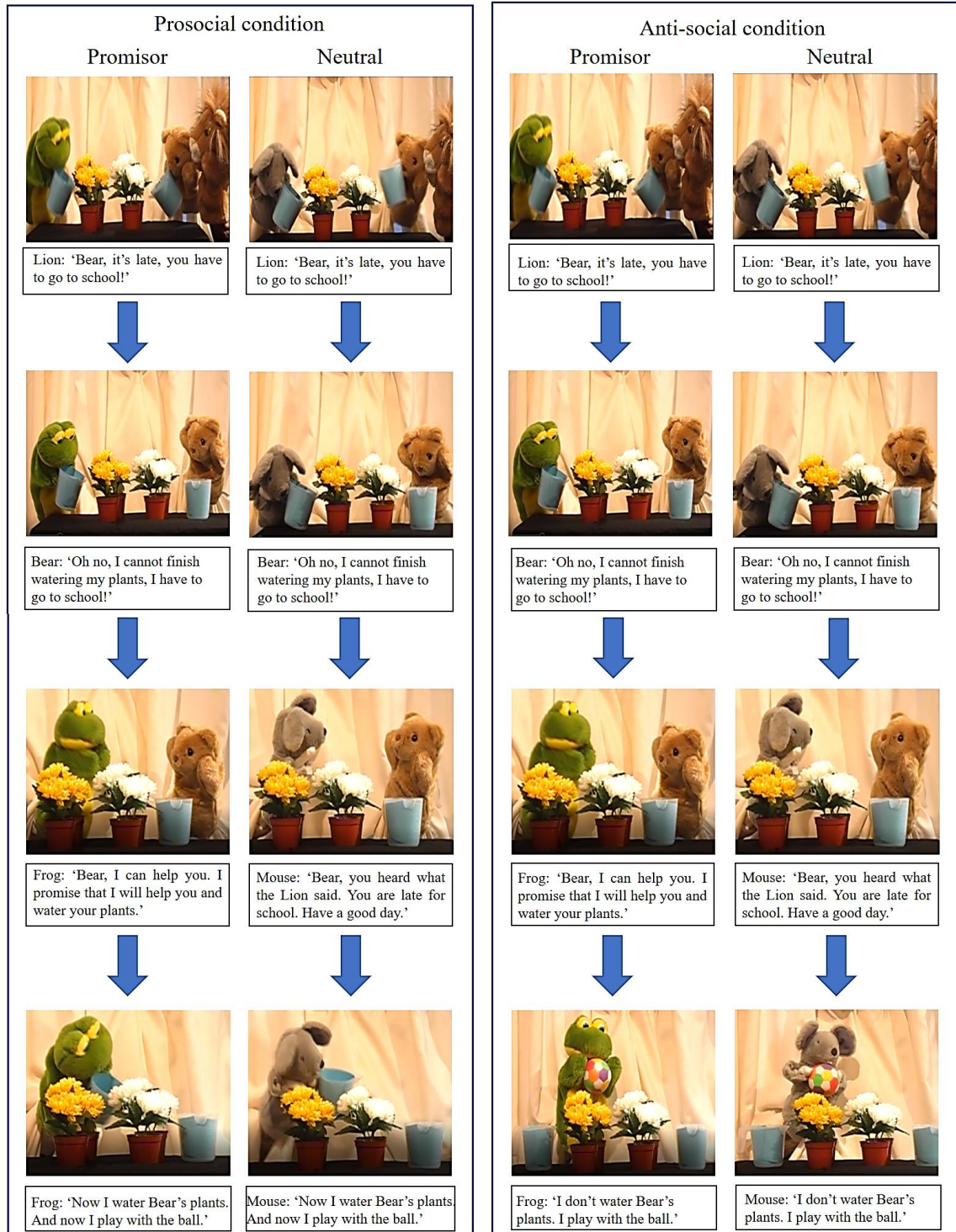


Fig. 5. Screen shots of video stimuli and script used in Study 3 (see the online supplementary material for the videoclips). The promisor-puppet (frog) promised to help a third-party (Bear), while the neutral-puppet (mouse) did not make any promise. In the pro-social condition, both puppets helped the third-party; in the anti-social condition, neither puppet helped.

