Making sense of conflicting information: A touchscreen paradigm to measure young children's selective trust

Jonas Hermes¹,² | Hannes Rakoczy¹,² | Tanya Behne¹,²

¹Institute of Psychology, Department of Developmental Psychology, University of Göttingen, Göttingen, Germany
²Leibniz ScienceCampus Primate Cognition, Göttingen, Germany

Abstract
Much recent research has shown that children from age 4 onwards reveal a robust preference for reliable over unreliable informants when choosing whom to trust and learn from. Findings concerning selective model choice in children younger than 4 years have mostly been mixed. The present study developed a new touchscreen-based paradigm with reduced task demands in order to test 2- and 3-year-old children (N = 48). Results showed that 3-year-olds selectively endorsed information from a previously reliable rather than a previously unreliable informant when searching for objects whereas 2-year-olds just followed the first hint even if provided by an unreliable informant. Whether the lack of selective model choice in 2-year-olds reflects competence or performance deficits remains to be clarified. But the present results do suggest that 3-years-olds have the basic competence to selectively choose reliable over unreliable informants that may have been masked in some previous studies by task demands.

Highlights
- The paper develops a novel touchscreen-based search paradigm to test young children's selective trust with reduced task demands.
- Three-year-olds chose selectively between conflicting hints, whereas 2-year-olds followed the first hint even if provided by an unreliable informant.

KEYWORDS
selective trust, social cognition, social learning
In their first years of life, young children learn much, if not the most, socially via testimony of others. Much recent research has shown that children do not trust anyone alike but are fairly selective in whom they trust. When confronted with two possible sources of information, children from 4 years onwards show robust capacities to accept new information selectively from, for example, the more reliable, the more knowledgeable, or the nicer informants (for reviews, see Harris, 2012; Mills, 2013; Robinson & Einav, 2014). But how do younger children choose between informants?

Research with children younger than 3 years generally used a one-informant design with each child encountering either a reliable or an unreliable informant. When responses to reliable versus unreliable informants are compared between subjects, recent research suggests that young children seem to trust the reliable informants more than the unreliable ones. Even infants followed the gaze of a face more often when it had previously proven to be reliable rather than unreliable (Tummeltshammer, Wu, Sobel, & Kirkham, 2014), and they looked longer at a person who labelled known things incorrectly than at one who did so correctly (Koenig & Echols, 2003). Young children imitated the actions of a model more often when the model had used familiar objects in conventional rather than unconventional ways (Zmyj, Buttelmann, Carpenter, & Daum, 2010), when the model had labelled known things accurately rather than inaccurately (Brooker & Poulin-Dubois, 2013), or when the model had previously shown appropriate affect (Poulin-Dubois, Brooker, & Polonia, 2011; Zmyj et al., 2010). Furthermore, children learned labels for novel objects at higher rates when the informant had previously labelled familiar objects accurately rather than inaccurately (Brooker & Poulin-Dubois, 2013; Koenig & Woodward, 2010). Thus, even young children's responses appear to be affected by the reliability of the informant they encountered. This could mean that young children already track the past behaviour of specific individuals and adjust their future responses accordingly. Alternatively, it could be that children responded to features of the general situation (i.e., learning more in "normal" situations than in "weird" ones) without tracking the behaviour of specific individuals.

Note that the one-informant design used in these studies differs from the standard two-informant, within-subject design that has been typically used to study preschoolers' selective trust (Mills, 2013). In the standard selective trust paradigm, children are confronted with a reliable and an unreliable informant who provide conflicting novel information. For example, children witness each informant offer a different label for a novel object, and then they are asked what they think the object is called. If children endorse and acquire the label provided by the reliable rather than the unreliable informant, it is possible to conclude that they must have tracked the past behaviour of specific individuals and selectively learn from the more reliable one. In contrast in the one-informant design, children are either, depending on condition, confronted with a relatively everyday situation in which they experience a person who labels known objects with appropriate terms and then uses a novel label to refer to a novel object. Or they are just confronted with a quite bizarre situation in which they, for example, experience a person consistently mislabelling known objects. If they learn less in the latter compared with the former situation, it could either be that they specifically learn less from unreliable informants, or it could be simply that they generally learn less in more bizarre circumstances. In one-informant, between-subject designs, this bizarreness selectively affects the unreliable but not the reliable conditions and can thus explain differences between conditions. In two-informant, within-subject designs, the bizarreness of someone mislabelling known objects does not affect conditions selectively because all children witness the demonstration of the unreliable informant. Therefore, in the two-informant design, selective learning from the reliable informants cannot be explained by differences between conditions and thus serves as a much clearer indication for children's understanding of more and less reliable informants.

