DOI: 10.1111/cdev.13758

EMPIRICAL ARTICLE

CHILD DEVELOPMENT

How children revise their beliefs in light of reasons

Hanna Schleihauf[1,2,3,4] | **Esther Herrmann**[5] | **Julia Fischer**[2,3,4] | **Jan M. Engelmann**[1]

[1]Department of Psychology, University of California, Berkeley, Berkeley, California, USA

[2]Cognitive Ethology Laboratory, German Primate Center—Leibniz Institute for Primate Research, Göttingen, Germany

[3]Department for Primate Cognition, Georg-August-University Göttingen, Göttingen, Germany

[4]Leibniz ScienceCampus Primate Cognition, Göttingen, Germany

[5]University of Portsmouth, Portsmouth, UK

Correspondence
Hanna Schleihauf, Department of Psychology, University of California, Berkeley, 2121 Berkeley Way, Berkeley, CA 94705, USA.
Email: schleihauf@berkeley.edu

Funding information
We thank the kindergartens, parents, and children that supported our work. This project has received funding from the European Union's Framework Program for Research and Innovation Horizon 2020 (2014–2020) under the Marie Sklowska-Curie Grant Agreement awarded to Hanna Schleihauf. Additionally, Jan Engelmann acknowledges support from the Jacobs Foundation.

Abstract

We investigate how the ability to respond appropriately to reasons provided in discourse develops in young children. In Study 1 ($N = 58$, Germany, 26 girls), 4- and 5-, but not 3-year-old children, differentiated good from bad reasons. In Study 2 ($N = 131$, Germany, 64 girls), 4- and 5-year-old children considered both the strength of evidence for their initial belief and the quality of socially provided reasons for an alternative view when deciding whether to change their minds. Study 3 ($N = 80$, the United States, 42 girls, preregistered) shows that 4- and 5-year-old children also consider meta-reasons (reasons about reasons) in their belief revision. These results suggest that by age 4, children possess key critical thinking capacities for participating in public discourse.

Humans, in contrast to other animals, reason interactively. Social discourse—the exchange of reasons with others—is a powerful tool to circulate reliable information, foster consensus, and increase community connection. Our reality, however, is also characterized by fake news, echo chambers, and increasing polarization. One prerequisite for healthy discourse is that children learn to support their claims with reasons and to evaluate reasons given by others, so that they mature into reasonable contributors to public exchange. Here, we study how children develop reason-responsiveness: the ability to respond appropriately to reasons provided in social discourse (Lord, 2014; Parfit, 2011; Raz, 2011; Schroeder, 2009).

We investigate the development of three foundational abilities of reason-responsiveness. *In Study 1, we focus on children's ability to distinguish good from bad reasons*. A reason is a consideration that counts in favor of a belief (Scanlon, 1998). A good reason is a consideration that makes it very likely that a given belief is true by ruling out alternative possibilities. Correspondingly, a bad reason does not support a given belief or does so only weakly, insofar as it leaves many, or indeed all, alternatives on the table. Suppose you inquire *Why do you believe that Mary is in the library?*, and I reply with the good reason: *Because she left a note saying she was going there*, a consideration that makes alternative beliefs unlikely.

Alternatively, I might answer with the bad reason: *Because I saw her leave the apartment*, a consideration which could lead to many alternative beliefs (she may be at the gym, bank, etc.).

In Studies 2 and 3, we undertake more demanding tests of children's reason-responsiveness: how children revise their beliefs in response to reasons in the context of disagreement; that is, when they have formed an initial belief and are then presented with reasons challenging this belief. Reasons can challenge beliefs in two ways (Pollock, 1987): (a) by supporting an alternative view (so-called rebutting reasons, our focus in Study 2), or (b) by undermining the initial reason for the belief (so-called undercutting or meta-reasons, our focus in Study 3).

In Study 2, we investigate children's ability to consider both the strength of the evidence for their initial belief and the quality of the reasons for an alternative belief provided by a social partner when deciding whether to maintain or revise their belief. Imagine you have the initial belief that your phone is in the car, supported by the reason that you cannot hear it ringing at home. If your partner offers the alternative belief that your phone is in the living room, supported by the reason that they saw it there, you need to weigh your own and your partner's reasons against one another. If you find your partner's reasoning more convincing than your own, you will likely revise your belief and check the living room, otherwise, you will likely maintain your initial belief and check the car. In social discourse, effective response requires fine-grained evaluation of the quality of both personal reasoning and alternative arguments. Adept reason-responsiveness grows from the epistemological matrix formed by such evaluation.

In Study 3, we investigate one of the most abstract abilities involved in reason-responsiveness: whether children consider not only reasons but also meta-reasons in the context of disagreement. Meta-reasons are reasons for or against other reasons and are often involved in reflective conversation (e.g., Kuhn et al., 2013). When we explicitly reason about reasons—deliberating whether reason X is sufficient to believe Y or saying things like *But that is not a good reason to believe Y, because…!*—we produce meta-reasons. When trying to change someone's mind, meta-reasons that prove their initial reasons invalid are particularly powerful. Coming back to the example of the lost phone, imagine if your partner added that not only did they see your phone in the living room, they also remember that you put your phone on silent and would therefore not hear it ringing (a meta-reason possibly invalidating your initial reason, that you can't hear your phone). Generally speaking, an individual should be most likely to change their mind if they are provided with a meta-reason weakening the initial reason for their belief *and* a strong reason for an alternative belief.

Combined, these three abilities, evaluating the quality of reasons, weighing personal and another's reasons against one another, and considering meta-reasons to support discernment, are fundamental to reason-responsiveness, and their mastery is necessary for participation in public discourse.

Prior research

Starting from an early age, children acquire vast amounts of information from testimony. By 2 years of age, children expect testimony to reflect a speaker's knowledge and treat it (if there are no reasons to think otherwise) as a reliable source of information (Galazka et al., 2016). The majority of prior research on children's understanding of testimony has investigated children's trust in the source: *who* information comes from. From around 4 years of age, for example, children selectively trust informants that have proven reliable and accurate in the past (for reviews, see Gelman, 2009; Harris, 2015; Harris & Lane, 2014; Harris et al., 2012, 2018; Koenig & Sabbagh, 2013; Poulin-Dubois & Brosseau-Liard, 2016; Ridge et al., 2018). Reason-responsiveness, in contrast, focuses not on the *who*, but the *what* (so-called trust in the content of a claim; Sperber et al., 2010). To participate in social discourse, children need to learn to evaluate whether claims of disagreeing interlocutors are supported by convincing reasons, independently of speaker reputation.

Several studies point to children having the first essential skill for reason-responsiveness: distinguishing good from bad reasons. Early observational studies suggest that starting around preschool age, children display sensitivity to the quality of reasons provided by their parents (e.g., Jipson et al., 2018; Kuczynski, 1984; Tizard & Hughes, 1984). Experimental work indicates that even 2-year-old children possess the capacity to distinguish strong from weak reasons presented in interpersonal discourse (Castelain et al., 2018). Children at this age were more likely to endorse the opinion of an interlocutor when they were presented with strong reasons (based on novel information) rather than weak reasons (circular reasons) for the interlocutor's view. But, as pointed out by the authors, additional information and the quality of reasoning are to some extent confounded in this case (as in many naturalistic cases). More robust evidence for the ability to differentiate reasons in terms of their relative strength appears in 3-year-old children. Koenig (2012) presented 3-, 4-, and 5-year-old children with two informants who gave different reasons for their opposing beliefs. Even the youngest children in this sample reliably trusted claims based on perceptual access, reliable testimony, and inference, relative to claims grounded in pretense, guessing, and desiring (see also Butler et al., 2018; Corriveau & Kurkul, 2014; Mercier et al., 2014). Children have also been shown to respond to good reasons when they engage in cooperative decision-making with a peer (Köymen et al., 2014, 2016; Mammen et al., 2018). The tendency to selectively respond to reasons in terms of their quality does not seem to be restricted to children growing up in Western cultures. Results from a traditional Maya population (Castelain et al., 2016) and from a Japanese population (Mercier et al., 2017) suggest that reason-responsiveness develops similarly across cultures.

The second essential skill investigated here is children's ability to revise existing beliefs following others' verbal input. Belief revision refers to the *process by which a rational agent changes [their] beliefs about a static world in light of new information* (Peppas et al., 1996, p. 1) and is to be distinguished from belief updating—the *process by which an agent keeps [their] beliefs up to date with an evolving world* (Peppas et al., 1996, p. 1; see also Hansson, 2022). In the context of reason-responsiveness, the development of children's ability to revise their beliefs is of particular interest. When you revise your beliefs (but not when you update your beliefs), you need to resolve a mental conflict: In order to take on the new belief you must accept that your prior belief was wrong, as regularly happens in contexts of disagreement.

Around the age of 2½ years, children *update* their beliefs following testimony and adjust their expectations about the world across different contexts (Ganea & Harris, 2010, 2013; Ganea et al., 2016; Özdemir & Ganea, 2020). By 3 years of age, children also consider a speaker's prior reliability when considering whether or not to update their beliefs (Ganea et al., 2011). What is known about children's willingness to *revise* their beliefs in social contexts, the focus of Study 2? While not focusing on children's responses to reasons offered by a social partner, a small number of previous studies have tested under what circumstances children engage in belief revision. These studies vary in the type of initial belief that participants formed and the type of counterevidence that was presented.

In Hagá and Olson (2017), children were told to pick the chartreuse-colored crayon out of a set of crayons. They were then asked how certain they were of their choice. Next, participants were shown a video of a peer picking a different crayon. The younger children in the sample (4- and 5-year-olds) displayed a counterintuitive pair of epistemic attitudes: they were simultaneously highly confident in their belief of what chartreuse meant and readily revised their view when exposed to a peer who had picked a different crayon. In Robinson et al. (1999), preschoolers guessed the identity of an object hidden in a container. They subsequently revised their beliefs when they received contradicting verbal information from an informed person (who had visual access to the object), but not when they received this information form an ignorant person (who had not seen the object, see also Miosga et al., 2020). Macris and Sobel (2017) asked children to form a hypothesis about which objects are needed to make a machine play music based on probabilistic, inconclusive evidence and then provided children with contradictory evidence (based either on observation or testimony). Forty-nine percent of children revised their hypothesis in response to observed counterevidence; 76% of children changed their mind in response to testimony (keeping message content constant). Ma and Ganea (2009) investigated children's willingness to change their mind following a false testimony statement contradicting what they had observed. Four- and 5-year-olds relied on what they had seen and disregarded the false verbal information, while 3-year-olds relied on their direct observation only when they could verify again what they had seen earlier. Another relevant line of research investigated children's willingness to change their minds in the domain of counterintuitive concepts. For example, young children are more accepting of an informant's counterintuitive claims about the identity of familiar objects (e.g., that a rock is a soap), when the informant explicitly stated that the objects were different from what they appeared to be—which can be considered a good reason to believe the claim (Lane et al., 2014).