Materials and procedure. The procedure was the same as in Study 2, with the exception that we modified the video stimuli to show promisors vs. neutral speakers and to show a pro-social or an antisocial outcome (between-subjects). The videos were changed as follows: (1) the promisor promised to help (*Bear, I can clean up your toys / water your plants. I promise that I will clean up your toys / water your plants*) whereas the neutral-speaker only uttered generic sentences (*Bear, you heard what the Lion said. You are late for school / catch the bus. Have a good day / trip!*), and (2) in the prosocial condition, both speakers helped (cleaned up the toys/watered the plants) after the bear had left the scene, whereas in the antisocial condition neither of the speakers helped. See Figure 5 for details of the video scripts.

Data coding and analyses. For each of the two trials, we assigned a score of 1 if the child chose the promisor and a choice score of 0 if the child chose the neutral speaker. The analyses were similar to Study 2: (1) we tested choice scores against chance (chance = 1) using exact Wilcoxon tests, and (2) we tested for the effect of age on choice scores using Mann-Whitney U-test. In addition, we tested the effect of condition on choice scores using Mann-Whitney U-test.

We also coded for children's recollection of the main events in the videos, according to the following categories:

- a) *Speech act recollection.* Children recalled what the promisor or the neutral speaker said to the third party. Examples for the promisor: 'Dass die das Spielzeug aufräumt.' (*That she puts away the toys.*), 'Bär ich verspreche dir, dass ich die Blumen zu Ende gieße.' (*Bear I promise, I will finish watering the plants.*); Examples for the neutral-speaker: 'Es ist schon spät, du musst zur Schule gehen.' (*It is late, you have to go to school.*), 'Gute Fahrt.' (*Have a good trip.*).

b) *Action recollection.* Children recalled whether the speaker performed the action or not. 'Der hat die Blumen gegießt' (*He has watered the plants.*).

Scoring and analyses for recollection were the same as in previous studies.

We coded children's justifications according to the following categories:

a) *Speech act justification.* Children justified their choice by referring to the speaker's promise or neutral speech act to the third party. For instance, 'Die Giraffe hat ja etwas gesagt, aber das war nicht, dass die das aufräumt. Und deswegen musste sie es auch nicht aufräumen. Aber die Kuh hat gesagt, dass die das aufräumt, aber die hats nicht gemacht' (*The giraffe said something, but it was not, that she will clean. And that is why she did not have to clean it. But the cow said that she will clean, but she did not do it.*), 'Weil die eine gute Fahrt gewünscht hat. Das ist schon ein bisschen freundlicher.' (*Because she wished a good trip. This is a bit friendlier.*).

b) *Action justification.* Children justified their choice by referring to the speaker's action. In the prosocial condition, we assigned a score of 1 if children said that the speaker did the action. For instance, 'Weil die Giraffe ihm hilft.' (*Because the giraffe helped him.*). In the antisocial condition, we assigned a score of 1 if they said the speaker did not do the action. For example, 'Weil die Kuh räumt das auch nicht auf.' (*Because the cow does not clean up, too.*).

c) *Socio-evaluative justification.* Same as in the previous studies. For example, 'Naja die Kuh ist eigentlich netter.' (*Well the cow is actually nicer.*).

A coder who had not worked on any of the prior studies scored all of the data. A second coder scored 25% of children's answers (an equal number of children per condition, age group, and gender). Agreement between coders was perfect on choice (Cohen's $\kappa = 1$) and recollection

scores (speech act: $\kappa = 1$; action: $\kappa = 1$), and perfect to average on justifications (speech act: $\kappa = 1$; action: $\kappa = .50$; socio-evaluation: $\kappa = .65$).

Results and discussion

In the prosocial condition (when both speakers helped), 5-year-olds chose the promisor significantly above chance levels ($M = 1.33$, $SD = 0.64$, $V = 65$, $p = .039$). However, 7-year-olds chose at chance levels ($M = 1.25$, $SD = 0.68$, $V = 58.5$, $p = .146$). In the antisocial condition (when neither speaker helped), neither 5-year-olds nor 7-year-olds had a significant preference for one of the speakers (5-year-olds: $M = 0.88$, $SD = 0.74$, $V = 35$, $p = .581$; 7-year-olds: $M = 0.96$, $SD = 0.86$, $V = 72$, $p > .999$). Five-year-olds chose the promisor significantly more often in the prosocial than in the antisocial condition ($W = 385$, $p = .041$). This replicates in a between-subjects design the findings of Study 2, namely, that children prefer a (prosocial) promise-keeper to a (anti-social) promise-breaker. Condition did not have a significant effect for 7-year-olds ($W = 343.5$, $p = .265$). Age did not have an effect on children's choice scores (prosocial condition: $W = 306$, $p = .766$; antisocial condition: $W = 273.5$, $p = .849$).

