

Children's selective trust decisions: rational competence and limiting performance factors

Jonas Hermes^{1,2} | Tanya Behne^{1,2} | Anna Elisa Bich¹ | Christa Thielert¹ | Hannes Rakoczy^{1,2}

¹Department of Developmental Psychology, Institute of Psychology, University of Göttingen, Germany

²Leibniz Science Campus Primate Cognition, Göttingen, Germany

Correspondence

Jonas Hermes, University of Göttingen, Institute of Psychology, Department of Developmental Psychology, Waldweg 26, D- 37073 Göttingen, Germany.
Email: jonas.hermes@psych.uni-goettingen.de

Funding information

Deutsche Forschungsgemeinschaft, Grant/Award Number: Ra-2155/1-1

Abstract

Recent research has amply documented that even preschoolers learn selectively from others, preferring, for example, reliable over unreliable and competent over incompetent models. It remains unclear, however, what the cognitive foundations of such selective learning are, in particular, whether it builds on rational inferences or on less sophisticated processes. The current study, therefore, was designed to test directly the possibility that children are in principle capable of selective learning based on rational inference, yet revert to simpler strategies such as global impression formation under certain circumstances. Preschoolers ($N = 75$) were shown pairs of models that either differed in their degree of competence within one domain (strong vs. weak or knowledgeable vs. ignorant) or were both highly competent, but in different domains (e.g., strong vs. knowledgeable model). In the test trials, children chose between the models for strength- or knowledge-related tasks. The results suggest that, in fact, children are capable of rational inference-based selective trust: when both models were highly competent, children preferred the model with the competence most predictive and relevant for a given task. However, when choosing between two models that differed in competence on one dimension, children reverted to halo-style wide generalizations and preferred the competent models for both relevant and irrelevant tasks. These findings suggest that the rational strategies for selective learning, that children master in principle, can get masked by various performance factors.

RESEARCH HIGHLIGHTS

- Young children's selective trust depends on their trait knowledge.
- Preschoolers generalize widely to unconnected domains when choosing between models that differ in overall competence.
- Preschoolers choose rationally between two models that are highly competent, yet in different domains, but only if they have correctly identified the models' traits.
- Young children are capable of rational selective trust decisions but fall back on global impression formation when this provides definite solutions.

1 | INTRODUCTION

When acquiring knowledge about the world, young children learn much, if not most, through the testimony of others. A growing body of recent research has shown that even preschoolers are selective in whom they learn from rather than trusting anyone indifferently (Clément, 2010; Harris, 2007; Mills, 2013; Robinson & Einav, 2014). For example, they prefer to learn from previously accurate labelers (Koenig, Clement, & Harris, 2004), from nice (Landrum, Mills, & Johnston, 2013) or attractive models (Bascandzief & Harris, 2014), from models who are familiar (Corriveau & Harris, 2009) or speak with

the same accent as the participant (Kinzler, Corriveau, & Harris, 2011) and from models who express certainty (Sabbagh & Baldwin, 2001).

1.1 | Cognitive foundations of selective social learning

But what cognitive foundations is such selective social learning based upon? Does it reflect adult-like rational inductive learning? That is, do children infer from a model's past behavior certain competencies/traits (e.g., that someone who provided accurate information is generally knowledgeable) that make the model a suitable source of information for certain future tasks? Alternatively, children's selective trust might not be based on such sophisticated reasoning about traits but rather on less differentiated global impression formation: If models differ with regard to a certain competence or trait, children might perceive models as somehow positive or negative in a global sense and choose the more positive one for a wide range of future problems – including those unrelated to the original competence or trait displayed.

The empirical findings so far cannot resolve this question given the somewhat mixed findings and the different structure of problems children were confronted with. In one set of studies, children were asked to choose between two models with similar degrees of competence yet in different domains. These studies suggest that children can differentiate between multiple epistemic cues in rational ways: that is, they infer specific model characteristics and competencies from observed behavior and consider these in their model choice in light of the task requirements. For example, when presented with experts from different domains of competence (e.g., toy labeler vs. toy fixer) children competently attributed knowledge to these experts when seeking information in the respective domains (Kushnir, Vredenburgh, & Schneider, 2013; Lutz & Keil, 2002; Sobel & Corriveau, 2010). And children understand that adults and children may be experts in different fields, trusting peers over adults for toy questions and adults over peers for food questions (VanderBorghet & Jaswal, 2009).

In a second set of studies children were presented with two models that differed in the degree of competence on the very same dimension (e.g., accurate vs. inaccurate labeling). In many such studies children preferred the more competent model also for tasks totally unrelated to the particular domain of competence the model displayed: For example, they expected a knowledgeable model to be nice too (Brosseau-Liard & Birch, 2010), a nice model to be smart and athletic (Cain, Heyman, & Walker, 1997), a strong model to be nice and smart (Fusaro, Corriveau, & Harris, 2011) or an attractive model to be knowledgeable (Bascandziev & Harris, 2014). However, children did not always show such a pattern of preferences: First, the inferences children draw from social to epistemic domains decrease once epistemic cues, that are more clearly predictive for epistemic problems, are made available: older preschoolers, for example, trust an accurate-unfamiliar over an inaccurate-familiar informant (Corriveau & Harris, 2009), a foreign-accented and accurate over a native-speaking and inaccurate speaker (Corriveau, Kinzler, & Harris, 2013) or a reliable dissenter over an unreliable majority (Bernard, Proust, & Clément, 2015). Second,

children's preference for a more competent model outside her domain of competence seems to be restricted to avoidance of incompetent models, but does not tend to occur when children choose between a neutral agent and an expert model (Koenig & Jaswal, 2011).

Taken together, children's cognitive strategies tend to differ as a function of the epistemic situations they find themselves in: When they are confronted with two experts with similar degrees of competence yet in different domains, they show rational inferences based on the ascription of specific competencies. In contrast, when choosing between two models that differ in overall competence on one dimension the picture is less clear.

1.2 | Rational competence masked by performance factors?

How can we make sense of these findings? One possibility is that children are capable, in principle, of rational selective learning and trust, but that this competence is masked in some situations by certain performance factors. One such limiting performance factor might be the structure and demands of the task. Children might use simpler strategies such as global impression formation as long as these yield unique determinate answers for a given task; and they might recruit their more rational strategies only in tasks for which the simpler strategies fail to yield unique answers. When models differ in overall competence in any domain, children might thus tend to prefer the more competent model even for tasks outside of the actual scope of her competence. In contrast, whenever models are similarly competent, yet in different domains, a global strategy would be inconclusive and not lead to a determinate solution. In such cases children might thus make use of their rational competence and recruit models in accordance with their perceived traits. To test for this possibility directly, one would have to manipulate the task structure and model contrasts in a within-subject design, and assess children's performance in tasks with the two structures, respectively – a methodological approach that, to our knowledge, no study has realized so far.