Although not focusing on belief revision, there is one final study that contributes important insights: Bridgers et al. (2016) investigated how children integrate conflicting information from observation and testimony when they form beliefs. In their study, when children were exposed to strong (but not conclusive) observational evidence supporting one hypothesis and testimony from a knowledgeable informant that supported an alternative view, half of the children based their decision on the observational evidence while the other half followed the testimony. When the testimony did not come from a knowledgeable informant, most children prioritized the observational evidence. In a second experiment, testimony was contrasted with conclusive observational evidence. In this case, all children based their choice on the observed evidence and ignored the testimony.

To our knowledge, whether preschoolers consider meta-reasons in their belief revision—a third essential skill for reason-responsiveness—has not been investigated in prior research. Köymen et al. (2020) investigated whether children *provide* meta-reasons. In a collaborative reasoning context, dyads of 3- and 5-year-old children were asked to figure out which of two boxes contained an item needed to walk through rain. Both children learned that one box contained an umbrella and one box contained rain boots. One child was given additional information: that the umbrella was broken. If the partner proposed to pick the box with the umbrella (*Let's take this box because it has an umbrella*), 5-year-old, but not 3-year-old, children often used the relevant piece of information to produce a meta-reason (*No, we shouldn't take that box because the umbrella in it is broken*). While this study focused on children's ability to *provide* meta-reasons, it has not yet been studied how children *respond* to meta-reasons.

The current research

Studies 1, 2, and 3 were designed with three goals in mind. The goal of Study 1 was to replicate prior findings showing that young children can differentiate between good and bad reasons. Thus, in Study 1, we explore whether 3-, 4-, and 5-year-old children favor a belief supported by a good reason over an alternative belief supported by a bad reason using a novel paradigm.

Study 2 expands previous findings by presenting the first systematic investigation of how 4- and 5-year-old children integrate the evidence for their initial view and the reasons provided by a disagreeing social partner for an alternative view when deciding whether to revise their beliefs. Note that while Bridgers et al. (2016) investigated children's integration of physical evidence and testimonial counterevidence when forming beliefs, they did not study children's belief revision, which is the focus of Study 2 (and Study 3).

Finally, the goal of Study 3 was to investigate whether children consider not only reasons, but also meta-reasons, when deciding whether or not to revise their beliefs.

STUDY 1

The aim of Study 1 was to determine the age at which children preferentially respond to good over bad reasons. We presented 3-, 4-, and 5-year-old children with two boxes and told them that a reward was hidden in one of them. They were given a good reason for believing the reward was in one of the boxes by one informant and a bad reason for believing it was in the other box by another informant (via video sequences). Aiming to replicate prior research (Koenig, 2012; Mercier et al., 2014), we expected children of all age groups to show a preference for good over bad reasons.

METHOD

Participants

The study was conducted in a medium-sized German city. Fifty-nine participants (27 girls) were recruited from an existing pool of children who had taken part in earlier, thematically non-related studies. The number of participants was based on prior related research (Koenig, 2012). The sample was made up of 22 three-year-old children (age range = 3.26–3.89 years; M_{age} = 3.57 years; 10 girls), 17 four-year-old children (age range = 4.32–4.83 years; M_{age} = 4.6 years; 7 girls) and 20 five-year-old children (age range = 5.27–5.83 years; M_{age} = 5.5 years; 9 girls). Most children participating were White and from mixed socioeconomic backgrounds. All children spoke German fluently. Children in that region typically receive high levels of direct child-centered pedagogy in dyadic or group settings, and are encouraged to make autonomous decisions, as required for the procedure of the current study.

One child (a 3-year-old girl) answered all control questions incorrectly; her data were excluded from the analysis. Four children (3 three-year-old girls and 1 five-year-old boy) did not answer the control questions correctly in at least one of four trials. Thus, for these children, not all trials were analyzed (taken together we excluded 9 trials of these 4 children). Thus, our final analysis included data of 223 trials of 58 children (26 girls).

Ethical statement

The study procedures were approved by the ethics committees at the University of California, Berkeley, USA, and the Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, that supported the data collection of Study 1 and Study 2 (project title Study 1: Argument Consideration, project title Study 2: Belief Revision, project title Study 3: Disconfirmation Consideration). The presented studies were noninvasive and strictly adhered to the legal requirements of the country in which they were conducted. Informed written consent was obtained from all parents and additional verbal consent was obtained from the children who participated in these studies.

Material

Boxes

Children were presented with 4 pairs of boxes (one pair in each trial). Two pairs matched in form but differed in color, and the other two pairs matched in color but differed in form (see Figure 1). Although children were told that only one box of each pair contained a reward, all boxes contained stickers (to ensure that all children would be rewarded at the end).

Video stimuli

Children participated in a total of four trials. In each trial, before choosing one of the two boxes, participants were shown a video sequence. The videos involved two puppets, each of them pointing to one of the two boxes. Each puppet gave a verbal reason for their box choice. One puppet formulated a good reason, the other puppet a bad reason. Across trials, two good reasons (based on eyewitness or testimony) and two bad reasons (based on individual preference or an unrelated description of the box) were presented. For details, please refer to Table 1. Good and bad reasons were combined in 4 different ways (testimony-unrelated description, eyewitness-unrelated description, testimony-preference, eyewitness-preference). In each video, a different combination was presented. Whether children first heard a good or a bad reason was counterbalanced within and between subjects. To account for possible color or form preferences of children, each video sequence existed in

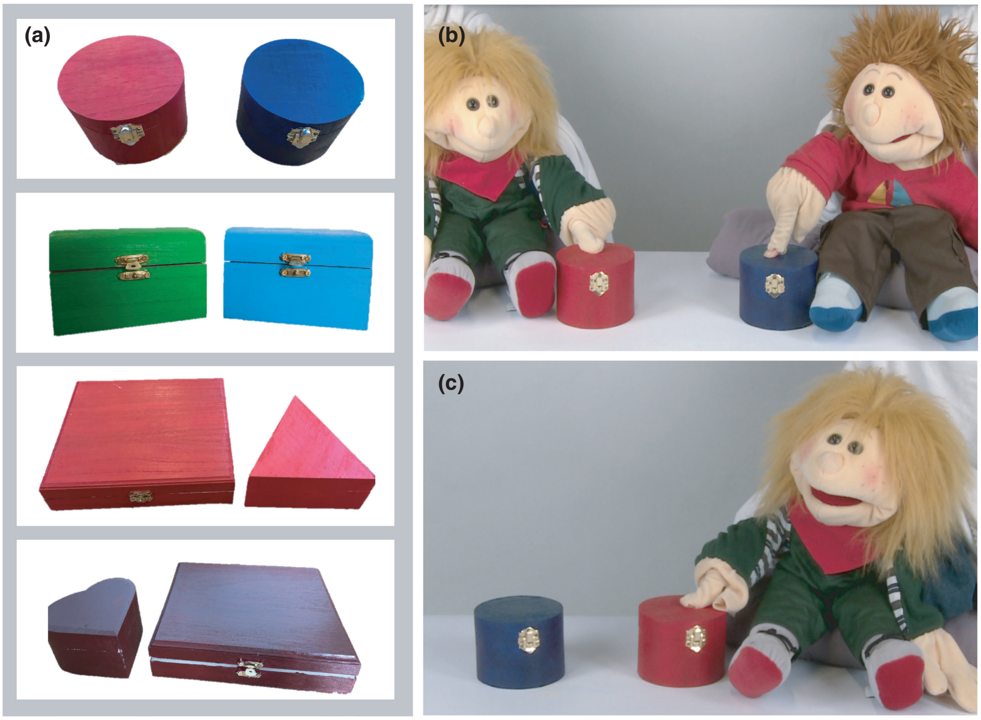

FIGURE 1 (a) In each of 4 test trials children were presented with a pair of boxes differing either in color or form. (b) Screenshot of video stimulus in Study 1. (c) Screenshot of video stimulus in Study 2

TABLE 1 Good and bad reasons presented to children in Study 1 and 2

Good reasons	Eyewitness	I believe that the reward is in the blue box because I looked inside earlier, and I saw that the reward was there. Take the blue box because that is where I have seen the reward
	Testimony	I believe that the reward is in the blue box because my kindergarten teacher told me that it is in that box. Take the blue box because my kindergarten teacher told me that it is in there
Bad reasons	Preference	I believe that the reward is in the red box because red is my favorite color. Take the red box because red is my favorite color
	Unrelated fact	I believe that the reward is in the red box because it is so red. Take the red box because it is so red

two versions, varying in which box color/form was combined with a good or a bad reason (counterbalanced between subjects). Finally, a new pair of puppets and boxes was shown on each trial. This was done to prevent children from thinking that one puppet is more or less reliable based on the previous statement. Stimuli can be found online.