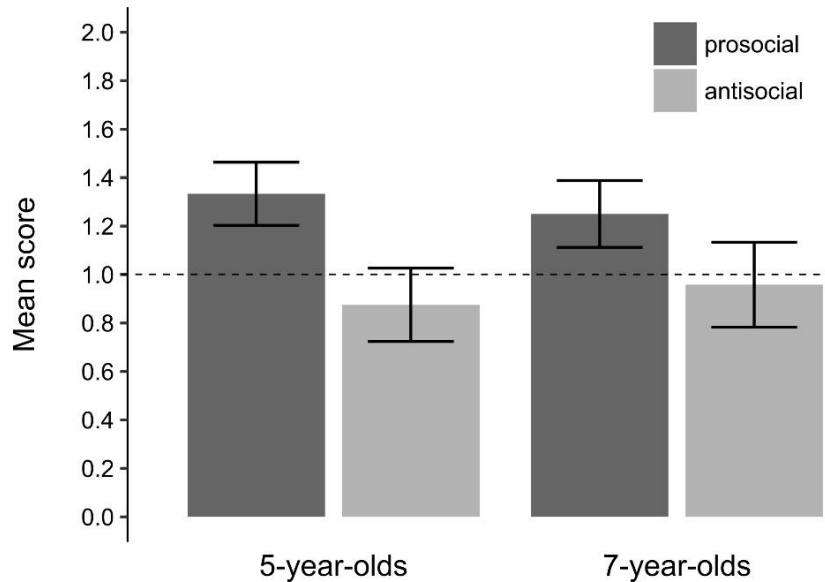


Fig. 6. Mean choice scores (\pm SE) of participants as a function of condition and age in Study 3. Chance (1) is indicated by the dotted line.

Across ages and conditions, few children justified their choices (16% valid justifications; see Supplementary Table 4 for details).

Across ages and conditions more children recalled the action (76%) than the (neutral or promise) speech act (50%; see Supplementary Table 2, for further details). In the prosocial condition, there was a significant positive correlation between choice scores in trial 2 and recollection scores for the action category for five-year-olds ($\phi = 0.426, p = .037$; phi-correlation), but not for seven-year-olds ($\phi = 0.107, p = .600$). Five-year-olds, who chose the promisor, tended to be better at recalling the speaker's actions. One possible explanation for this is that the helping-action was more salient for the promisor than for the neutral speaker. There was no correlation for the speech act category for both age groups (five-year-olds: $\phi = -0.324, p = .113$; seven-year-olds: $\phi = 0.204, p = .318$). In the antisocial condition, there was a significant positive correlation for the speech act category for seven-year-olds ($\phi = 0.438, p = .032$), but not five-year-olds ($\phi = 0.324, p = .113$). Seven-year-olds who chose the promisor were better at recalling the speech act – possibly,

because the promise was more salient than the utterance of the neutral-speaker in the absence of any helping action. There was no significant correlation for the action category for both age groups (five-year-olds: $\phi = -0.175, p = .391$; seven-year-olds: $\phi = -0.048, p = .814$).

We found that 5-year-olds preferred the promisor in the prosocial condition, but did not prefer any speaker in the antisocial condition. This suggests that children's use of speaker-consistency is context dependent. Specifically, speaker consistency seems to matter for positive, prosocial outcomes but not for negative, antisocial outcomes. This finding is in line with previous work showing that positive vs. negative information affects children's reasoning and behavior differently (see, Doebel & Koenig, 2013, for a review). It is possible that children viewed the two non-helpers in the anti-social condition as malevolent and thus as unreliable and untrustworthy (see Mascaro & Sperber, 2009 for evidence that young children use cues about benevolence/malevolence when selecting whom to trust). Consequently, children may have failed to trust either character to keep their promise.

Nevertheless, we found that, in prosocial contexts, children did use consistency in their trust decisions. This is in contrast to previous work suggesting that children are only able to use speaker consistency from 7 years of age (Rotenberg, 1980). We may have found an earlier emergence in our study, because we used a less challenging task (i.e., children chose between two characters, rather than providing trust ratings and justifications).

Seven-year-olds showed no significant preferences in any of the conditions – although the behavior in the pro-social condition was leaning in the direction of what we found for 5-year-olds. To date, we can only speculate why this is the case. One possibility is that 7-year-olds lacked the motivation to choose carefully between the speakers. For instance, for these older and taller children the bag with the missing pieces may have appeared within their reach and they may have

thought they could complete the puzzle without help. Future studies could modify the set-up so that it was more evident that children needed help to retrieve the missing pieces (i.e., the pieces are locked in a box and only the puppets have the keys to the box).