Besides task structure, what other limiting factors may mask children's capacity for rational inferences in selective trust? Another candidate that comes to mind is (lack of) background knowledge (Sobel & Kushnir, 2013). Imagine a child is confronted with models differing in a certain behavior that is in fact diagnostic for a certain trait (e.g., lifting objects as an indicator of strength). Now, the child can only see the behavior as evidential basis for ascription of the trait in question, and thus as a basis for predicting future behavior expressive of the trait if she has some background knowledge about the domain and some mastery of the relevant trait concept. Just as someone without any background knowledge about football cannot see an event as 'offside', a child without any background knowledge about physical strength and thus without the trait concept 'strong' cannot see events of lifting, etc. as expressions of an agent's strength. In the absence of the relevant background knowledge and trait concepts, all that children can do, it seems, is to engage in global impression formation.

Evidence compatible with this idea comes from those studies reviewed above that show that children differentiate between experts

from different domains – since such differentiation plausibly was based on conceptual knowledge about the specific kinds of expertise (Kushnir et al., 2013; Lutz & Keil, 2002; VanderBorghet & Jaswal, 2009). However, this evidence is rather indirect, since background trait knowledge was not independently assessed in these studies. More direct preliminary support comes from a recent study that did assess children's selective learning and their trait knowledge independently and showed that only those children with the requisite trait knowledge made rational model choices (Hermes, Behne, & Rakoczy, 2015). Converging evidence for a direct relation between conceptual knowledge and selective model choice comes from another recent study in which children's own understanding of false belief, for example, predicted how much they endorsed information from a protagonist with sophisticated false belief understanding over that of a protagonist with no such understanding (Van Reet, Green, & Sobel, 2015).

In addition to the structure of the tasks and the availability of background trait knowledge, two more factors might potentially interfere with children's rational competence for selective learning: pragmatic confusion and executive function. Concerning pragmatic confusion, in the studies reviewed above, in which children seem to generalize widely to unrelated dimensions, the format of the test questions may have misled them: When children are asked to choose between models for a task that requires a certain trait, and no information concerning that trait is provided about the models (e.g., a knowledge question after a strength demonstration), the most rational answer would be indifference ('I do not know', 'both', 'none'). A standard forced choice question that provides alternatives (e.g., 'Who knows ... model A or model B?') but excludes the most rational one ('neither'/'both') is pragmatically misleading and might mask children's competence and seduce them into giving answers based on their global impressions (since we know from a large body of research that young children are highly prone to suggestive cues; Ceci & Bruck, 1993). If this were the case, it should be easier for children to express their rational inferences if questions were less suggestive.

Concerning executive function (EF), we know that EF is related to rational selective trust decisions. Young children have a robust bias to trust testimony in the absence of competing information even if the informant was highly unreliable previously (Jaswal, Croft, Setia, & Cole, 2010). However, if they do have competing first-hand experience (e.g., about the location of a hidden object), children can overcome this uncritical bias to trust, and the degree to which they do so strongly correlates with their EF (Jaswal & Pérez-Edgar, 2014; Jaswal et al., 2014). Thus, preschoolers' use of rational strategies of selective trust, more generally, may be closely related to their developing EF: Advanced EF might help to inhibit global reasoning as a primary response when two models differ in overall competence, or to identify the relevant trait for a given task when confronted with two models with similar degrees of competence in different domains.

1.3 | Rationale of the present study

In sum, then, the present empirical picture regarding young children's selective trust is mixed, with some studies suggesting that this

capacity is based on sophisticated reasoning about traits whereas other studies tend to suggest that it is based on global impression formation. How can we make sense of this? The main rationale of the present study was to test whether the general idea of a rational competence account (that children's selective trust is based on the cognitive competence for rational inductive learning, masked in certain situations by certain limiting performance factors) can explain the diverging patterns of performance in different selective learning tasks. More specifically, we tested the influence of potential limiting performance factors, focusing especially on different task structures and background trait knowledge.

To test whether children's rational selective model recruitment is limited by the task structure, we systematically varied and contrasted (within-subject) tasks involving two types of model contrasts:

First, in the *degree-of-competence contrasts*, children were confronted with two models that differed in competence in one dimension (strength or knowledge) and were asked to decide between the models for strength-related tasks and knowledge-related tasks. If availability/lack of trait knowledge is the only limiting performance factor, children with the corresponding trait knowledge should prefer a competent over an incompetent model only for tasks within the scope of the model's competence (e.g., strong model for strength tasks) but choose indifferently for tasks outside the scope of her competence (knowledge-related tasks after strength manipulation). However, if – in addition to trait knowledge – task structure is a limiting factor (such that children revert to global impression formation as long as it yields unique answers), children should generalize widely in the degree-of-competence contrasts (e.g., preferring the strong model for both knowledge- and strength-related tasks) irrespective of their trait knowledge.

Secondly, in the *domain-of-competence contrast*, children encountered two models that were both highly competent, but in different domains (strong vs. accurate model). In this contrast we expect children to choose models rationally in accordance with the models' competencies, as long as they do not lack the corresponding trait knowledge. In this case the task structure should not limit rational reasoning since the predominant strategies based on global impression formation are inconclusive when both models show similar degrees of competence. The within-subject design allowed us to test not only for children's performance in each model contrast, but also for the relation of performance across the contrast types. In particular, we were able to test whether – as predicted by a rational competence account that considers task structure a crucial performance factor – children who rationally chose models in the domain-of-competence contrast were still subject to wide generalizations based on global impression formation in the degree-of-competence contrasts.

To test for the effect of background trait knowledge, we directly measured children's trait knowledge using explicit questions (as in Hermes et al., 2015). If only those children who answer the trait questions correctly recruit models rationally, this would speak for a general rational competence account, with lack of trait knowledge as a limiting factor.

Furthermore, in addition to task structure and trait knowledge, we explored the pragmatics of suggestive questioning and executive

function as two additional, potentially limiting performance factors on children's rational selective trust decisions.

2 | METHOD

2.1 | Participants

Seventy-five children were included in the final sample; 37 children were 4-year-olds (48–58 months, $M = 53.2$ months, $SD = 3.39$, 17 girls) and 38 children were 5-years-olds (60–71 months, $M = 66.3$ months, $SD = 3.48$, 19 girls). We chose this age range based on recent research showing that children from 4 to 5 years perform competently in selective recruitment when confronted with two models with equivalent competence, yet in different domains (Hermes et al., 2015; Li, Heyman, Xu, & Lee, 2014). Eight additional children took part but were not included in the final sample, because they did not complete the test session ($n = 3$) or failed to answer the comprehension questions correctly ($n = 5$, for details see below).

Children were recruited from a database of families who had agreed to participate. They were from mixed socioeconomic backgrounds and some children had acquired more than one language. As we had no data about their proficiency in the language used for testing, a vocabulary screening was conducted (vocabulary test from the Kaufman Assessment Battery for Children; see Kaufman, Kaufman, Melchers, & Preuss, 1994). Since 13 children scored significantly below average (two standard deviations below the mean for normally developing 48–51-month-olds), all statistical analyses were also run separately with these children excluded. The results of these control analyses replicated the results for the whole sample.

2.2 | Design and outline

The selective trust study followed a 2×2 design, examining the effect of *question format* (standard question vs. reduced suggestibility) and *model contrast* (*degree-of-competence* or *domain-of-competence*). The factor *question format* was varied between subjects: half of the children were asked the standard closed questions, and for the other half the suggestiveness of the questions was reduced by making the answer 'both the same' an equally available option.