Procedure

The study took place at the children's local daycare centers in a quiet testing room. The child and the experimenter sat down opposite each other at a table. The experimenter explained to the child that they were going to play a game together and that the goal of the game was to find a reward. The experimenter put one pair of boxes on the table and explained the game: *Look, here are two boxes: a red box and a blue box. There is a reward in only one of the two boxes. You can only pick one of them. If you pick the box with the reward, you can keep it. If you pick the box without the reward, you will not get a reward. You may not touch or shake the boxes, but before you make your decision, I will show you a video in which you will get some advice as to which box contains the reward. Once you have selected a box, we will put it aside and at the end of the game we are going to open it to see if there is a reward inside*. The boxes were only opened at the end of the experiment to avoid learning effects. Then, the experimenter showed the child the start screen of the video, and introduced the puppets shown in the video (*Look! These are Luca and Lee. They were here earlier and would like to tell you something.*) Next, the experimenter pointed out that the boxes in the video were the same boxes as those present on the table (*Look! That's the red box* [experimenter pointed to the box in the video]. *It's this one* [experimenter pointed to the box on the table]. *And that's the blue box* [experimenter pointed to the box in the video]. *It's this one right here* [experimenter pointed to the one on the table]). Then the video was started. After the first puppet had given their reason for believing that the reward was in one of the boxes, the experimenter paused the video and asked the child where the puppet thought that the reward was hidden. This control question was introduced to ensure that children had understood what the puppet had said. If children pointed to the wrong box, the video was repeated, and children were asked again. If children did not answer this question correctly after a second repetition, the data of this trial were excluded from analysis (this was the case for a total of 13 trials). Following the control question, the video was continued, and the second puppet presented their reason for believing that the reward was in the other box. After the second control question, the experimenter summarized the puppets' reasons once more and asked the child to pick one of the boxes. This procedure was repeated with all pairs of boxes for a total of four trials. After the 4th trial, children were allowed to open all the chosen boxes and to retrieve their rewards. Children also received a certificate for their participation.

Coding and reliability

Whether or not children picked the box which was associated with a good reason was coded from tape by the first author. We applied a binary coding system. For each trial, in which children picked the box which was associated with a good reason they received a score of 1. For each trial, in which children picked the box which was associated with a bad reason, they received a score of 0. Furthermore, we coded how often the video sequences were repeated during the control questions and excluded the trials during which children still pointed to the wrong box after the second repetition of the video sequence. Additionally, a research assistant, who was blind regarding study design and hypotheses, coded 25% of all trials. According to Cohen's kappa, inter-rater reliability was excellent, $\kappa = 1.00$.

Statistical analysis

To determine the age at which children showed a preference for good over bad reasons, we applied a logistic Generalized Linear Mixed Model fitted via maximum likelihood (Baayen et al., 2008). We used the statistical program R (version 3.4.3; R Core Team, 2019) together with the function *glmer* of the package *lme4* (Bates et al., 2015). We used the package *emmeans* (Lenth et al., 2018) for the post hoc analyses, when necessary.

We included children's *age group* (3-, 4-, or 5-year-olds) as the main predictor and added the factor *reason combination* (eyewitness-unrelated description, eyewitness-preference, testimony-unrelated description, testimony-preference), to control for the possibility that one type of good reason (eyewitness or testimony) was more convincing than one type of bad reason (unrelated description or preference). As a second control predictor, we added the variable *child's gender*. To account for repeated measures and for potential learning effects within one subject (since in each trial a good and a bad reason is presented, it could be easier for children to spot the good reasons from trial to trial), we included the random effect *individual identity* with the random slope of *trial* (z-transformed to avoid convergence issues). To avoid an increased type 1 error risk due to multiple testing, we first tested the overall effect of all test predictors. Therefore, we compared the full model's deviance with that of a null model comprising only the random effect and random slope to examine whether the inclusion of the test predictors provided a better fit to the data than *participant identity* alone. To determine the effects of each predictor alone, we further compared the full model with the corresponding reduced models that lacked the predictor of interest. Please refer to the Supporting Information for more details on the statistical analysis. Finally, we performed post hoc pairwise comparisons (Tukey method) for the significant predictors.

TABLE 2 Study 1: Results of the likelihood ratio test for the comparisons between the full model and the reduced models lacking the predictors of interest

	χ^2	*df*	*p*
Age group	13.409	2	.001*
Reason combination	5.801	3	.122
Child's gender	0.360	1	.548

*$p < .05$.

Results

Children differed in their probability to respond to the good versus bad reason as a function of age. This finding was supported by full-reduced model comparisons that revealed a significant effect for the age group (see Table 2; Figure 2). Four- and 5-year-old children were more likely to choose the box that was supported by good reasons. Three-year-old children, on the other hand, showed no such tendency. They chose the box supported by good reasons with the same probability as they chose the box supported by bad reasons. This conclusion was further based on pairwise comparisons which revealed that the overall effect of the age group was driven by the difference between 3-year-old children and the other age groups (see Table 3). Children's probability to choose the right box was independent of their gender or the combination of reasons that were presented (see Table 2).

Discussion

In Study 1, 4- and 5-year-old children selectively attended to good over bad reasons, whereas 3-year-old children did not show such receptivity. This finding replicates previous results for 4- and 5-year-old children (Koenig, 2012; Mercier et al., 2014). Our result that

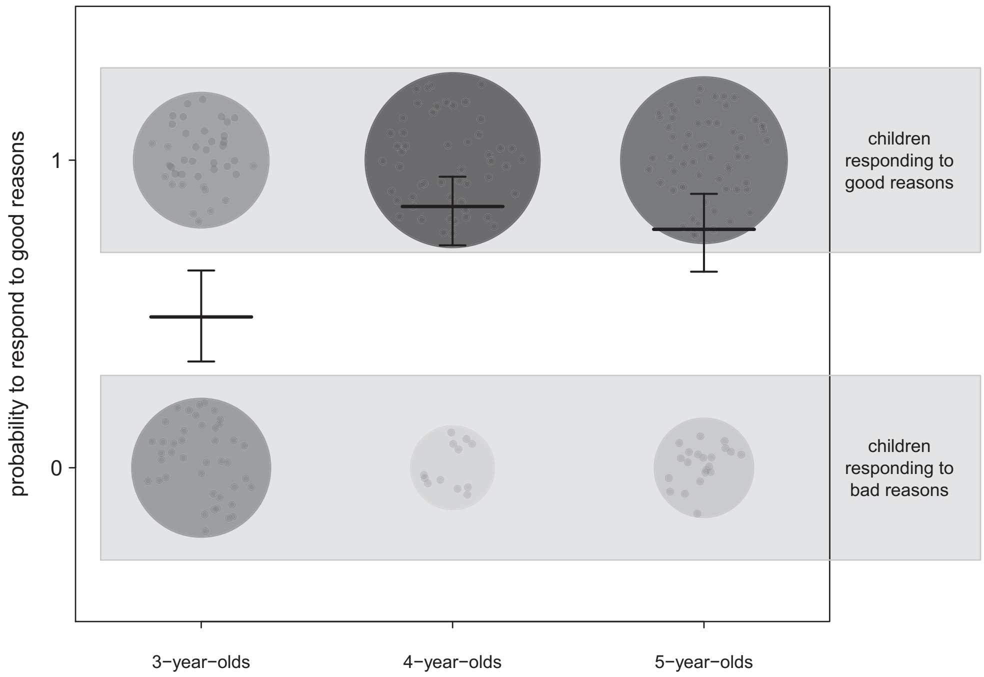

FIGURE 2 Probability to respond to the good versus bad reasons, separated by *age group*. The number of trials on which children responded to the good reason (1) or the bad reason (0) is represented by the size of the large circles as well as by the number of small circles (each small circle represents one trial). Lines represent the point estimates of the Generalized Linear Mixed Model (centered for the factors *argument combinations* and *child's gender*) with the corresponding 95% confidence intervals

TABLE 3 Study 1: Post hoc pairwise comparisons of two age groups at a time with the Tukey method

	p
3-year-olds–4-year-olds	.003*
3-year-olds–5-year-olds	.023*
4-year-olds–5-year-olds	.612

*$p < .05$.

3-year-old children do not differentiate good from bad reasons, however, is not in line with earlier work. For example, Mercier et al. (2014) found that 3-year-old children favored non-circular arguments over circular arguments. Similarly, Koenig (2012) reports that 3-year-old children appropriately judged verbal reasons based on looking, a teacher's testimony, and inference to be better reasons for a belief than pretense, guessing, and desiring. The divergent results for 3-year-old children might be explainable in terms of methodological differences between studies. In contrast to Koenig's study, children in the current set-up were not explicitly asked to judge a reason to be good or bad but to pick one of the two boxes that were linked with these reasons. We do not believe that 3-year-olds' failure to follow good instead of bad reasons can be explained by potential working memory limitations, since children were reminded of the puppets' arguments right before they made their final choice.

Since only the 4- and 5-year-old children considered the quality of reasons in their behavioral decisions, Studies 2 and 3 focused on these two age groups. In Studies 2 and 3, children were presented with a stronger test of reason-responsiveness which investigated whether they appropriately respond to new reasons when these reasons contradict their prior beliefs.

STUDY 2

The goal of Study 2 was to investigate whether children rationally revise their prior beliefs when presented with new reasons by a disagreeing social partner for an alternative view. The study followed a 2 × 2 mixed design, with the between-subjects factor *type of initial belief* (based either on no evidence or strong evidence) and the within-subjects factor *quality of new reasons* (which could be either bad or good). As in Study 1, children had to determine the location of a reward which was hidden in either of two boxes. In contrast to Study 1, children were not provided with conflicting reasons by two different informants. Instead, children first formed an initial belief regarding the reward's location via individual exploration of the two boxes and were then asked to state their belief. Children's initial belief was either based on no evidence (both boxes were identical) or on strong evidence (one box was heavier or made a noise). Then, children were exposed to a disagreeing informant who gave a good or bad reason for the opposing belief (that the reward was in the other box). Subsequently, children were again asked for their belief (i.e., children could decide to stick to their initial belief or to change their mind).

If children consider both the evidence for their prior belief and the reasons for an alternative belief, we expect the following pattern of results. When presented with no evidence for their initial belief and a bad reason for the alternative belief, we expect children's belief revision rates to be at chance level (or slightly above). When presented with no evidence and a strong reason, children should revise their beliefs on most trials. When children have formed an initial belief based on strong evidence and are then confronted with a bad reason for an opposing belief, they should stick to their initial choice. Lastly, when children based their initial belief on strong (but not definite) evidence and were then confronted with strong reasons for a contradicting belief, their belief revision rates should be on chance level.

Method

Participants

The study was conducted in a medium-sized German university town. In total, 147 four- and-five-year-old children (71 girls) participated. The recruitment process was identical to that of Study 1. Participants were 72 four-year-old children (age range = 4.18–4.83 years; M_{age} = 4.54 years; 32 girls) and 75 five-year-old children (age range = 5.2–5.82; M_{age} = 5.52 years; 39 girls). Most children were White and from mixed socioeconomic backgrounds. All children spoke German.