General Discussion

Across three studies, we investigated German children's selective trust in promises. We found that from 6 years of age children selectively trusted speakers who had previously kept a promise (Study 1). When presenting children with promises to help, children from 5 years of age selectively trusted speakers who had kept a promise to help rather than speakers who broke a promise to help (Study 2). In Study 3, we found that 5-year-olds preferred speakers who kept a promise to help to neutral speakers who merely helped, but did not selectively avoid speakers who broke a promise to help (as compared to neutral speakers who failed to help). Taken together, these findings suggest that children from age 5 to 6 selectively trust promises of reliable speakers, using consistency (i.e., keeping promises), prosociality (i.e., helping) or both as reliability cues.

Our findings contribute to the literature on children's understanding of promises. During early and middle childhood, children begin to understand that speakers are responsible for keeping their promises (Kanngiesser et al., 2017; Mant & Perner, 1988), start to keep their own promises (Heyman, Fu, Lin, Qian, & Lee, 2015; Kanngiesser et al., 2017; Lyon & Dorado, 2008; Lyon, Malloy, Quas, & Talwar, 2008; Talwar, Lee, Bala, & Lindsay, 2002), and evaluate others based on their past promise-behaviour (Rotenberg, 1980). We show that by age five to six, children also *use* their evaluations of speaker's promises and behaviour to decide whether to trust new promises by those speakers. While some authors have suggested that preschoolers only focus on outcomes and fail to distinguish between the speech act of promising (e.g., a promise to help) and the

performance of the promised act (e.g., helping) (Astington, 1988b; Kalish & Cornelius, 2007; Rotenberg, 1980), our findings in Study 3 question this claim. Specifically, five-year-olds distinguished between speakers that promised and helped and those that only helped, indicating that, under some conditions, children take both the promise speech act and the action into account. In line with other recent work (Lohse, Gräfenhain, Behne, & Rakoczy, 2014; Rakoczy & Tomasello, 2009), we show that a more sophisticated understanding of speech acts emerges already in the (late) preschool years and thus earlier than previously claimed (Astington 1988a; Astington, 1988b; Rotenberg, 1980).

In our studies, characters intentionally broke (or kept) their promises. Yet, how and why promises are broken may also affect children's selective trust. From six to seven years of age children evaluate speakers who broke their promises or commitments intentionally more negatively than speakers who broke them by accident (Maas & Abbeduto, 2001; Mant & Perner, 1988). A similar relation may hold for children's selective trust: they may trust speakers who broke promises intentionally less than those who broke them by accident. Moreover, Hussar and Horvath (2013) found that six- to ten-year-olds evaluated promise-breaking in the personal or socio-conventional domain as less severe than promise-breaking in the moral domain. Likewise, children's selective trust decisions may vary across these domains. Finally, children from six years of age rated individuals that lie to benefit others as more trustworthy than those that lie to benefit themselves (Fu, Heyman, Chen, Liu, & Lee, 2015). Plausibly, reasons for promise breaking (e.g., prosocial vs. selfish reasons) may factor into children's trust decisions.

Moreover, socialization seems to play a role in the development of trust: Rotenberg (1995) found that mothers' (self-reported) promise-keeping correlated with their children's trust in others. This finding has interesting implications for studies on trust in promises. Conceivably, children,

who have frequently experienced promise-breaking in the family context, may view others as generally unpredictable and show lower levels of selective trust in a paradigm like ours. Children, who have rarely experienced promise-breaking, may be more sensitively attuned to others' reliability and show higher levels of selective trust.

Our findings also contribute to the literature on children's selective trust. There has been extensive work in the context of selective trust in others' testimony, which has found that children selectively trust others from three to four years of age (Harris et al., 2012; Koenig et al., 2004; Mascaro & Morin, 2014; Stephen, Suarez, & Koenig, 2015). This is about two years earlier than we found in our study using a similar paradigm. One explanation for this later developmental onset is that selective trust in promises may be more taxing for younger children than selective trust in assertions (i.e., the speech act commonly used in testimony studies): (1) Selective trust in promises requires children to integrate information about speakers' past behaviour, their present intentions, and expectations about their future behaviour, while selective trust in assertions only requires an integration of past and present information. Only from the late pre-school years onwards are children able to imagine, plan and predict future events (Suddendorf, 2017), which may explain why younger children struggle to selectively trust promises, but not assertions. (2) While assertions express beliefs, promises express intentions. Young children may think that the fulfillment of promises depends solely on speakers' intentions and may be (overly) optimistic about their willingness to keep their promises. In fact, research has shown that pre-schoolers find it difficult to distrust speakers with deceptive intentions (Mascaro & Sperber, 2009; Vanderbilt et al., 2011). (3) Assertions hold universally (even though they might be addressed to a specific hearer), while promises are person-specific – for instance, only the recipient of a promise can release the speaker from their obligation (Searle, 1969). Younger children may not yet understand that speakers'

promises and behaviour to third-parties can contain relevant information for their own trust decisions.