The factor *model contrast* was varied within subjects: In the *degree-of-competence contrasts*, all children saw two pairs of models that differed in overall competence in the domains strength (successful vs. unsuccessful lifting) and knowledge (expert labeling vs. inaccurate labeling) respectively, whereas in the *domain-of-competence contrast*, children were confronted with the two highly competent models (strong and accurate). The degree-of-competence contrasts were shown first to ensure that children understood that the degree of competence in strength and accuracy might vary between models and would not assume that all these puppets were highly competent in both dimensions when they saw the domain-of-competence contrast. After the demonstrations of each of the three model contrasts, children were asked to select a model when they needed to obtain the labels for novel objects (knowledge tasks) and when they faced a

problem that required physical strength (strength tasks). In addition, trait questions were asked as a measure of background trait knowledge and for analysis the sample will be divided into those who answered trait questions correctly and those who answered incorrectly.

2.3 | Material and procedure

Children were tested in quiet rooms in their day care centers, with test sessions lasting approximately 30 minutes. The test sessions started with the vocabulary screening followed by a practice game to train question formats according to condition. The main selective trust study consisted of two degree-of-competence contrasts, followed by a test block each, and one domain-of-competence contrast, followed by a third test block. After this children answered trait questions and preference questions about the models, and at the end they did the executive function test.

2.3.1 | Question format training

The question formats used in the selective trust study were introduced in a short animal quiz: Children were shown three pairs of animals (one pair at a time, e.g., a cow and a pig) and asked to attribute characteristics to the animals. The characteristics, the question format used and the answer options provided differed between the two question format conditions: In the 'reduced suggestibility' condition, children were asked to attribute characteristics that applied to only one of the animals (e.g., who is pink?) and additionally a characteristic that applied to both animals (e.g., who lives on a farm?). Children were provided with an answer card that depicted three answer options, (1) just animal A, (2) just animal B, or (3) animal A and B; and the question format also highlighted this (e.g., Who is pink, only the pig, only the cow, or both?). Children were asked to answer by pointing to the answer card and were given feedback on their response. In the 'standard questions' condition, children were shown the same pairs of animals but were only asked to attribute characteristics that applied to only one of the animals with standard questions (e.g., Who is pink, only the pig or only the cow?). They answered without answer cards and were given feedback on their response. For the rest of the test sessions, all questions in the 'reduced suggestibility' condition in which children were asked to decide between models ended with '... model A, model B, or both?' and children were given answer cards that depicted the three options. In the 'standard questions' condition, all the following questions ended with '... model A or model B?' and children answered without answer cards.

2.3.2 | Main selective trust study

Children were shown pairs of models on a notebook pc. First, they saw model contrasts with an overall discrepancy in competence (degree-of-competence contrasts) in the domains strength (strong vs. weak model) and labeling accuracy (accurate vs. inaccurate model). Then in the domain-of-competence contrast the two highly competent models (strong and accurate) were combined. Each demonstration phase

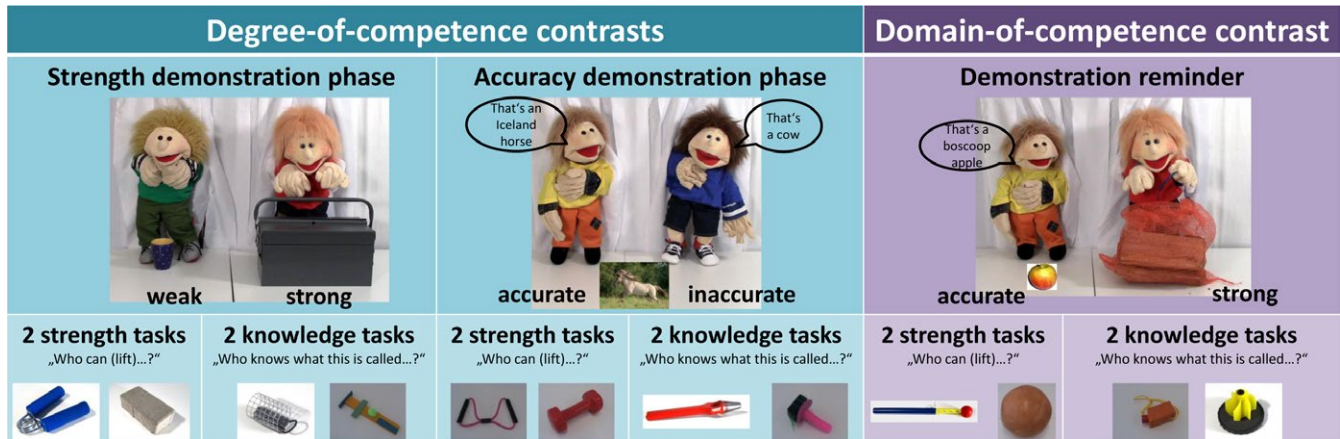


FIGURE 1 Procedure of the main selective trust study. Note that in the degree-of-competence contrasts, the order of the accuracy and strength phase was counterbalanced between subjects

was followed by a test block with four trials (two knowledge trials and two strength trials). See Figure 1 for a summary of the procedure.

Degree-of-competence contrasts

Children participated in degree-of-competence contrasts in both the accuracy and the strength domain, with the order of presentation counterbalanced across subjects. For each domain, a pair of models was introduced in a demonstration phase, followed by a test block with four trials (two knowledge trials and two strength trials).

Accuracy demonstration phase (4 trials). Children were introduced to two puppet models (Ms. Yellow and Ms. Blue; degrees of competence counterbalanced). In each demonstration trial, the models were presented with a picture of a known object in front of them (e.g., a jacket, see Supporting Information A). Children were given the same picture and were asked what it depicted. They normally provided the common label (e.g., jacket). The experimenter double-checked the label on a list and provided an expert label (e.g., ski jacket). Then children were shown a video in which one model provided the same expert label as previously read from the list whereas the other model provided a wrong label (e.g., pullover). Thus the accurate model was presented as more knowledgeable than the child herself and the inaccurate model as less knowledgeable. Expert labels always included the common labels (e.g., ski jacket) so that children would not mistake them as inaccurate. As the inaccurate model was supposed to be perceived as inaccurate rather than as bizarre or deceptive, the wrong labels were always part of the same category (e.g., pullover for jacket). After each trial, the experimenter repeated what the models had said and after the four trials children were asked a comprehension question: Who was good at labeling these things ...? (See above at *question format training* for how questions differed between conditions here and for all following questions.) If the child did not answer correctly, an additional trial was presented followed by a repetition of the comprehension question. If children still did not answer correctly, they were excluded from the final sample (for details see below).

Strength demonstration phase (4 trials). Children were shown another pair of puppet models (Ms. Green and Ms. Red; degrees of competence counterbalanced). At the start of each demonstration trial, one model had a light object (e.g., a cup) lying in front of her and the other model

a heavy object (e.g., a potato sack, see Supporting Information A for details). On video children then saw one puppet always failing to lift the light objects saying 'I can't manage this!' and the other puppet always succeeding in lifting the heavy objects saying 'I'm good at this' (order counterbalanced across trials). The experimenter repeated what had happened after each trial ('Ms ... was able to lift ... and Ms ... was not able to lift ...'). The objects to be lifted were chosen so that the weak puppet was perceived as weaker than the child herself and the strong puppet as stronger. After four trials, we asked a comprehension question: 'Who was good at lifting these things ...?'. If children did not answer correctly, they were shown one more demonstration trial and asked the comprehension question again. Children who failed to answer the repeated comprehension questions correctly, either following the accuracy demonstration ($n = 3$) or the strength demonstration ($n = 1$) or after both demonstrations ($n = 1$) were excluded from the final sample.