We aimed to test 55 children per condition, with four trials per child. The number of participants per condition was based on prior related research (Macris & Sobel, 2017). However, some children in the *strong evidence* condition did not recognize the evidence in single trials (the boxes had to be lifted or shaken to notice the manipulation, which in some trials children did not do) and, for that reason, did not choose the box which was heavier or rustling. Therefore, in the *strong evidence* condition, we continued data collection until our sample involved 55 children who consistently (in all 4 trials) picked the target box in the first phase of the study. In total, we tested 86 children in the *strong evidence* condition and 56 children in the *no evidence* condition. Additionally, the data of five children had to be excluded due to experimenter error.

To ensure that we only include trials in the *strong evidence* condition, in which children really formed a strong prior belief, we excluded trials in which children either did not pick the correct box or did not give an appropriate justification (for details see Figure S1). After the second repetition of the stimuli, the control question was always answered correctly, therefore no further trials had to be excluded. Using this procedure, we ended up with 224 trials of 56 children in the *no evidence* condition and 228 trials of 75 children in the *strong evidence* condition. Thus, the final sample included the data of 131 children (64 girls).

Importantly, we also ran the analysis without excluding any data points. This revealed identical results, which speaks for the stability of our findings. This additional analysis is reported in Supporting Information.

Material

Boxes

Study 2 used the same 4 pairs of boxes as Study 1. In the *strong evidence* condition, one box of each pair was manipulated so that it was either heavier (iron weights under a double bottom) or rustling (beads under a double bottom). This manipulation functioned as a clue for the location of the reward. All boxes were stuffed with cotton wool to avoid the reward itself making any noise when children shook or lifted the boxes. In the *no evidence* condition, both boxes in a pair were equally heavy and did not make any noise.

Video stimuli

Each child saw four different video sequences, one on each trial. The video sequence showed a puppet pointing to one of the two boxes (see Figure 1c). Crucially, in each trial, a new puppet presented the reasons. This was done to prevent a puppet being perceived as unreliable due to giving reasons that are inconsistent in their quality. Each puppet provided either a good or a bad reason for believing that the reward was in the respective box. In two trials, the puppet gave a good reason (eyewitness, testimony), in the other two trials the puppet gave a bad reason (preference, unrelated description; see Table 1). For each pair of boxes, eight videos were recorded. We recorded two videos for each of the four different reasons; in one video the puppet gave a reason for one box (e.g., the blue box), in the other video the puppet gave a reason for the other box (e.g., the green box). This procedure allowed us to always present the video in which the puppet gave a reason for the box which was not initially chosen by the child. The order in which the videos with these different reasons were presented was counterbalanced between subjects. All stimuli are available online.

Procedure

All children were tested at their daycare centers in a quiet room. The general experimental setup and the introduction were the same as in Study 1. However, in contrast to Study 1, the children in Study 2 were allowed to touch and lift the boxes at the beginning of each trial. Instead of placing the two boxes directly on the table, the experimenter presented the children with a basket which contained the two boxes and asked them to take the boxes out. This procedure was chosen so that children in the strong evidence condition could feel that one of the boxes was heavier or rustling. However, some children turned the basket over and slid the boxes toward them on the surface of the table without lifting or shaking them. Therefore, they did not feel that one of the boxes was heavier or rustling. This was also represented in children's belief about the reward's hiding location and their justification for their beliefs. Thus, such trials were excluded from the analysis (see participant description and Figure S1).

After children removed the boxes from the basket, they were asked where they believed the reward was hidden. In the *strong evidence* condition, where one box of each pair was manipulated to be heavier or rustling, children could form a strong prior belief of where the reward might be hidden. Importantly, children had no information about what kind of reward was hidden, so they could not know that the observed weight or noise was not a good indicator for a hidden sticker. In the *no evidence* condition, where both boxes were equal, children could only guess the location of the reward. After the children had stated their initial belief, the experimenter asked them whether they had a reason for this belief (*Why do you believe this?*). Depending on each child's initial belief about where the reward was hidden, the experimenter chose which video to present. Before the experimenter started the video, she introduced the puppet, and pointed out that the boxes in the video were identical to the boxes on the table (same as in Study 1). In the video, the puppet gave a reason for the box the child did not pick. In contrast to Study 1, children listened to only one reason per trial, which was either good or bad (see Table 1). Following the video sequence, to ensure that the children had understood what the puppet had said, the experimenter asked where the puppet thought that the reward was hidden. In case children failed to answer this control question correctly, the video was repeated up to two times. Next, the experimenter summarized the children's reasons for the chosen box, repeated the puppet's reason for the other box, and asked them to make their final decision. The selected box was then put aside to be opened later. After this procedure had been repeated for all four trials, children could open the boxes and retrieve their reward (all boxes contained a sticker). Children also received a certificate for their participation.

Coding and reliability

We used a binary coding of belief revision. When children revised their prior belief in line with the reason provided by the puppet, they received a score of 1; when children decided to stick with their prior belief, they received a score of 0. Additionally, for the *strong evidence* condition, we coded whether children had initially selected the box for which they had evidence and whether they had justified their belief by stating that one of the boxes was heavier or rustling. As in Study 1, we further coded how often children failed to answer the control questions. All children answered the control question correctly after the second repetition. A second coder, who was blind to study design and hypotheses, coded 25% of all trials. Inter-rater reliability was very high (Cohen's kappa: $\kappa = .95$).

Statistical analysis

To investigate whether children consider the type of initial belief and the quality of new reasons when revising their beliefs, we conducted a logistic Generalized Linear Mixed Model to analyze the data of Study 2. A detailed description of the analyses and the respective assumption tests can be found in Supporting Information. We examined whether the *type of children's initial belief* (based on no evidence vs. strong evidence), the *quality of new reasons* (good reasons vs. bad reasons) contradicting those beliefs, *children's age group* (4- or 5-year-olds) and *children's gender* had an effect on the probability that children revised their prior beliefs. We additionally examined whether possible interactions between *type of initial belief*, *quality of new reasons*, and *age group* influenced children's change of belief.

We added these variables as fixed effects to the model. We also included a random intercept for children's *individual identity*. Initially, we had also included two random slope terms in the model: the random slope of *quality of new reason* within *individual identity* and the random slope of *trial* within *individual identity*. In our final model, however, both random slope terms were removed. The random slope for *trial* within *individual identity* was unidentifiable. Additionally, due to very little within subject variance, our statistical model overestimated the values for the predictors. We aimed to reduce this overestimation by removing the random slope of *quality of new reason*. However, the model estimates were still extreme and disregarding random slopes induces the risk of unreliable estimates and non-generalizable effects (Barr et al., 2013; Schielzeth & Forstmeier, 2009). To nevertheless ensure the reliability of our results, we conducted an additional analysis in which we analyzed only the first trial of each child to avoid the necessity of the random effect structure. Importantly, only effects that were found in both analyses were considered to be reliable. As in Study 1, we performed full-null model and full-reduced model comparisons to determine the effects of our predictors.

Results

Our results show that children consider both the strength of evidence for their initial belief and the quality of reasons for an opposing belief when deciding whether or not to change their minds. This was supported by significant main effects for both manipulated factors *type of initial belief* and *quality of new reason*. Children were more likely to change their mind when there was no evidence for their initial belief. They were also more likely to revise their belief after being exposed to a good reason compared to a bad reason (see Table 4; Figure 3). Crucially, these main effects were also found in the alternative first-trial analysis. Interestingly, when comparing the strength of the statistical effects, whether children's initial belief was based on evidence (factor: type of initial belief) had a stronger impact on their belief revision than the quality of reasons (factor: quality of new reason) for the alternative view. None of the interaction effects reached significance, and there was no effect of *age group*. We also found an unexpected significant effect of child's gender, with boys being more likely to revise their initial beliefs than girls. However, since child's gender was only included as a control predictor and we had no prior hypotheses regarding this effect it will not be interpreted further.

Discussion

The results of Study 2 indicate that 4- and 5-year-old children revise their beliefs in selective ways: they consider both the strength of the evidence for their initial beliefs as well as the quality of reasons provided by a social partner for an alternative view when deciding whether or not to revise their rationale. When children's initial belief was supported by evidence, they were less likely to change their minds than when there was no evidence to base their belief on (i.e., when they guessed). Children's propensity to change their mind was also influenced by the quality of the reasons for the alternative view, with good reasons (*Take the blue box, because that is where I have seen the reward*) prompting more changes of mind than bad reasons (*Take the blue box, because blue is my favorite color*).

To explore children's pattern of belief revision in more detail, we focus on the more reliable estimates of the first-trial analysis (Figure 3, lower plot). When children had no evidence for their initial belief and were presented with a bad reason for the alternative belief, we expected that children revise their beliefs in approximately 50% of trials. One could argue that sticking to one's guns is the appropriate choice in such contexts; after all, the person providing the bad reason also seems to be clueless. On the other hand, changing one's mind, even when doing so is based on a bad reason, might appear less risky. What we found, was that children changed their minds with a 66% probability, which was, in line with our hypotheses, not significantly different from chance level (the confidence interval includes .5).

When children formed a strong initial belief and were subsequently exposed to a good reason for the opposite belief, we also expected children's belief revision rates to be at (or close to) chance level. We assumed the effect of a strong initial belief to balance out the effect of a strong reason for an alternative belief. This was the case. The probability for belief revision was at 47% and not significantly different from chance. Focusing on these two conditions, it seems like children weigh self-perceived evidence and verbal reasons given by a social partner equally when making up their minds.

How can this conclusion be reconciled with the finding that the statistical effect of *type of prior belief* was stronger than the effect of *quality of new reason* (see Table 4)? Let's take a look at children's belief revision rates in the two remaining conditions. When children had no evidence for their initial choice and were subsequently presented with a good reason to change their mind, we expected the highest rates of belief revision. Indeed, the probability for belief revision was at 79%, which was significantly above chance.

TABLE 4 Study 2: Results of the likelihood ratio test for the comparisons between the full model of Study 2 and the reduced models lacking the predictors of interest

	Analysis considering all trials of each child			Analysis considering only the first trial of each child		
	χ^2	*df*	*p*	χ^2	*df*	*p*
Type of initial belief	27.108	1	<.001*	14.219	1	<.001*
Quality of new reason	18.647	1	<.001*	4.951	1	.026*
Age group	0.009	1	0.924	0.001	1	.973
Gender	7.402	1	.007*	4.618	1	.031*
Type of initial belief × quality of new reason	0.530	1	.467	0.032	1	.858
Type of initial belief × age group	1.605	1	.205	1.069	1	.301
Quality of new reason × age group	1.664	1	.197	0.037	1	.848
Type of initial belief × quality of new reason × age group	1.259	1	.262	1.069	1	.301

*$p < .05$.