Alternatively, we may have found a later developmental onset due to methodological factors. We presented children only with a single event, in which a speaker behaved unreliably, while previous studies presented three or four events (Koenig & Harris, 2005; Birch et al., 2008). There is some evidence that the number of events matters: 3-year-olds showed selective trust in testimony when they were presented with four events (Pasquini et al., 2007), but not when they were presented with three events (Koenig & Harris, 2007). Moreover, in a study that presented children with a single reliability event, 7- but not 4-year-olds selectively trusted the accurate speaker (Fitneva & Dunfield, 2010). Witnessing multiple reliability events may thus help younger children to selectively trust speakers – irrespective of whether speakers utter assertions or promises. To investigate this possibility, future studies could directly compare assertions and promises and manipulate the number of reliability events.

Children in our study gained information about others' reliability by observing interactions between third parties (in line with studies on trust in testimony). Recently, Kidd, Palmeri, and Aslin (2013) found that young children already respond to other's word keeping/breaking in first party-situations. While first-party situations, like the one used by Kidd and colleagues, offer children socially richer interactions, they have some disadvantages. Specifically, we wanted to avoid that children choose based (i) on arbitrary characteristics of live experimenters (e.g., gender, tone of voice), or (ii) on having previously experienced a rewarding and pleasant interaction with the experimenter.

In Study 3, we found, to our surprise, that seven-year-olds performed at levels that did not differ significantly from chance when (in)consistent speakers were contrasted with neutral

speakers. We suggest that this finding could be due to methodological reasons: seven-year-olds may have thought that they could complete the task by themselves and thus may have been less motivated to identify the reliable speaker. Largely, research on selective trust in testimony has focused on preschoolers (Harris, Koenig, Corriveau, & Jaswal, 2018) and we know comparatively little about the development of selective trust beyond the preschool years. Researchers have studied trust in older children and adolescents using anonymous trust games and have found that trust increases linearly until early adulthood (Sutter & Kocher, 2007; van den Bos, Westenberg, van Dijk, & Crone, 2010) – possibly because adolescents and young adults have more experience interacting with strangers. Future research could investigate whether the development of selective trust in strangers' testimony follows a similar developmental trajectory.

There are a number of limitations to our findings. First, we used a forced choice measure and contrasted two speakers with different histories of promise behaviour (or neutral behaviour; Study 3). Young children show selective trust when presented with the testimony of two different speakers, but struggle to reject the testimony of a single inaccurate speaker (Heyman, Sritanyaratana, Kimberly, & Vanderbilt, 2013; Vanderbilt, Heyman, & Liu, 2014; Vanderbilt, Liu, & Heyman, 2011). Based on these findings, Vanderbilt, Heyman, and Liu (2014) suggested that younger children judge speakers relative to each other and not in absolute terms. Our paradigm does not allow us to tease apart whether children showed *relative* selective trust in speakers' promises or *absolute* selective trust. Future studies could investigate at what age children reject the promise of a single speaker with a history of promise-breaking.

Second, we cannot conclusively tell whether children trusted reliable speakers to *keep their promise to help*, or whether they simply trusted reliable speakers to *help*. As an experimental measure, we always asked children who they would like to receive help from, as we did not think

it (methodologically) feasible to have children spontaneously choose a partner. The explicit helping-question may have directed children's attention to choosing a partner that helps rather than a partner who keeps their promise to help. Future studies could examine this possibility by asking children explicitly whom they *trust to keep their promise*.

Third, the characters in studies 2 and 3 spontaneously promised without an explicit reason for doing so. This may have left children wonder why a promise was uttered at all. Future studies could explore different reasons for promising. For example, one character could express doubt about another's reliability ('How do I know you will really do it? Last time you forgot.') and the speaker could utter a promise as a re-assurance ('I promise, this time I will do it.'). Research with adults has shown that promises re-establish trust in economic games after deceptive or non-cooperative actions (Schweitzer, Hershey, & Bradlow, 2006; Schniter, Sheremeta, & Sznycer, 2013) and it would be worthwhile to investigate whether children would treat promises in a similar way.