Test blocks in the degree-of-competence contrasts. After each of the strength and the accuracy demonstration blocks, children were presented with a test block. Each test block consisted of two strength trials and two knowledge trials (order counterbalanced). At the start of each trial the models had an object lying in front of them and the experimenter said: 'Look what they have! I brought the same for you.' Children were given the same objects to act on (see Supporting Information B for objects and exact test questions). In the knowledge trials the objects were unknown to the children and the children were asked as a test question: 'Who knows what this is called ...?' If children claimed to know the object, the label they provided was doubted by the experimenter and if they insisted, a spare object was introduced. In the strength trials, the objects required physical strength to act on them (e.g., to lift a brick). After children acted on each object, they were asked which model would be good at performing the action. For each task, we accepted pointing to the puppet (or the answer card, if applicable) or saying the color of the puppets as answers.

Domain-of-competence contrast

For the domain-of-competence contrast, the two competent models, i.e., the strong and the accurate one from the degree-of-competence contrasts were presented.

Demonstration reminder (4 trials). Children witnessed the strong and the accurate model in four demonstration trials that generally followed the same procedure as in the degree-of-competence contrasts. In two trials the strong puppet succeeded in lifting a heavy object while the accurate puppet did nothing. In the other two trials the accurate puppet provided an expert label for a known object while the strong puppet said nothing (see Supporting Information A for objects). The experimenter asked two reminder questions: 'Ms ..., was she good or not so good at lifting these objects?' and 'Ms ..., was she good or not so good at labeling these objects?' Children who did not answer both questions correctly were shown two more demonstration trials (one puppet performing an action each) and asked the reminder questions again. Seventeen children still did not answer both questions correctly. Since these children correctly identified the strong and the accurate model in the degree-of-competence contrasts, they were not excluded from analysis. However, all results for the domain-of-competence contrast were replicated with the smaller subsample (with children excluded who erred in the reminder questions), with one exception described below.

Test block in the domain-of-competence contrast. The following test block was the same as described above for the degree-of-competence contrasts. That is, children again participated in two strength trials and two knowledge trials and were asked the same questions (for details see above and Supporting Information B).

Trait questions and preference questions

At the end of the selective trust study, children saw the strong model and the accurate model on the screen and were asked two trait questions ('Who is strong ...?' and 'Who is smart ...?') and two preference questions ('Who is nice ...?' and 'Who would share her sweets ...?')

2.3.3 | Executive functions task

At the end of the test session, children performed the Head-Toe-Knee-Shoulder task (HTKS) (McClelland et al., 2014; Ponitz, McClelland, Matthews, & Morrison, 2009; Ponitz et al., 2008) to test their capacity to inhibit a predominant response. That is, on hearing a request to perform a particular action (e.g., 'Touch your head') children were required to perform an opposite action (e.g., to touch their foot). Children performed on 30 trials in three stages which increased in cognitive complexity (adding additional rules in stage 2 and changing rules in stage 3; see McClelland et al., 2014).

2.4 | Coding procedure

Two coders watched half of the videos each. For the degree-of-competence contrasts, it was coded whether children chose the competent model (accurate or strong) or the incompetent model or expressed indifference. For the domain-of-competence contrast and the trait and preference questions, it was coded whether children chose the strong model or the accurate model or expressed indifference (see Supporting Information C for descriptions). An indifference response was scored when children

answered 'both' (86% of all cases), 'none' (1%), or 'I don't know' (13%) (see Supporting Information D for details). In each trial in the HTKS task children received two points if they spontaneously responded correctly, one point if they initially did a wrong action, but then corrected themselves and zero points if they performed a wrong action and did not correct it. A sum score was calculated across all (up to 30) trials (range: 0–60 points).

2.4.1 | Inter-rater reliability

One additional coder who was blind to the hypotheses coded a subset of 20% of the data ($n = 17$ participants). Agreement for all test questions and comprehension questions in the main selective trust study was perfect ($K_s = 1$). There were deviations in two data points in the reminder question resulting in $K = .82^{**}$ for the accuracy reminder questions and $K = .62$ for the second strength reminder question for those children who answered the first incorrectly. Inter-rater agreement concerning the HTKS sum score was high, as indicated by a one-way random intraclass correlation ($ICC = .93$, $F(14,15) = 13.72$, $p < .01$).

3 | RESULTS

3.1 | Focal analysis: patterns of selective model choice in the degree-of-competence contrasts and in the domain-of-competence contrast

We were interested in children's model choice and patterns of generalizations across domains when given a choice between two models that differed in overall competence (degree-of-competence contrasts) and when given a choice between two models with high competence yet in different domains (domain-of-competence contrast). To explore these, we aggregated children's responses across the two trials of each task. Specifically, in the degree-of-competence contrasts we aggregated children's responses to four *preference scores*, one each for the strength tasks and the knowledge tasks after both the strength demonstration and the knowledge demonstration: for each of these we determined the number of trials (0–2) in which the child chose the competent model (strong or knowledgeable) and the number of trials (0–2) in which she chose the incompetent model (weak or unknowledgeable), and computed the preference score by subtracting the latter from the former. Thus, the preference scores could range from –2 to 2, with high values indicating a preference for the more competent model. In the domain-of-competence contrast, children's responses across the two strength trials and the two knowledge trials were aggregated to two *expert scores* in the following way: For each type of question, we determined the number of trials (0–2) in which the child chose the model with the relevant expertise (strong model for strength trials, knowledgeable model for knowledge trials), the number of trials (0–2) in which she chose the model with the irrelevant expertise (e.g., strong model for knowledge trials) and computed the expert score by subtracting the latter from the former. Thus the expert scores could range from –2 to 2 with