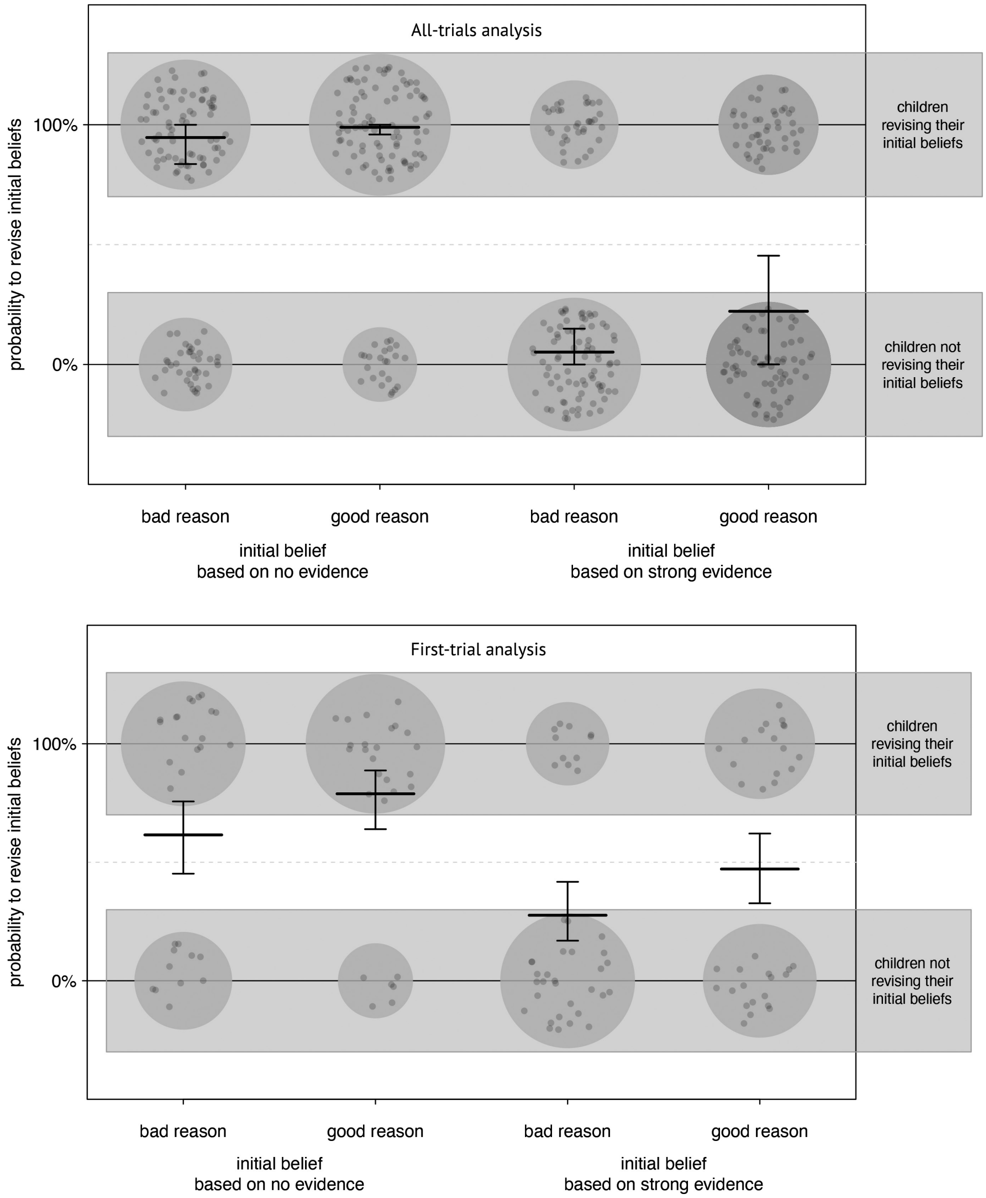

FIGURE 3 Probability of children to revise their beliefs as a function of the integration of the main effects *type of initial belief* and *quality of new reason*. The upper graph is based on the estimates of the all-trial analysis, the lower graph is based on the estimates of the first-trial analysis. The number of trials on which children kept their initial belief (0) or changed their initial belief (1) is represented by the size of the large circles as well as by the number of small circles (each small circle represents one trial). Lines represent the point estimates for the main effects of the Generalized Linear Mixed Model (centered for the factors *child's gender*, and *age group*) with the corresponding 95% confidence intervals that were calculated with parametric bootstraps for the all trial analysis and with the function *confint* of the package *stats* for the first trial analysis

When one's initial belief is based on no evidence, it seems rational to revise this belief when presented with credible reasons for an alternative view. One might wonder why belief revision rates were not even higher. It is possible that some children were driven by instrumental motivations, such as the wish to be right. Having publicly committed to a belief might have made some children less likely to revise that belief later, even in the face of good reasons for an alternative view. However, children were stating their prior belief in all conditions. Thus, such instrumental motivations should be present in all conditions.

When children had strong evidence for their initial belief and were then presented with a weak reason for an alternative view, we expected the lowest rates of belief revision. As predicted, in this scenario, children were least likely to change their mind (27% probability). And in this case, one might question why revision rates were not even lower. One possible explanation is that children responded not only to the content of the message but also to the pragmatics of the interaction with the experimenter. Children might have assumed that the experimenter showed the puppet's statement to them for a reason, for example, because their initial statement might have been incorrect. While such pragmatic considerations likely influenced children's epistemic practices, this does not challenge our main finding in this context, since such considerations would have been present in all conditions. A more likely explanation for the smaller effect size of *quality of new reason* is that the need to justify their beliefs (the experimenter asked children for a justification, see methods) led to an increase in confidence in their views in the *strong evidence* condition and to a decrease in the *no evidence* condition. Lastly, differentiating between good and bad reasons might have been particularly difficult for children in Study 2. While they were presented with two contrasting reasons in Study 1 (as in all previous studies on children's ability to differentiate good from bad reasons), in Study 2 they were presented with only one reason per trial, making it harder to categorize them as good or bad.

STUDY 3

Studies 1 and 2 show that children respond to reasons in appropriate ways when making up their minds. In Study 3 (preregistered), we asked whether children additionally consider meta-reasons in their belief revision. As in Study 2, children were first presented with evidence for an initial belief, and were asked to state their belief and were then, second, presented with good reasons for an alternative belief. In contrast to Study 2, children were also given meta-reasons that either confirmed or disconfirmed their initial reasons. In a within-subject design, children participated in three trials in each condition (confirming condition and disconfirming condition). We predicted that children show sensitivity to meta-reasons, revising their beliefs more often when their initial reason was confirmed by a meta-reason compared to when it was disconfirmed by a meta-reason.

Method

Participants

The study was conducted online with children living in the Bay Area of the United States. Eighty-two participants (42 girls) were recruited from an existing database of children whose parents confirmed their interest to participate in developmental research. Participants were 41 four-year-old children (age range = 4.02–4.99 years; M_{age} = 4.51 years; 22 girls) and 41 five-year-old children (age range = 5.01–5.99; M_{age} = 5.41 years; 20 girls). The number of participants was based on a power simulation expecting a 90% probability for belief change in the disconfirming condition and a 20% probability for belief change in the conforming condition. Across different simulated random slopes and random effects, this led to an average power of $1 - \beta = .75$ (for details see preregistration). Most children were Asian American or White and from middle to high socioeconomic backgrounds. All children spoke English fluently. Parents in the Bay Area typically emphasize their children's psychological autonomy from an early age. In the year the study was conducted the majority of the children participated in preschool online (due to the COVID19 pandemic). Children were thus used to screen-based interactions with adults, as required for the procedure of the current study. Data of two children (5-year-old boys) had to be excluded completely due to their parents' interference during the study. For 21 of the children, single trials had to be excluded from the analysis; in total 27 trials. In 26 of these trials, children did not choose the intended hiding location or did not detect the evidence. In one trial, the child refused to choose one of the hiding locations. Thus, our final sample included 453 trials of 80 children (42 girls), with 227 trials in the *confirming* condition and 226 trials in the *disconfirming* condition.

Material

Stimuli

Children were presented with picture-book like stories presented on PowerPoint slides (Figure 4). Children saw a total of six stories, three stories in the *confirming* condition and three stories in the *disconfirming* condition, presented in alternating order. Each story started with the introduction of two agents (always of the same gender), whose pet had run away. On the next slide, children saw two possible hiding locations (e.g., a bush with red berries and a bush with purple berries) and evidence leading to one of them (e.g., footprints). While this slide was shown, children were asked to state where they thought the animal was hiding and why. When children stated the depicted evidence as a reason for their belief, the next slide was presented. On this slide, one of the agents confirmed or disconfirmed the child's initial reason with a meta-reason. In the *confirming* condition, the agent verbally confirmed that the detected evidence is a good reason for the drawn conclusion (e.g., *These footprints look like bird footprints. Look bird footprints look like this [showing a picture of bird footprints] and these footprints here look just like that*). In the *disconfirming* condition, the agent stated that the observed evidence is not a good reason to draw a conclusion about the animals hiding location (e.g., *These footprints don't look like bird footprints. Look bird footprints look like this [showing a picture of bird footprints] and these footprints here don't look like that*). The type of evidence varied between trials but was kept constant for both conditions. The evidence was presented in the form of footprints, lost objects (e.g., a duck feather of either the same [confirming condition] or a different color [disconfirming condition] as the lost duck), and an animal's body part (e.g., a mouse's tail which was the same [confirming condition] or a different color [disconfirming condition] as the lost mouse). Once children's reason was either confirmed or disconfirmed, the second agent appeared on