In conclusion, we find that children show selective trust in promises from five years of age. Our findings contribute to the literature on children's understanding of promises by showing that late preschoolers, under some conditions, take both speech acts and actions into account. They also contribute to the literature on selective trust, which has almost exclusively focused on assertions, by mapping the developmental trajectory of a different type of speech act: promises.

References

- Astington, J. W. (1988a). Children's understanding of the speech act of promising. *Journal of Child Language*, *15*, 157-173. doi:10.1017/S0305000900012101
- Astington, J. W. (1988b). Promises: words or deeds? *First Language*, *8*, 259-270.
doi:10.1177/014272378800802404
- Austin, J. L. (1975). *How to do things with words*. Cambridge, MA: Harvard University Press.
- Belot, M., Bhaskar, V., & van de Ven, J. (2010). Promises and cooperation: Evidence from a TV game show. *Journal of Economic Behavior & Organization*, *73*, 396-405.
doi:10.1016/j.jebo.2010.01.001
- Belot, M., Bhaskar, V., & Van De Ven, J. (2012). Can observers predict trustworthiness? *Review of Economics and Statistics*, *94*, 246-259. doi:10.1162/REST_a_00146
- Bicchieri, C. (2002). Covenants without swords: Group identity, norms, and communication in social dilemmas. *Rationality and Society*, *14*, 192-228.
doi:10.1177/1043463102014002003
- Birch, S. A., Vauthier, S. A., & Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition*, *107*, 1018-1034.
doi:10.1016/j.cognition.2007.12.008
- Brandom, R. (1994). *Making it explicit*. Cambridge, MA: Harvard University Press.
- Chen, J., & Houser, D. (2017). Promises and lies: can observers detect deception in written messages. *Experimental Economics*, *20*, 396-419. doi:10.1007/s10683-016-9488-x

- Clément, F., Koenig, M., & Harris, P. (2004). The ontogenesis of trust. *Mind & Language, 19*, 360-379. doi:10.1111/j.0268-1064.2004.00263.x
- Corriveau, K., & Harris, P. L. (2009). Choosing your informant: Weighing familiarity and recent accuracy. *Developmental Science, 12*, 426-437. doi:10.1111/j.1467-7687.2008.00792.x
- Corriveau, K. H., Meints, K., & Harris, P. L. (2009). Early tracking of informant accuracy and inaccuracy. *British Journal of Developmental Psychology, 27*, 331-342.
doi:10.13481026151008X310229
- Doebel, S., & Koenig, M. A. (2013). Children's use of moral behavior in selective trust: Discrimination versus learning. *Developmental Psychology, 49*, 462-469.
doi:10.1037/a0031595
- Einav, S., & Robinson, E. J. (2011). When being right is not enough: Four-year-olds distinguish knowledgeable informants from merely accurate informants. *Psychological Science, 22*, 1250-1253. doi:10.1177/0956797611416998
- Fitneva, S. A., & Dunfield, K. A. (2010). Selective information seeking after a single encounter. *Developmental Psychology, 46*, 1380-1384. doi:10.1037/a0019818
- Friedrich, D., & Southwood, N. (2011). Promises and trust. In H. Sheinman (Eds.), *Promises and Agreements: Philosophical Essays*, 277-294.
- Fu, G., Heyman, G. D., Chen, G., Liu, P., & Lee, K. (2015). Children trust people who lie to benefit others. *Journal of Experimental Child Psychology, 129*, 127-139.
- Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development, 26*, 30-39. doi:10.1016/j.cogdev.2010.09.001