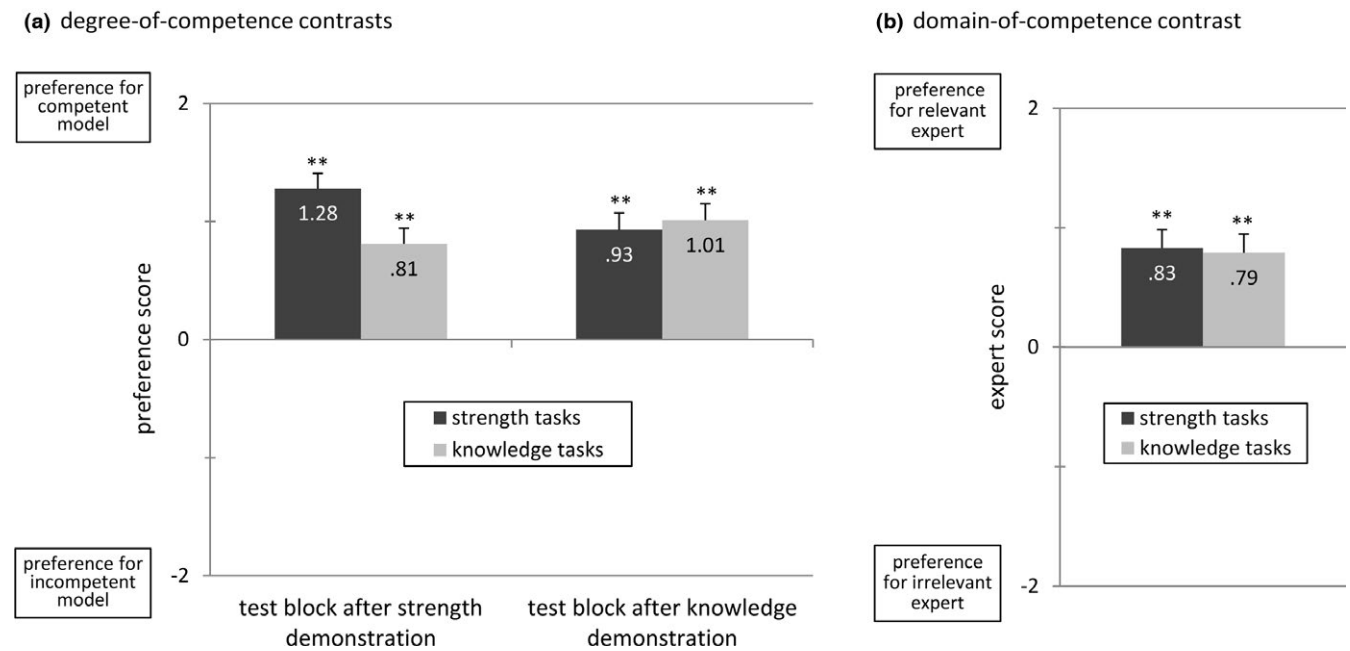


FIGURE 2 Children's model preference in the strength tasks and the knowledge tasks. Figure 2a depicts children's preference scores in the strength and the knowledge tasks in the degree-of-competence contrasts. Positive values indicate a preference for the competent (i.e., strong or knowledgeable model, according to demonstration) over the incompetent (i.e., weak or unknowledgeable) models and negative values indicate a preference for the incompetent over the competent models. Figure 2b depicts children's expert scores in the strength and the knowledge tasks in the domain-of-competence contrast. Positive values indicate a preference for the model with the relevant expertise (e.g., strong model for the strength tasks) over the model with the irrelevant expertise (e.g., knowledgeable model for the strength tasks) and negative values indicate the reverse preference. The line at 0 reflects chance level. Asterisks show deviation from chance level (Wilcoxon signed-rank tests, $**p < .0017$ [Bonferroni-corrected $\alpha = .01$ for 6 tests]). Error bars show standard errors

high values indicating a preference for the model with the relevant expertise.¹

Figure 2 depicts these preference scores and expert scores for each test block in each group of tasks. In the degree-of-competence contrasts children preferred the more competent models significantly above chance for all tasks (see Figure 2a). This indicates that children generalized widely, selecting the more competent models for all tasks, including those irrelevant to the competencies displayed (e.g., for knowledge tasks after strength demonstration). In the domain-of-competence contrast, children preferred the models with the relevant expertise (i.e., the strong model for the strength tasks and the knowledge model for the knowledge tasks) significantly above chance (see Figure 2b). Thus in the domain-of-competence contrast, children rationally chose the models in accordance with their respective competencies and the requirements of the given tasks.²

3.1.1 | Aggregation across strength and knowledge tasks

We explored whether the degree of children's preference for the strong model differed from that for the knowledgeable model. In the degree-of-competence contrasts, there was no difference in children's degree of preference for the competent model (strong or knowledgeable), when the domain of demonstration and task matched (i.e., preference scores for strength task after strength demonstration and for knowledge task after knowledge demonstration, Wilcoxon

signed-rank test, $Z = 1.81, p > .05$), and when the domain of demonstration and task did not match (i.e., preference score for strength task after knowledge demonstration and for knowledge task after strength demonstration, $Z = 0.78, p > .05$). Similarly, in the domain-of-competence contrast the expert score of the strength tasks was not different from that of the knowledge tasks ($Z = 0.12, p > .05$). Thus, children's model preferences did not differ when they were reasoning about knowledge and strength.

For further statistical analyses we, thus, aggregated the data across strength tasks and knowledge tasks to scores indicating whether children chose the relevant models in the relevant and irrelevant tasks in the following way: In the degree-of-competence contrasts we aggregated children's preference scores by whether or not the specific competence demonstrated by the models was relevant to a given task. To this end, we created a *relevant tasks score* by adding the preference score of the strength trials after strength demonstration and the one for the knowledge trials after knowledge demonstration, yielding an aggregate score ranging from -4 to 4 , with high values indicating a preference for models that were competent (rather than incompetent) in domains relevant to the tasks at hand. And we created an *irrelevant tasks score* by adding the preference score of the strength trials after knowledge demonstration and the one of the knowledge trials after the strength demonstration, yielding an aggregate score ranging from -4 to 4 , with high values indicating a preference for models that were competent (rather than incompetent) in domains irrelevant to the task at hand. In the domain-of-competence contrast we aggregated the

expert score for the strength tasks and the one for the knowledge tasks to a *combined relevant model score* ranging from -4 to 4 , with high values indicating preferences for models whose competence was relevant (rather than irrelevant) for a given task.

3.1.2 | The role of task relevance in the degree-of-competence contrasts

In the degree-of-competence contrasts children preferred the competent over the incompetent models for all tasks including those irrelevant to the specific competence of a model (see Figure 2). To assess whether the degree of children's preference for the more competent model was higher for tasks relevant to the specific competence the model had demonstrated than for those irrelevant to the specific model competence, we compared the relevant tasks score and the irrelevant tasks score. Children did choose the competent model more often in tasks relevant to her competence than in tasks irrelevant to her competence (Wilcoxon signed-rank test, $Z = 2.99$, $p < .01$, $r = .35$). This shows that although children did show halo-patterns (choosing the competent model more often than the incompetent one, even when the competence displayed was irrelevant to the task), there was still some differentiation such that they chose a given model less often when her competence was irrelevant than when it was relevant.

3.2 | Additional analyses

To explore whether children's patterns of selective model choice were related to their age, their trait knowledge, the question format they were exposed to, or their degree of executive function, we analyzed how these factors predicted the relevant tasks score and the irrelevant tasks score in the degree-of-competence contrasts and the combined relevant model score in the domain-of-competence contrast. Trait knowledge, age group, and question format condition were included as categorical predictors in ordinal regressions on each of these three scores. Due to missing data in the HTKS task,³ we assessed the effect of children's executive function separately in correlation analyses.