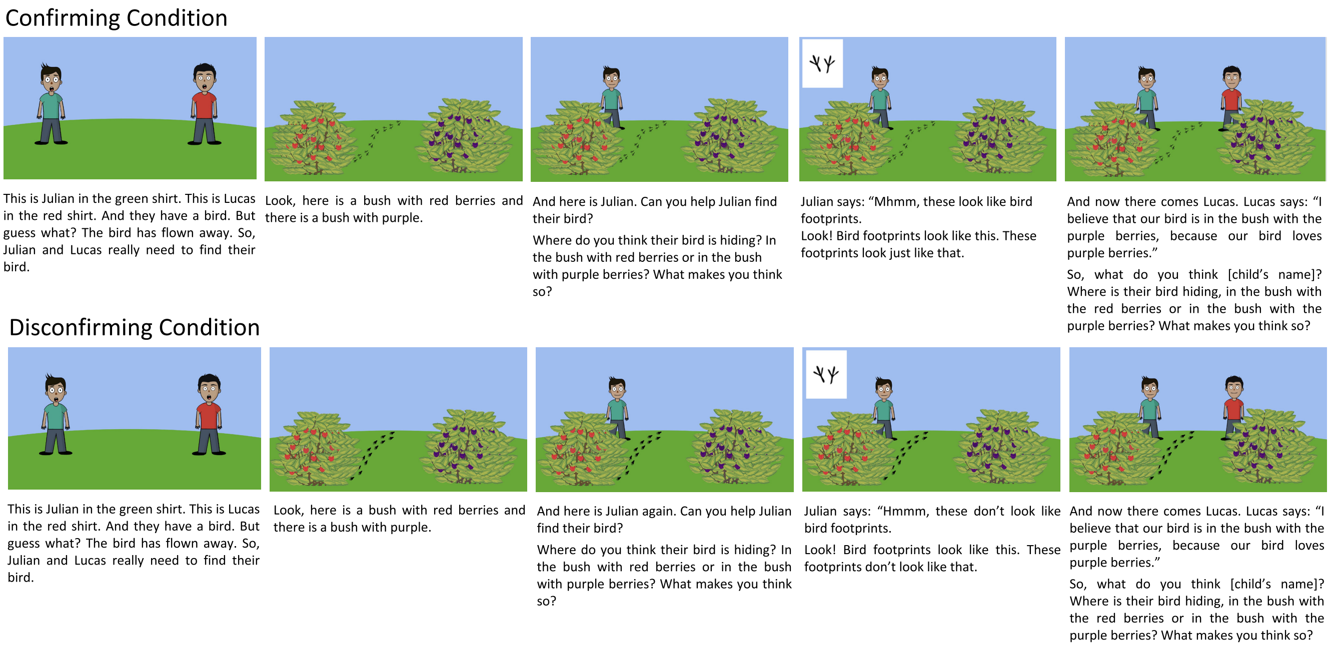

FIGURE 4 Example of the bird story in the *confirming* condition and the *disconfirming* condition of Study 3

the next slide and gave a good reason to believe the animal was hiding in the respective other location. The good reasons were based on the animals' preferences (e.g., *I believe the bird is hiding in the bush with the purple berries because our bird loves purple berries*). Subsequently, children could make their final choice (they were asked again where they thought the animal was hiding). In each of the six stories a different animal was lost, a different pair of agents gave reasons, and different evidence pointed to one of the hiding locations. We used two versions of the stimuli and counterbalanced which story was presented in which condition between children. This way, a particular story (e.g., the bird story) was shown for half of the children in the *confirming*, and for the other half in the *disconfirming* condition. Since the stories were presented in a fixed order, half of the children started with the *confirming*, the other half with the *disconfirming* condition. Within-subjects we counterbalanced the side (left or right hiding location) of the presented evidence, as well as the gender of the agents. All stimuli are available online.

Procedure

All children were tested online over the video-communication software Zoom. After welcoming the parent and the child, the experimenter guided the parent through a calibration process, which ensured that they were seeing the demonstration in full-screen mode, that they could see the experimenter's but not their own video, and that the experimenter's video was floating on top of the slides. Before the experiment started, a short warm-up game was played, in which the child was reminded that animals can leave traces or lose objects. On the first warm-up trial, the child saw frog footprints on the first slide, and then on the second slide a tractor and a frog. The child was then asked who they thought had left the tracks. On the second trial, the child saw a chewy bone on the first slide, and then a slide that depicted a dog and a bunny. The child was then asked who they thought had left the chewy bone. All children answered these questions correctly.

Following this warm-up game, the experiment started. Children saw six stories in total (as described above). In each story, children were asked to state their initial belief and their reason for this belief. If children chose the cued hiding location but did not mention the evidence, the experimenter asked them whether there was anything in the picture that made them think that the animal was hiding in the chosen location. If children did not mention the evidence after this prompt, or if children chose the wrong hiding location, the trial ended and was excluded from the analysis (this was the case for 26 trials, which is approx. 5% of all trials). Once children had chosen the correct location and had referred to the evidence as a reason for their choice, their reason was confirmed or disconfirmed by one of the agents in the story. Then, the second agent gave a good reason to believe the animal was hiding in the other location. To measure belief revision, children were asked again where they believed the animal was hiding and why. At the end of each trial, the experimenter told them that after the game was complete, they would find out whether they were right. The experimenter then proceeded to the next story. After the 6th story, children were presented with a slide in which all agents were depicted with their pets and children were told that they were always right and helped all agents find their pets. After the study, parents and children were given a short summary of the purpose of the study and could ask questions. Children received a certificate for their participation.

Coding and reliability

We used a binary coding of belief revision. When children changed their initial belief, they received a score of 1; when children decided to stick with their initial choice, they received a score of 0. As in Study 1 and 2, a second coder, who was blind to the hypotheses of the study, coded 25% of all trials. Inter-rater reliability was very high (Cohen's kappa: $\kappa = .96$).

Statistical analysis

Again, we used a logistic Generalized Linear Mixed Model to analyze the data of Study 3 (details and assumption tests can be found in Supporting Information). We were interested in whether 4- and 5-year-old children are sensitive to meta-reasons that confirm or disconfirm their initial reasoning, and whether such meta-reasons make it less or more likely that children subsequently change their mind in light of a good reason for an alternative belief. Thus, we examined whether the *valence of the meta-reason* (whether it was confirming or disconfirming the initial reason), *children's age group* (4- or 5-year-olds), and *children's gender* influenced children's belief revision. Additionally, we tested for a possible interaction effect between *valence of the meta-reason* and *age group*. Besides these fixed effects, we added a random intercept for children's *individual identity* with the random slopes for *condition* and *trial number*. However, as in Study 2, we had little within-condition variance per subject, which potentially led to an overestimation of the random slopes and therefore to unreliable estimates. Thus, following our preregistered analysis, we tried to reduce this potential overestimation by following the same procedure as in Study 2 and kept only the random intercept of *individual identity* in the random effect structure. As in Study 2, being aware of the risk that is associated with disregarding random slopes (Barr et al., 2013; Schielzeth & Forstmeier, 2009), we conducted an additional analysis in which we analyzed only the first trial of each child to avoid the necessity of the random effect structure. Again, we performed full-null model and full-reduced model comparisons to determine the effects of our predictors.

Results

We found significant main effects for the factor *valence of meta-reason* (Figure 5; Table 5). When the reason that children based their initial belief on was disconfirmed by a meta-reason, children were more likely to change their mind. In contrast, when the reason for their initial belief was confirmed by a meta-reason, children were less likely to change their mind.

Discussion

The results of Study 3 indicate that children consider meta-reasons when deciding whether to revise their beliefs. When children formed an initial belief based on a reason and were then confronted not only with a strong reason for an alternative belief, but also a meta-reason that spoke against their initial reason, they changed their minds on most trials. In contrast, when the meta-reason supported the children's initial reason, they mostly maintained their initial belief.

Given that children in Study 3 received strong evidence for an initial belief followed by a strong reason for an alternative belief, one might compare children's pattern of belief revision in Study 3 with their pattern in Study 2 in the analogous condition (strong evidence for initial belief followed by a good reason for the alternative belief). In Study 2, children's belief revision probability was at chance. In Study 3, an additional disconfirming meta-reason increased the probability that children changed their minds to 87%, which was significantly above chance (based on the first-trial analysis).

It is important to highlight that the meta-reasons offered in Study 3 did not directly address children's initial beliefs. For example, the social partner did not say: *I don't think the bird is hiding behind this bush, because these footprints don't look like bird footprints*. Otherwise, the pattern of results might simply be explainable in terms of an agent either disagreeing (disconfirming condition) or agreeing (confirming condition) with the children. Instead, the reasons offered by the partner were true meta-reasons: reasons for or against the reasons produced by children in the first step. That is, the social partner simply said: *These footprints don't look like bird footprints*. Thus, by the age of 4, children seem to understand how reasons and beliefs are interconnected.

GENERAL DISCUSSION

The current studies investigated the development of a psychological capacity essential for participation in rational discourse: reason-responsiveness. Across three studies, we tested children's ability to respond appropriately to reasons provided by a social partner. In Study 1, we investigated children's ability to distinguish strong from weak reasons. We found that by the age of 4 (but not at age 3), children reliably preferred views supported by good instead of bad reasons. Studies 2 and 3 presented children with a more challenging test: a scenario where their initial views were contradicted by a disagreeing partner. In Study 2, we tested children's willingness to revise existing beliefs in light of reasons for an alternative view. We found that