- Harris, P. L., Corriveau, K. H., Pasquini, E. S., Koenig, M., Fusaro, M., & Clément, F. (2012). Credulity and the development of selective trust in early childhood. In M. Beran, J. L. Brandl, J. Perner, & J. Proust (Eds.), *Foundations of Metacognition* (pp. 193-210). Oxford, England: Oxford University Press.
doi:10.1093/acprof:oso/9780199646739.001.0001
- Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. *Annual Review of Psychology*, *69*, 251-273. doi: 10.1146/annurev-psych-122216-011710
- Hermes, J., Behne, T., & Rakoczy, H. (2015). The role of trait reasoning in young children's selective trust. *Developmental Psychology*, *51*, 1574-1587. doi: 10.1037/dev0000042
- Heyman, G. D., Fu, G., Lin, J., Qian, M. K., & Lee, K. (2015). Eliciting promises from children reduces cheating. *Journal of Experimental Child Psychology*, *139*, 242-248. doi: 10.1016/j.jecp.2015.04.013
- Heyman, G. D., & Legare, C. H. (2013). Social cognitive development: Learning from others. In Carlston, D. C. (Ed.), *Oxford Handbook of Social Cognition* (pp. 749-766). Oxford, UK: Oxford University Press. doi:10.1037/0012-1649.41.4.636
- Heyman, G. D., Sritanyaratana, L., & Vanderbilt, K. E. (2013). Young children's trust in overtly misleading advice. *Cognitive Science*, *37*, 646-667. doi:10.1111/cogs.12020
- Hussar, K.M. & Horvath, J. (2013). But you promised: Children's judgments of broken promises. *Psychology*, *4*, 1046-1050. doi: 10.4236/psych.2013.412152

- Kalish, C., & Cornelius, R. (2007). What is to be done? Children's ascriptions of conventional obligations. *Child Development, 78*, 859–878. doi:10.1111/j.1467-8624.2007.01037.x
- Kanngiesser, P., Köymen, B., & Tomasello, M. (2017). Young children mostly keep, and expect others to keep, their promises. *Journal of Experimental Child Psychology, 159*, 140-158. doi:10.1016/j.jecp.2017.02.004
- Kidd, C., Palmeri, H., & Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition, 126*, 109-114. doi: 10.1016/j.cognition.2012.08.004
- Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76*, 1261-1277. doi:10.1111/j.1467-8624.2005.00849.x
- Koenig, M. A., & Harris, P. L. (2007). The basis of epistemic trust: Reliable testimony or reliable sources? *Episteme, 4*, 264-284. doi:10.3366/E1742360007000081
- Koenig, M. A., & Woodward, A. L. (2010). 24-month-olds' sensitivity to the prior inaccuracy of the source: possible mechanisms. *Developmental Psychology, 46*, 815-826. doi:10.1037/a0019664
- Krogh-Jespersen, S., & Echols, C. H. (2012). The influence of speaker reliability on first versus second label learning. *Child Development, 83*, 581-590. doi:10.1111/j.1467-8624.2011.01713.x
- Liu, D., Vanderbilt, K. E., & Heyman, G. D. (2013). Selective trust: Children's use of intention and outcome of past testimony. *Developmental Psychology, 49*, 439. doi:10.1037/a0031615

Lohse, K., Gräfenhain, M., Behne, T., & Rakoczy, H. (2014). Young children understand the normative implications of future-directed speech acts. *PloS One*, *9*, e86958.

doi:10.1371/journal.pone.0086958

Lyon, T. D., & Dorado, J. S. (2008). Truth induction in young maltreated children: The effects of oath-taking and reassurance on true and false disclosures. *Child Abuse & Neglect*, *32*,

738–748. doi: 10.1016/j.chiabu.2007.08.008

Lyon, T. D., Malloy, L. C., Quas, J. A., & Talwar, V. A. (2008). Coaching, truth induction, and young maltreated children's false allegations and false denials. *Child Development*, *79*,

914–929. doi: 10.1111/j.1467-8624.2008.01167.x

Maas, F., & Abbeduto, L. (2001). Children's judgments about intentionally and unintentionally broken promises. *Journal of Child Language*, *28*, 517–529. doi:

10.1017/S0305000901004743

Mant, C. M., & Perner, J. (1988). The child's understanding of commitment. *Developmental Psychology*, *24*, 343-351. doi:10.1037/0012-1649.24.3.343

Mascaro, O., & Morin, O. (2014). Gullible's travel: How honest and trustful children become vigilant communicators. In L. Robinson & S. Einav (Eds.), *Trust and skepticism: Children's selective learning from testimony* (pp. 69-83). Psychology Press.

doi:10.4324/9781315849362

doi:10.4324/9781315849362

Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition*, *112*, 367-380.