3.2.1 | Trait knowledge

We tested whether children's trait knowledge, as indicated in their explicit answers to the trait questions at the end of the selective trust study, affected their patterns of model choice. Preliminary analyses showed that children correctly identified the 'smart' model and the 'strong' model at above chance levels. When asked 'who is smart?' more children chose the accurate model rather than choosing the strong model ($\chi^2(1) = 20.25$, $p < .01$). Similarly, when asked 'who is strong?', more children selected the strong model rather than choosing the accurate model ($\chi^2(1) = 37.88$, $p < .01$) (see Supporting Information C for details). For ordinal regression analyses, we aggregated the data over both trait questions and compared the sub-sample of children who answered both trait questions correctly ($n = 46$) with

the sub-sample of children who answered at least one trait question incorrectly ($n = 29$). Figure 3 shows children's model choice separately for these two sub-samples.

For the degree-of-competence contrasts, trait knowledge did not predict the relevant tasks score or the irrelevant tasks score. As can be seen in Figure 3a, all children including those who made mistakes answering the trait question and those who answered the trait questions correctly preferred the competent over the incompetent models above chance in all tasks. Thus trait knowledge did not help children restrict their generalizations to relevant tasks. In contrast, for the domain-of-competence contrast, children's trait understanding did predict their combined relevant model score: Those children who attributed both traits correctly were 9.75 (95% CI, 3.54 to 26.87) times more likely to select the model with the relevant expertise for a given task than those children who did not attribute both traits correctly (Wald $\chi^2(1) = 19.37$, $p < .001$). As can be seen in Figure 3b, children who made at least one mistake in attributing the traits chose between the models at chance level. In contrast, those children who attributed both traits correctly selectively chose the model with the relevant competence for a given task (i.e., the strong model for strength tasks and the accurate model for knowledge tasks).

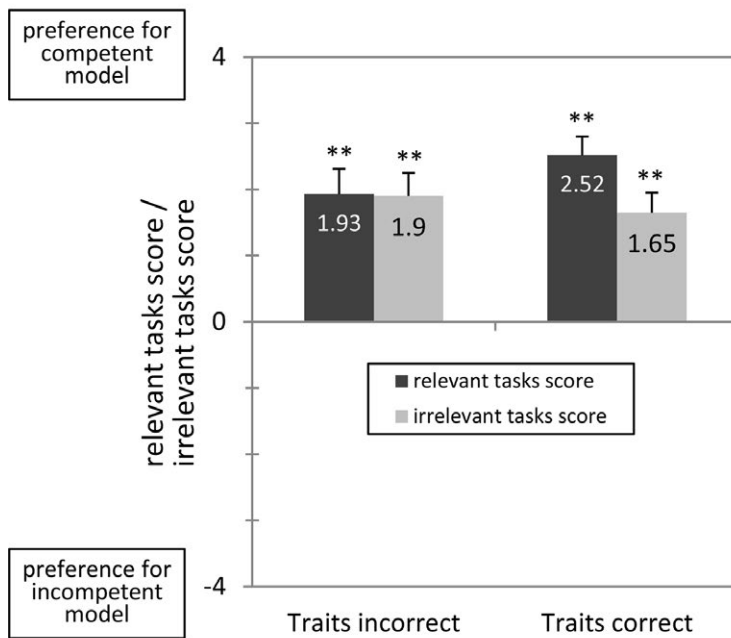
3.2.2 | Age group

Children's age group did not significantly predict the relevant tasks score or the irrelevant tasks score in the degree-of-competence contrasts. In the domain-of-competence contrast, children's age group did predict the combined relevant model score: 5-year-olds were 2.54 (95% CI, 1.05 to 6.13) times more likely than 4-year-olds to select the model with the relevant expertise for a given task (Wald $\chi^2(1) = 4.31$, $p < .05$). However, the combined relevant model score was significantly above chance level for both 4-year-olds (combined relevant model score = 0.89; Wilcoxon signed-rank test, $Z = 2.61$, $p < .01$, $r = .43$) and 5-year-olds (combined relevant model score = 2.32; Wilcoxon signed-rank test, $Z = 4.68$, $p < .001$, $r = .76$). Thus, both 4-year-olds and 5-year-olds chose the model with the relevant expertise for a given task above chance, but 5-year-olds did so even more frequently than 4-year-olds.

3.2.3 | Question format

In a preliminary analysis we compared the sum of indifference answers children provided across the 12 test trials as a function of the question format (for details see Supporting Information D). Children in the Reduced Suggestibility condition provided more indifference answers ($M = 3.62$) than those in the Standard Questions condition ($M = 0.74$; Mann-Whitney test, $U(38, 37) = 290$, $p < .001$, $r = .53$). Yet in the ordinal regression analyses, the question format children had encountered did not predict their model choice in either the degree-of-competence contrasts or the domain-of-competence contrast. Therefore, although those children who were asked less suggestively provided more indifference answers than those who were asked standard questions, this difference in frequency was unsystematic.

(a) degree-of-competence contrasts



(b) domain-of-competence contrast

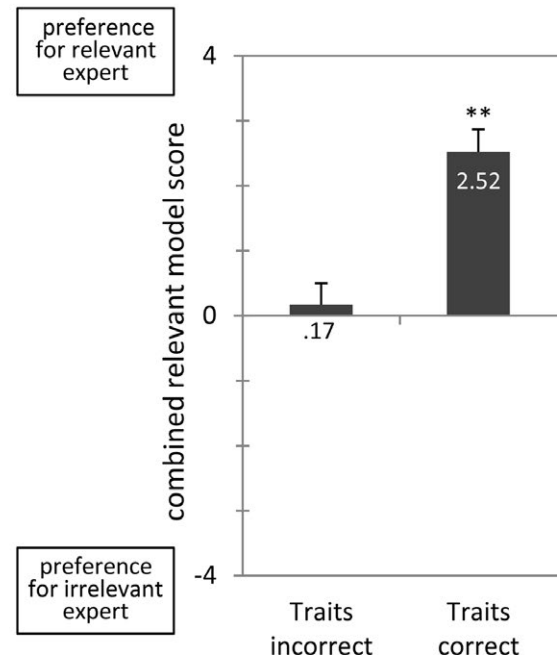


FIGURE 3 Children's model preferences separated by trait understanding: For the degree-of-competence contrasts, Figure 3a depicts children's relevant tasks scores (positive values indicating children's preference for the more competent model in tasks relevant to the model's competence) and irrelevant tasks scores (positive values indicating children's preference for the more competent model in tasks irrelevant to the model's competence) separated by their trait understanding. Figure 3b depicts children's combined relevant model score (positive values indicating a preference for the model with the expertise relevant to the task, over the model with the expertise irrelevant to the task). The line at 0 reflects chance level. Asterisks show deviation from chance level (Wilcoxon signed-rank tests, $**p < .0017$ [Bonferroni-corrected $\alpha = .01$ for 6 tests]). Error bars show standard errors

Children in the Reduced Suggestibility condition did not express indifference particularly often in cases where indifference was justified (when choosing models for tasks for which their demonstrated competence was irrelevant; irrelevant tasks score in the degree-of-competence contrasts).

To assess potential masking effects of suggestive questions in more detail, the responses of children with trait knowledge are of special interest: Would these children refrain from wide generalizations when asked less suggestively? The crucial subsample to test this includes only those children in the reduced suggestibility condition who answered both trait questions correctly ($n = 21$). Planned analysis of this subsample in the degree-of-competence contrasts showed that even these children drew wide generalizations, preferring the more competent model above chance even for tasks outside the model's specific competence (irrelevant tasks score = 2.24; Wilcoxon signed-rank test, $Z = 3.57$, $p < .001$, $r = .78$).