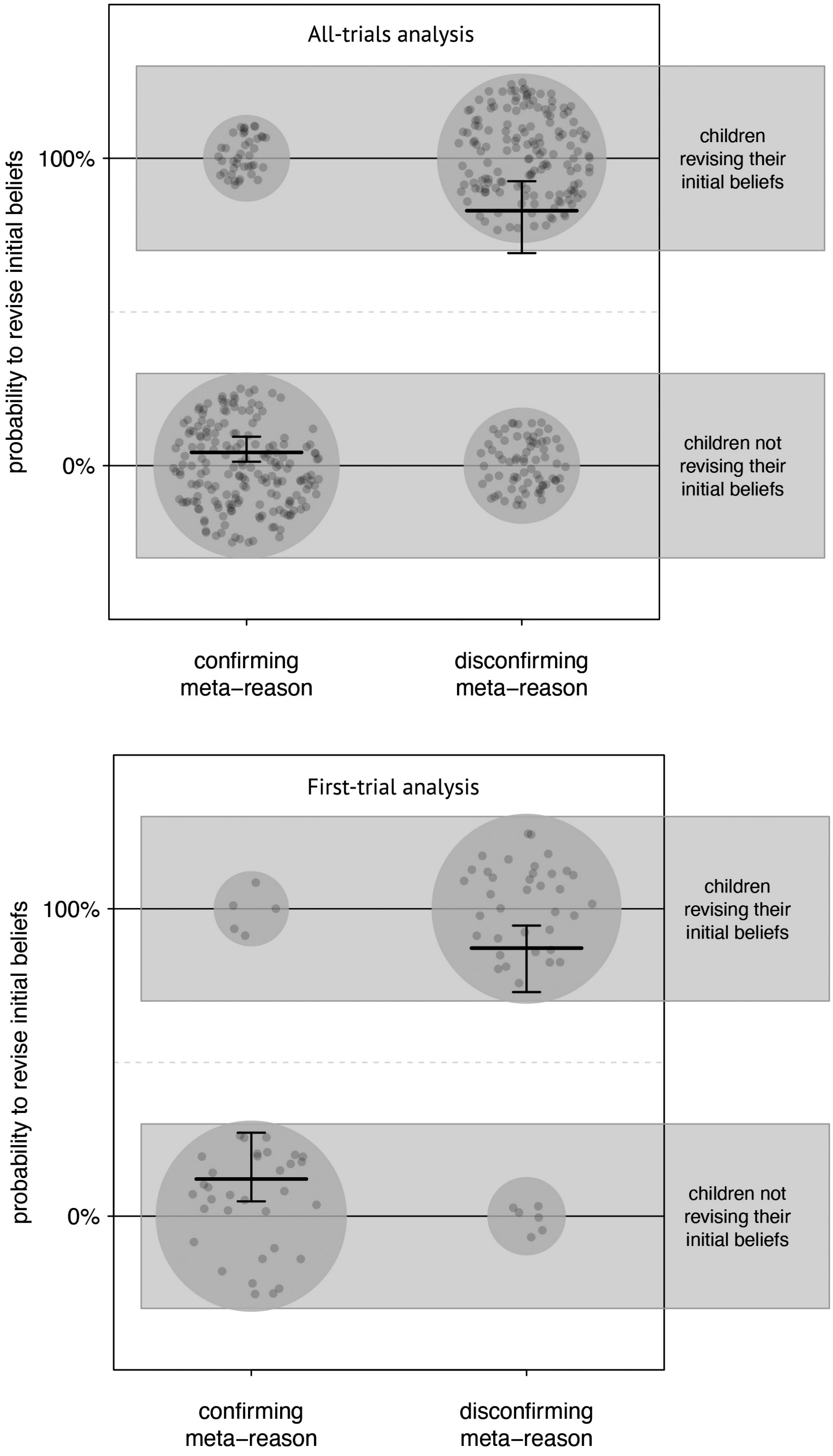

FIGURE 5 Probability of children to revise their beliefs as a function of the main effect *valence of meta-reason*. The number of trials on which children kept their initial belief (0) or changed their initial belief (1) is represented by the size of the large circles as well as by the number of small circles (each small circle represents one trial). Lines represent the point estimates for the main effects of the Generalized Linear Mixed Model (centered for the factors *child's gender*, and *age group*) with the corresponding 95% confidence intervals that were calculated with parametric bootstraps for the all trial analysis and with the function *confint* of the package *stats* for the first trial analysis

TABLE 5 Study 3: Results of the likelihood ratio test for the comparisons between the full model of Study 3 and the reduced models lacking the predictors of interest

	Analysis considering all trials of each child			Analysis considering only the first trial of each child		
	χ^2	*df*	*p*	χ^2	*df*	*p*
Valence of meta-reason	190.386	1	<.001*	47.798	1	<.001*
Age group	1.721	1	.190	0.422	1	.516
Child's gender	1.522	1	.217	0.027	1	.868
Valence of meta-reason × age group	0.518	1	.472	2.719	1	.099

*$p < .05$.

4- and 5-year-old children consider both the strength of the evidence for their initial belief and the reasons given to them by social partners for the alternative view when making up their minds. In Study 3, we tested whether children attend not only to reasons for beliefs, but also to so-called meta-reasons—reasons that speak for or against reasons—in their revision decisions. We found that children appropriately respond to meta-reasons: 4- and 5-year-old children were more likely to revise their initial beliefs when they were provided with meta-reasons contradicting their initial reasons than when the meta-reasons supported their initial reasons. These findings expand previous results showing that in addition to discerning the strength of reasons, children are able to apply this evaluation when deciding whether to maintain or revise their own beliefs, and can incorporate meta-reasons into their decision-making.

While the goal of Study 1 was to replicate prior research, Study 2 expands previous findings on belief revision. Earlier research suggests that 4- and 5-year-old children, following a guess, readily change their mind when presented with a disagreeing partner—even though they are confident in their choice (Hagá & Olson, 2017). In contrast, when 4- and 5-year-olds have conclusive evidence for their belief, they are unlikely to change it when presented with conflicting testimony (Ma & Ganea, 2009). Using a 2 × 2 design in Study 2 allowed us to paint a more nuanced picture of children's belief revision: 4- and 5-year-olds appropriately revise existing beliefs by considering both the strength of evidence for their prior belief and the quality of reasons supporting an alternative belief. This also extends prior research showing that children integrate the strength of observational evidence and the quality of testimonial counterevidence when *forming beliefs* (Bridgers et al., 2016) to the domain of *belief revision*. Revising existing beliefs in light of good reasons offered by a disagreeing social partner is an especially important skill for participating in rational discourse. Thus, in Study 2, children were asked to state their belief, and were then presented with counterevidence. Importantly, once a belief is publicly expressed, not only epistemic motivations (*I hold a belief because I have strong evidence for it*) but also instrumental motivations (*I hold a belief because it is beneficial for me*; see, e.g., Kahan et al., 2011; Kelly, 2002; Over et al., 2017) influence what we believe—making belief revision fundamentally different from belief formation. Wanting to be right, to win an argument, or to signal group membership via one's beliefs are all examples of instrumental motivations that might lead otherwise reasonable agents on irrational paths. While our results cannot rule out the existence of such motives for some individuals, our overall findings represent a belief revision pattern that is in line with epistemically rational practices. The extent to which children's belief formation and revision practices are influenced by instrumental motivations represents an exciting direction for future research.

Study 3 is the first investigation to show that children are sensitive to meta-reasons: reasons that speak for or against other reasons. Children revised their beliefs when they learned that the reasons supporting those beliefs were invalid, and they maintained their beliefs when they learned that their underlying reasons were valid. These results indicate that children understood that their belief was based on a certain reason, and that their belief was not supported any more once the initial reason was disconfirmed by a meta-reason (or that it was still supported if the initial reason was confirmed by a meta-reason). The emergence of this capacity represents a crucial step in the development of rational reasoning skills: children, by age 4, understand how reasons support beliefs. The ability to represent beliefs and the underlying reasons for those beliefs allows children to engage in one of the most fundamental forms of critical thinking, namely to explicitly evaluate how well presumed reasons count in favor of a given belief.

In the current studies, 4-year-olds, but not 3-year-olds, distinguished between strong and weak reasons (in Study 1 and Study 2) and between confirming and disconfirming meta-reasons (in Study 3). What experiences drive the development of such rational reason-responsiveness in young children? One hypothesis is that children may learn skills of reason-responsiveness by engaging in discourse, particularly discourse that involves disagreement (Heyes, 2018; Köymen & Tomasello, 2020; O'Madagain, 2019; Tomasello, 2019; Vygotsky, 1978). When children are confronted with different views—much like in Study 2 and 3—they are naturally prompted to weigh reasons against each other and to ask how strongly different reasons support contrasting views. It is thus possible that 4-year-olds in the current study distinguished between strong and weak reasons, while 3-year-olds did not, because 4-year-olds have had more experience with reason-based discourse. There is tentative support for this view: 3-year-olds become more competent at identifying and producing meta-reasons after a short discourse-based training session in which they are presented with a disagreeing partner who provides reasons for their view (Köymen et al., 2020).

Taken together, our results suggest that by 4 years of age, children have developed reason-responsiveness: they respond appropriately to reasons and meta-reasons provided in social discourse. These findings highlight the importance of considering reasoning not only as an individual process, but also as a social activity. Children's beliefs are influenced and shaped by engaging with others in the practice of giving and asking for reasons.

ACKNOWLEDGMENTS

We thank Katharina Haberl, Susanne Mauritz, Jana Jurkat, Emily Chau, and Alissa Gomez for their great support during data collection. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors have no competing interests to declare.

DATA AVAILABILITY STATEMENT

The data and stimuli for all three studies are made publicly available on OSF:
Study 1: https://osf.io/wkmje/?view_only=378bc468cbe9483db64e630a9e305eac
Study 2: https://osf.io/tsbfr/?view_only=378bc468cbe9483db64e630a9e305eac
Study 3: https://osf.io/vbqrp/?view_only=378bc468cbe9483db64e630a9e305eac
Link to preregistration for Study 3: https://osf.io/hjcmk.

ORCID

Hanna Schleihauf https://orcid.org/0000-0003-4344-9070

REFERENCES

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bridgers, S., Buchsbaum, D., Seiver, E., Griffiths, T. L., & Gopnik, A. (2016). Children's causal inferences from conflicting testimony and observations. *Developmental Psychology*, *52*, 9–18. https://doi.org/10.1037/a0039830

Butler, L. P., Schmidt, M. F. H., Tavassolie, N. S., & Gibbs, H. M. (2018). Children's evaluation of verified and unverified claims. *Journal of Experimental Child Psychology*, *176*, 73–83. https://doi.org/10.1016/j.jecp.2018.07.007

Castelain, T., Bernard, S., & Mercier, H. (2018). Evidence that two-year-old children are sensitive to information presented in arguments. *Infancy*, *23*, 124–135. https://doi.org/10.1111/infa.12202

Castelain, T., Girotto, V., Jamet, F., & Mercier, H. (2016). Evidence for benefits of argumentation in a Mayan indigenous population. *Evolution and Human Behavior*, *37*, 337–342. https://doi.org/10.1016/j.evolhumbehav.2016.02.002

Corriveau, K. H., & Kurkul, K. E. (2014). "Why does rain fall?": Children prefer to learn from an informant who uses noncircular explanations. *Child Development*, *70*, 1827–1835. https://doi.org/10.1111/cdev.12240

Fox, J., & Weisberg, S. (2011). *Multivariate linear models in R. An appendix to an R companion to applied regression*. Sage.