doi:10.1016/j.cognition.2009.05.012

- Mills, C. M. (2013). Knowing when to doubt: developing a critical stance when learning from others. *Developmental Psychology, 49*, 404-418. doi:10.1037/a0029500
- Orbell, J. M., Van de Kragt, A. J., & Dawes, R. M. (1988). Explaining discussion-induced cooperation. *Journal of Personality and Social Psychology, 54*, 811–819.
doi:10.1037/0022-3514.54.5.811
- Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American Political Science Review, 86*, 404-417.
doi:10.2307/1964229
- Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology, 43*, 1216-1226.
doi:10.1037/0012-1649.43.5.1216
- Rakoczy, H., & Tomasello, M. (2009). Done wrong or said wrong? Young children understand the normative directions of fit of different speech acts. *Cognition, 113*, 205–212.
doi:10.1016/j.cognition.2009.07.013
- Rakoczy, H., Warneken, F., & Tomasello, M. (2009). Young children's selective learning of rule games from reliable and unreliable models. *Cognitive Development, 24*, 61-69.
doi:10.1016/j.cogdev.2008.07.004
- Robinson, E.J., and E. Nurmsoo. 2009. When do children learn from unreliable speakers? *Cognitive Development, 24*, 16–22. doi:10.1016/j.cogdev.2008.08.001
- Rotenberg, K. J. (1980). 'A Promise Kept, a Promise Broken': Developmental Bases of Trust. *Child Development, 51*, 614-617. doi:10.2307/1129307

- Rotenberg, K. J. (1995). The socialisation of trust: Parents' and children's interpersonal trust. *International Journal of Behavioral Development, 18*, 713-726. doi: 10.1177/016502549501800408
- Ruble, D. N., & Dweck, C. S. (1995). Self-perceptions, person conceptions, and their development. In N. Eisenberg (Ed.), *Review of personality and social psychology, 15. Social development* (pp. 109-139). Thousand Oaks, CA, US: Sage Publications, Inc.
- Sally, D. (1995). Conversation and cooperation in social dilemmas: a meta-analysis of experiments from 1958 to 1992. *Rationality and Society, 7*, 58-92. doi:10.1177/1043463195007001004
- Scanlon, T. (1998). *What we owe to each other*. Harvard University Press.
- Schniter E., Sheremeta R.M., Sznycer D. (2013) Building and rebuilding trust with promises and apologies. *Journal of Economic Behavior & Organization, 94*, 242-256. doi:10.1016/j.jebo.2012.09.011
- Schweitzer M.E., Hershey J.C., Bradlow E.T. (2006) Promises and lies: Restoring violated trust. *Organizational Behavior and Human Decision Processes, 101*, 1-19. doi:10.1016/j.obhdp.2006.05.005
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language* (Vol. 626). Cambridge University Press. doi:10.1017/CBO9781139173438
- Searle, J. R., & Vanderveken D. (1985) *Foundations of Illocutionary Logic*. Cambridge: Cambridge University Press.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, *25*, 359-393. doi:10.1111/j.1468-

0017.2010.01394.x

Stephens, E., Suarez, S., & Koenig, M. (2015). Early Testimonial Learning: Monitoring Speech Acts and Speakers. *Advances in Child Development and Behavior*, *48*, 151-183.

doi:10.1016/bs.acdb.2014.11.004

Suddendorf, T. (2017). The emergence of episodic foresight and its consequences. *Child Development Perspectives*, *11*, 191-195. doi: 10.1111/cdep.12233

Sutter, M., & Kocher, M. G. (2007). Trust and trustworthiness across different age groups.

Games and Economic Behavior, *59*, 364-382. doi:10.1016/j.geb.2006.07.006

Talwar, V., Lee, K., Bala, N., & Lindsay, R. C. L. (2002). Children's conceptual knowledge of lying and its relation to their actual behaviors: Implications for court competence examinations. *Law and Human Behavior*, *26*, 395-415. doi: 10.1023/A:1016379104959

Vaish, A., Carpenter, M., & Tomasello, M. (2010). Young children selectively avoid helping people with harmful intentions. *Child Development*, *81*, 1661-1669. doi:10.1111/j.1467-

8624.2010.01500.x

Van den Bos, W., Westenberg, M., van Dijk, E., & Crone, E. A. (2010). Development of trust and reciprocity in adolescence. *Cognitive Development*, *25*, 90-102.

doi:10.1016/j.cogdev.2009.07.004

Vanderbilt, K. E., Heyman, G. D., & Liu, D. (2014). In the absence of conflicting testimony young children trust inaccurate informants. *Developmental Science, 17*, 443-451. doi:10.1111/desc.12134

Vanderbilt, K. E., Liu, D., & Heyman, G. D. (2011). The development of distrust. *Child Development, 82*, 1372-1380. doi:10.1111/j.1467-8624.2011.01629.x

Watson, G. (2004). Asserting and promising. *Philosophical Studies, 117*, 57-77. doi:10.1023/b:phil.0000014525.93335.9e