3.2.4 | Executive functions

In order to test whether executive function was specifically related to rational rather than simpler, impression-based processes, we correlated the HTKS scores with children's decisions in both model contrasts: First, for those tasks in which children saw models that differed in overall competence (degree-of-competence contrasts), the HTKS scores did

not correlate with the irrelevant tasks score ($r_s = -.06$, $p = .65$) or the relevant tasks score ($r_s = -.03$, $p = .81$). Hence there was no evidence that higher inhibitory control predicted more restrained generalizations or higher rates of rational choice in the degree-of-competence contrasts. Second, for the domain-of-competence contrast, the combined relevant model score correlated positively with the HTKS score ($r_s = .27$, $p < .05$).⁴ Thus, children with advanced executive functions showed higher degrees of rational choice in the domain-of-competence contrast than did those children low in executive functions.

3.2.5 | Preference questions

Children's model choices in the preference questions were compared to chance using χ^2 tests (see Supporting Information C for details). Children had no significant preference for the strong or the accurate model when asked 'who will share her toys ...?' ($\chi^2(1) = .82$, $p = .37$) or 'who is nice ...?' ($\chi^2(1) = 0.5$, $p = .48$). Thus the preference questions indicate no general preference for one of the models.

4 | DISCUSSION

The aim of the present study was to explore the cognitive foundations of children's selective learning. In particular, we tested the general idea

put forward by rational competence accounts that children's selective learning generally builds on a capacity for rational inductive inference (Sobel & Kushnir, 2013) that may get masked under certain circumstances by limiting performance factors or task demands. The main focus was on the systematic variation of two such potential limiting performance factors. The first was the availability/ lack of requisite background knowledge, here conceptualized as trait knowledge and assessed through direct trait questions. The idea is that only those children with sufficient trait knowledge should be able to recruit models rationally and refrain from global impression formation. As a second factor that potentially limits children's use of their rational strategies, we systematically varied the structure of tasks and compared two different kinds of model contrasts that differ in the degree to which they encourage simpler strategies such as global impression formation. In the degree-of-competence contrasts, in which two models differed in overall competence in one domain, global impression formation would have yielded unique determinate answers. In the domain-of-competence contrast where two models showed the same degree of competence yet in different domains, however, global impression formation would not have yielded a unique determinate answer.

The main findings were the following. First of all, trait knowledge did play a key role: When, in the domain-of-competence contrast, children were confronted with a strong and a knowledgeable model (who differed in domain, but not in overall degree of competence), only those children who correctly identified the smart and the strong model in their answers to trait questions showed rational patterns of model choice: they preferred the strong model for the strength tasks and the knowledgeable model for the knowledge tasks. In contrast, those children who did not correctly attribute traits chose models at chance. These findings clearly speak in favor of the general idea of rational competence accounts.

The second main finding was that trait knowledge, although presumably necessary, was not sufficient for rational inductive answer patterns. Rather, task structure played an additional key role. When presented with models that differed in competence within one domain, children (regardless of their correct or incorrect trait attributions) recruited the more competent model for all tasks above chance – irrespective of whether domains of competence and the requirements of tasks did or did not match. Note that these wide generalizations do not reflect a general halo effect since the same children who generalized widely when models differed in overall competence chose rationally when two models were both highly competent, but in different domains. Neither do these wide generalizations reflect children's general reluctance to provide indifference answers, as children frequently expressed indifference in the Reduced Suggestibility Condition, in particular in response to the preference questions. These findings are in line with previous studies in which children showed wide generalizations when there was a discrepancy in overall competence between models (Bascandziev & Harris, 2014; Brosseau-Liard & Birch, 2010; Fusaro et al., 2011).⁵

In more exploratory fashion, two further potential limiting factors in children's early performance were investigated, the pragmatics of suggestive questioning and executive function. Concerning the former, children in the Reduced Suggestibility Condition provided more indifference answers than those in the Standard Questions Condition,

yet this difference was unsystematic with regard to children's aggregated model preferences: When models differed in overall competence, children did not make less wide generalizations when asked less suggestively – showing that the wide generalizations between domains found in many selective trust studies are not just an artifact of the question format. Finally, executive function was expected to be related to rational inductive answer patterns (Jaswal et al., 2014). In the present study, however, such relations were not very strong: High levels of executive function did go along with more rational choice in cases where the two models were both competent, but in different domains (domain-of-competence contrast), yet it did not predict less wide generalizations when models differed in competence in the same domain (degree-of-competence contrasts).

All in all, thus, the present findings speak in favor of general rational competence accounts of selective learning: Children are capable, in principle, of rational inductive inferences for selective learning, yet this capacity is sometimes masked by performance factors. More generally, the present results suggest that young children have both rational and global processes at their disposal, yet use them in different types of tasks. So, what exactly might the relation between the two kinds of processes be? One interesting, yet at this stage clearly speculative possibility, to be explored in future research, is that some kind of dual-process theory might best describe the relation of simpler processes, such as global impression formation, and more rational processes, such as trait-based inductive inferences, in the realm of selective learning (see also Heyes, 2016). Despite the considerable differences in details in various dual-process theories in cognitive, social and developmental psychology, a common assumption underlying most of these accounts is the following (Evans, 2008): Type I processes are fast, operate relatively inflexibly, based on heuristic and associative (rather than inferential) processes, and are implicit, unconscious and independent from general cognitive resources. Type II processes, in contrast, are relatively slow, operate in controlled and flexible ways, are explicit and conscious, and dependent upon general cognitive resources. The present data seem compatible with such a dual-process picture in several respects: first of all, children may have used simpler heuristic processes as a default in tasks in which they yielded determinate answers (degree-of-competence contrasts), yet in tasks in which such processes did not provide definite solutions (domain-of-competence contrast) children's more rational Type II-like processes intervened. Second, children's rational performance in the domain-of-competence contrast (but not their reasoning based on global impression formation in the degree-of-competence contrasts) showed some signatures of Type II processes: children's rational answers depended upon their background knowledge (indicated in the answers to explicit trait questions), were more frequent in older children and showed some relation to executive function. But of course, the present data taken by themselves are compatible with rather than anything close to proper evidence for a dual-process account of children's early selective learning. Nonetheless, such theories may present an interesting and productive framework for future research on the cognitive foundations and development of selective learning.