Galazka, M. A., Gredebäck, G., & Ganea, P. A. (2016). Mapping language to the mind: Toddlers' online processing of language as a reflection of speaker's knowledge and ignorance. *Cognitive Development*, *40*, 1–8. https://doi.org/10.1016/j.cogdev.2016.07.003

Ganea, P. A., Fitch, A., Harris, P. L., & Kaldy, Z. (2016). Sixteen-month-olds can use language to update their expectations about the visual world. *Journal of Experimental Child Psychology*, *151*, 65–76. https://doi.org/10.1016/j.jecp.2015.12.005

Ganea, P. A., & Harris, P. L. (2010). Not doing what you are told: Early perseverative errors in updating mental representations via language. *Child Development*, *81*, 457–463. https://doi.org/10.1111/j.1467-8624.2009.01406.x

Ganea, P. A., & Harris, P. L. (2013). Early limits on the verbal updating of an object's location. *Journal of Experimental Child Psychology*, *114*, 89–101. https://doi.org/10.1016/j.jecp.2012.04.013

Ganea, P. A., Koenig, M. A., & Gordon Millet, K. (2011). Changing your mind about things unseen: Toddlers' sensitivity to prior reliability. *Journal of Experimental Child Psychology*, *109*, 445–453. https://doi.org/10.1016/j.jecp.2011.02.011

Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology*, *60*, 115–140. https://doi.org/10.1146/annurev.psych.59.103006.093659

Hagá, S., & Olson, K. R. (2017). Knowing-it-all but still learning: Perceptions of one's own knowledge and belief revision. *Developmental Psychology*, *53*, 2319–2332. https://doi.org/10.1037/dev0000433

Hansson, S. O. (2022). Logic of belief revision. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Spring 2022 Edition. https://plato.stanford.edu/entries/logic-belief-revision/

Harris, P. L. (2015). *Trusting what you're told: How children learn from others*. Harvard University Press.

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. J. Beran, J. Brandl, J. Perner, & J. Proust (Eds.), *Foundations of Metacognition* (pp. 193–210). Oxford University Press. https://doi.org/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*(1), 251–273. https://doi.org/10.1146/annurev-psych-122216-011710

Harris, P. L., & Lane, J. D. (2014). Infants understand how testimony works. *Topoi*, *33*, 443–458. https://doi.org/10.1007/s11245-013-9180-0

Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Harvard University Press. https://doi.org/10.4159/9780674985155

Jipson, J. L., Labotka, D., Callanan, M. A., & Gelman, S. A. (2018). How conversations with parents may help children learn to separate the sheep from the goats (and the robots). In M. M. Saylor & P. A. Ganea (Eds.), *Active learning from infancy to childhood: Social motivation, cognition, and linguistic mechanisms* (pp. 189–212). Springer. https://doi.org/10.1007/978-3-319-77182-3_11

Kahan, D. M., Jenkins-Smith, H., & Braman, D. (2011). Cultural cognition of scientific consensus. *Journal of Risk Research*, *14*, 147–174. https://doi.org/10.1080/13669877.2010.511246

Kelly, T. (2002). The rationality of belief and some other propositional attitudes. *Philosophical Studies*, *110*, 163–196. https://doi.org/10.1023/a:1020212716425

Koenig, M. A. (2012). Beyond semantic accuracy: Preschoolers evaluate a speaker's reasons. *Child Development*, *83*, 1051–1063. https://doi.org/10.1111/j.1467-8624.2012.01742.x

Koenig, M. A., & Sabbagh, M. A. (2013). Selective social learning: New perspectives on learning from others. *Developmental Psychology*, *49*, 399–403. https://doi.org/10.1037/a0031619

Köymen, B., Mammen, M., & Tomasello, M. (2016). Preschoolers use common ground in their justificatory reasoning with peers. *Developmental Psychology*, *52*, 423–429. https://doi.org/10.1037/dev0000089

Köymen, B., O'Madagain, C., Domberg, A., & Tomasello, M. (2020). Young children's ability to produce valid and relevant counter-arguments. *Child Development*, *91*, 685–693. https://doi.org/10.1111/cdev.13338

Köymen, B., Rosenbaum, L., & Tomasello, M. (2014). Reasoning during joint decision-making by preschool peers. *Cognitive Development*, *32*, 74–85. https://doi.org/10.1016/j.cogdev.2014.09.001

Köymen, B., & Tomasello, M. (2020). The early ontogeny of reason giving. *Child Development Perspectives*, *14*(4), 215–220. https://doi.org/10.1111/cdep.12384

Kuczynski, L. (1984). Socialization goals and mother–child interaction: Strategies for long-term and short-term compliance. *Developmental Psychology*, *20*, 1061. https://doi.org/10.1037/0012-1649.20.6.1061

Kuhn, D., Zillmer, N., Crowell, A., & Zavala, J. (2013). Developing norms of argumentation: Metacognitive, epistemological, and social dimensions of developing argumentive competence. *Cognition and Instruction*, *31*, 456–496. https://doi.org/10.1080/07370008.2013.830618

Lane, J. D., Harris, P. L., Gelman, S. A., & Wellman, H. M. (2014). More than meets the eye: Young children's trust in claims that defy their perceptions. *Developmental Psychology*, *50*, 865–871. https://doi.org/10.1037/a0034291

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). *Package "emmeans"*. CRAN.

Lord, E. (2014). The coherent and the rational. *Analytic Philosophy*, *55*, 151–175. https://doi.org/10.1111/phib.12034

Ma, L., & Ganea, P. A. (2009). Dealing with conflicting information: Young children's reliance on what they see versus what they are told. *Developmental Science*, *13*, 151–160. https://doi.org/10.1111/j.1467-7687.2009.00878.x

Macris, D. M., & Sobel, D. M. (2017). The role of evidence diversity and explanation in 4- and 5-year-olds' resolution of counter-evidence. *Journal of Cognition and Development*, *18*, 358–374. https://doi.org/10.1080/15248372.2017.1323755

Mammen, M., Köymen, B., & Tomasello, M. (2018). The reasons young children give to peers when explaining their judgments of moral and conventional rules. *Developmental Psychology*, *54*, 254–262. https://doi.org/10.1037/dev0000424

Mercier, H., Bernard, S., & Clément, F. (2014). Early sensitivity to arguments: How preschoolers weight circular arguments. *Journal of Experimental Child Psychology*, *125*, 102–109. https://doi.org/10.1016/j.jecp.2013.11.011

Mercier, H., Sudo, M., Castelain, T., Bernard, S., & Matsui, T. (2017). Japanese preschoolers' evaluation of circular and non-circular arguments. *European Journal of Developmental Psychology*, *15*, 493–505. https://doi.org/10.1080/17405629.2017.1308250

Miosga, N., Schultze, T., Schulz-Hardt, S., & Rakoczy, H. (2020). Selective social belief revision in young children. *Journal of Cognition and Development*, *21*, 513–533. https://doi.org/10.1080/15248372.2020.1781127

O'Madagain, C. (2019). Is reasoning culturally transmitted? *Teorema: Revista Internacional De Filosofía*, *38*(1), 107–120. https://www.jstor.org/stable/ https://doi.org/10.2307/26746237

Over, H., Eggleston, A., Bell, J., & Dunham, Y. (2017). Young children seek out biased information about social groups. *Developmental Science*, *21*, e12580–e12644. https://doi.org/10.1111/desc.12580

Özdemir, B., & Ganea, P. A. (2020). Variability in toddlers' ability to verbally update their mental representations of absent objects. *Journal of Experimental Child Psychology*, *196*, 104843. https://doi.org/10.1016/j.jecp.2020.104843

Parfit, D. (2011). *On what matters: Two-volume set.* (Vol. 1). Oxford University Press.

Peppas, P., Nayak, C., Pagnucco, M., & Foo, N. Y. (1996). Revision vs. update: Taking a closer look. In W. Wahlster (Ed.), *12th European Conference on Artificial Intelligence* (pp. 95–99). John Wiley & Sons Ltd.

Pollock, J. L. (1987). Defeasible reasoning. *Cognitive Science*, *11*(4), 481–518. https://doi.org/10.1016/s0364-0213(87)80017-4

Poulin-Dubois, D., & Brosseau-Liard, P. (2016). The developmental origins of selective social learning. *Current Directions in Psychological Science*, *25*, 60–64. https://doi.org/10.1177/0963721415613962

Powell, M. J. D. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. Report DAMTP 2009/NA06.

R Core Team. (2019). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.r-project.org

Raz, J. (2011). *From normativity to responsibility*. Oxford University Press.

Ridge, K. E., Pesch, A., Suárez, S., & Koenig, M. (2018). Insights into children's testimonial reasoning. In M. M. Saylor & P. A. Ganea (Eds.), *Active learning from infancy to childhood: Social motivation, cognition, and linguistic mechanisms* (pp. 131–146). Springer. https://doi.org/10.1007/978-3-319-77182-3_8

Robinson, E. J., Champion, H., & Mitchell, P. (1999). Children's ability to infer utterance veracity from speaker informedness. *Developmental Psychology*, *35*, 535–546. https://doi.org/10.1037/0012-1649.35.2.535

Scanlon, T. (1998). *What we owe to each other*. Belknap Press of Harvard University Press. https://doi.org/10.1111/j.1933-1592.2003.tb00249.x

Schielzeth, H., & Forstmeier, W. (2009). Conclusions beyond support: Overconfident estimates in mixed models. *Behavioral Ecology*, *20*(2), 416–420. https://doi.org/10.1093/beheco/arn145

Schroeder, M. (2009). Means-end coherence, stringency, and subjective reasons. *Philosophical Studies*, *143*, 223–248. https://doi.org/10.1007/s11098-008-9200-x

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, *25*, 359–393. https://doi.org/10.1111/j.1468-0017.2010.01394.x

Tizard, B., & Hughes, M. (1984). *Young children learning.* Fontana Original.

Tomasello, M. (2019). *Becoming human—A theory of ontogeny.* Harvard University Press. https://doi.org/10.4159/9780674988651

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Schleihauf, H., Herrmann, E., Fischer, J., & Engelmann, J. M. (2022). How children revise their beliefs in light of reasons. *Child Development*, 93, 1072–1089. https://doi.org/10.1111/cdev.13758