NOTES

- ¹ Since these sum scores were not normally distributed we used nonparametric statistical analyses.
- ² These patterns of results for the degree-of-competence contrasts and the domain-of-competence contrast were replicated with trial-wise χ^2 analyses for all single test trials (with an exception in one trial only, for details see Supporting Information C).
- ³ Five children did not complete the HTKS task and five additional children were excluded from this analysis because they had problems understanding the instructions and received a value of zero. If the latter five children are included, the results do not change.
- ⁴ Note that this correlation failed to reach significance when children who made mistakes in the reminder questions were excluded from analysis (remaining sample: $n = 59$), $r_s = .25$, $p = .065$
- ⁵ In Fusaro et al. (2011), children showed local differences in their patterns of generalization: they were presented with either a pair of models that differed in strength (strong/weak) or in accuracy in labeling objects (accurate/inaccurate). In subsequent test trials, children preferred the accurate model in labeling tasks but not strength tasks, whereas they preferred the strong puppet in both labeling and strength tasks. The current study did not replicate such local differences in generalizations between domains. One possible reason is that in Fusaro et al. (2011) there was a confound such that the strength manipulation was more powerful than the knowledge manipulation: the weak puppet was not only weak but also inaccurate, falsely announcing 'I will lift this' whereas in the knowledge condition, models differed only in accuracy. In the current study, the contrasts between the two models in the strength and the knowledge demonstrations were deconfounded and more carefully matched, with the competent model scoring above and the incompetent scoring below the child's level of competence on just one dimension.

ACKNOWLEDGEMENTS

Thanks to Marlen Kaufmann and Konstanze Schirmer for help with recruiting children and organizing the study. We also like to thank all the children who participated as well as their parents and preschool teachers. This work was supported by the German Science Foundation (DFG), grant Ra-2155/1-1.

REFERENCES

- Bascandzief, I., & Harris, P.L. (2014). In beauty we trust: Children prefer information from more attractive informants. *British Journal of Developmental Psychology*, *32*, 94–99.
- Bernard, S., Proust, J., & Clément, F. (2015). Four- to six-year-old children's sensitivity to reliability versus consensus in the endorsement of object labels. *Child Development*, *86*, 1112–1124.
- Brosseau-Liard, P.E., & Birch, S.A.J. (2010). 'I bet you know more and are nicer too!': What children infer from others' accuracy. *Developmental Science*, *13*, 772–778.
- Cain, K.M., Heyman, G.D., & Walker, M.E. (1997). Preschoolers' ability to make dispositional predictions within and across domains. *Social Development*, *6*, 53–75.
- Ceci, S.J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, *113*, 403–439.
- Clément, F. (2010). To trust or not to trust? Children's social epistemology. *Review of Philosophy and Psychology*, *1*, 531–549.
- Corriveau, K., & Harris, P.L. (2009). Choosing your informant: Weighing familiarity and recent accuracy. *Developmental Science*, *12*, 426–437.
- Corriveau, K.H., Kinzler, K.D., & Harris, P.L. (2013). Accuracy trumps accent in children's endorsement of object labels. *Developmental Psychology*, *49*, 470–479.
- Evans, J.S.B.T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278.
- Fusaro, M., Corriveau, K.H., & Harris, P.L. (2011). The good, the strong, and the accurate: Preschoolers' evaluations of informant attributes. *Journal of Experimental Child Psychology*, *110*, 561–574.
- Harris, P.L. (2007). Trust. *Developmental Science*, *10*, 135–138.
- Hermes, J., Behne, T., & Rakoczy, H. (2015). The role of trait reasoning in young children's selective trust. *Developmental Psychology*, *51*, 1574–1587.
- Heyes, C. (2016). Who knows? metacognitive social learning strategies. *Trends in Cognitive Sciences*, *20*, 204–213.
- Jaswal, V.K., Croft, A.C., Setia, A.R., & Cole, C.A. (2010). Young children have a specific, highly robust bias to trust testimony. *Psychological Science*, *21*, 1541–1547.
- Jaswal, V.K., & Pérez-Edgar, K. (2014). Resolving conflicts between observation and testimony: The role of inhibitory control. In E.J. Robinson & S. Einav (Eds.), *Trust and skepticism: Children's selective learning from testimony* (pp. 110–122). New York: Psychology Press.
- Jaswal, V.K., Pérez-Edgar, K., Konrad, R.L., Palmquist, C.M., Cole, C.A., & Cole, C.E. (2014). Can't stop believing: Inhibitory control and resistance to misleading testimony. *Developmental Science*, *17*, 965–976.
- Kaufman, A.S., Kaufman, N.L., Melchers, P., & Preuss, U. (1994). *K-ABC: Kaufman-Assessment Battery for Children*. Amsterdam: Swets & Zeitlinger.
- Kinzler, K.D., Corriveau, K.H., & Harris, P.L. (2011). Children's selective trust in native-accented speakers. *Developmental Science*, *14*, 106–111.
- Koenig, M.A., Clement, F., & Harris, P.L. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science*, *15*, 694–698.
- Koenig, M.A., & Jaswal, V.K. (2011). Characterizing children's expectations about expertise and incompetence: Halo or pitchfork effects? *Child Development*, *82*, 1634–1647.
- Kushnir, T., Vredenburg, C., & Schneider, L.A. (2013). 'Who can help me fix this toy?' The distinction between causal knowledge and word knowledge guides preschoolers' selective requests for information. *Developmental Psychology*, *49*, 446–453.
- Landrum, A.R., Mills, C.M., & Johnston, A.M. (2013). When do children trust the expert? Benevolence information influences children's trust more than expertise. *Developmental Science*, *16*, 622–638.
- Li, Q.-G., Heyman, G.D., Xu, F., & Lee, K. (2014). Young children's use of honesty as a basis for selective trust. *Journal of Experimental Child Psychology*, *117*, 59–72.
- Lutz, D.J., & Keil, F.C. (2002). Early understanding of the division of cognitive labor. *Child Development*, *73*, 1073–1084.
- McClelland, M.M., Cameron, C.E., Duncan, R., Bowles, R.P., Acock, A.C., Miao, A., & Pratt, M.E. (2014). Predictors of early growth in academic achievement: The head-toes-knees-shoulders task. *Frontiers in Psychology*, *5*, 599.
- Mills, C.M. (2013). Knowing when to doubt: Developing a critical stance when learning from others. *Developmental Psychology*, *49*, 404–418.
- Ponitz, C.E.C., McClelland, M.M., Jewkes, A.M., Connor, C.M., Farris, C.L., & Morrison, F.J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly*, *23*, 141–158.
- Ponitz, C.C., McClelland, M.M., Matthews, J.S., & Morrison, F.J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology*, *45*, 605–619.
- Robinson, E.J., & Einav, S. (Eds.) (2014). *Trust and skepticism: Children's selective learning from testimony*. New York: Psychology Press.
- Sabbagh, M.A., & Baldwin, D.A. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development*, *72*, 1054–1070.

- Sobel, D.M., & Corriveau, K.H. (2010). Children monitor individuals' expertise for word learning. *Child Development, 81*, 669–679.
- Sobel, D.M., & Kushnir, T. (2013). Knowledge matters: How children evaluate the reliability of testimony as a process of rational inference. *Psychological Review, 120*, 779–797.
- VanderBorgh, M., & Jaswal, V.K. (2009). Who knows best? Preschoolers sometimes prefer child informants over adult informants. *Infant and Child Development, 18*, 61–71.
- Van Reet, J., Green, K.F., & Sobel, D.M. (2015). Preschoolers' theory-of-mind knowledge influences whom they trust about others' theories of mind. *Journal of Cognition and Development, 16*, 471–491.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Hermes J, Behne T, Bich AE, Thielert C, and Rakoczy H. Children's selective trust decisions: rational competence and limiting performance factors. *Dev Sci.* 2017; 00: e12527. <https://doi.org/10.1111/desc.12527>