In fact, such standard two-informant designs have been used with 3-year-old and older children, but the 3-year-olds' findings have been mixed. Some studies have found that 3-year-olds do already choose selectively between models providing conflicting information, trusting, for example, an accurate or knowledgeable informant over an inaccurate or ignorant one (Birch, Vauthier, & Bloom, 2008; Corriveau & Harris, 2009; Corriveau, Meints, &
Harris, 2009; Jaswal & Neely, 2006; Koenig, 2012; Koenig & Harris, 2005, exp. 2). Other studies using a two-informant design, however, have found that 3-year-olds do not selectively choose between the two informants when confronted with, for example, a previously accurate and an inaccurate (Clément, Koenig, & Harris, 2004; Koenig & Harris, 2005, exp. 1), an accurate and a neutral (Corriveau et al., 2009) or an honest and a deceptive (Li, Heyman, Xu, & Lee, 2014) informant.

So how can we make sense of and reconcile these findings both regarding the mixed results with the 3-year-olds and the divergence that arises when one compares these findings with 3-year-olds to the seemingly more robust findings with younger children? One possibility is that children are able to selectively choose reliable over unreliable informants from a young age but that high task demands have sometimes masked their competence (see also Clément, 2010). First, demands on children's working memory may have masked children's competence in some studies with 3-year-olds: When the informants' location (i.e., whether they appeared on the left or the right side of the screen) varied across trials, 3-year-olds did not choose selectively (Koenig & Harris, 2005), whereas when informants' location were held constant, they did choose selectively (e.g., Corriveau & Harris, 2009). And more generally, studies with two informants (conducted with 3-year-olds), in which children need to track and remember the reliability of two models, obviously pose higher demands on their working memory than studies with one model only (conducted with younger children). This is in line with the pattern that children have shown less consistent selective trust in studies with two informants (e.g., Clément et al., 2004; Koenig & Harris, 2005) than those with one informant (see e.g., Koenig & Woodward, 2010, for positive findings with 2-year-olds), and it is probably the reason why children younger than 3 years were mostly confronted with a single informant only.

Second, verbal demands might have sometimes masked children's competence. Another difference between the methods used in studies with 3-year-olds and those with younger children is that 3-year-olds were usually required to provide explicit verbal answers (and sometimes did not choose selectively, e.g., Corriveau et al., 2009; Koenig & Harris, 2005), whereas studies indicating potential comprehension of model reliability in even younger children used non-verbal measures, for example, object choice (Koenig & Woodward, 2010) or looking times (Tummeltshammer et al., 2014).

Third, high demands on inhibitory control may have masked children's competence. In fact, 2- and 3-year-olds often show high rates of trust in the verbal information of a single unreliable source (e.g., Jaswal et al., 2010; Krogh-Jespersen & Echols, 2012), even after repeated instances of misleading information. It has been argued that information is automatically accepted temporarily before it might be rejected (Gilbert, 1991); therefore, it needs cognitive effort, probably inhibitory control, to override this temporary acceptance. And indeed, children with higher inhibitory control show lower rates of trust in previously unreliable informants (Jaswal et al., 2014, but see also Heyman, Sritanyaratana, & Vanderbilt, 2013).

Against this background, the aim of the present study was to systematically explore 2- and 3-year-olds' capacity to selectively choose between reliable versus unreliable informants using the same method for both age groups. To this end, we developed a novel test with radically reduced task demands, combining the advantages of methods previously used with 2-year-olds (reduced verbal demands) with the advantages of methods used in older children (less ambiguous interpretation due to the two-informant design). In a very simple procedure, children were confronted with and had to choose between two differentially reliable informants who made conflicting claims about the location of objects to be searched. This builds on an established search paradigm in which children from 12 months readily take into account the communicative cues of single informants in locating objects (Behne, Liszkowski, Carpenter, & Tomasello, 2012). We measured children's direct search behaviour, thereby reducing the verbal task demands. In order to additionally reduce the inhibitory task demands posed by the need to ignore the communicative cue provided by one of the informants, we used such types of cues that have been found in previous work to elicit the least bias to trust (less than language and deictic gestures), namely, placing markers at target locations (Couillard & Woodward, 1999; Jaswal et al., 2010). Because we needed children to discriminate between two markers, we used iconic markers (similar as the use of pictures in Palmquist, Burns, & Jaswal, 2012).
In the present study, children saw videos with two informants on a touchscreen device. In demonstration trials, one informant proved reliable by providing correct cues to the locations of hidden objects, whereas the other proved unreliable by providing incorrect cues. In subsequent test trials, both informants provided conflicting cues to the locations of hidden objects, and children were encouraged to actively search for the targets by exploring a location via touching it on a touchscreen. The crucial dependent measure was whether children searched at the locations cued by the previously reliable rather than the unreliable informant.

2 | METHOD

2.1 | Participants

Forty-eight children in three age groups (30–35 months, \(M = 32.1\); 36–41 months, \(M = 39\); 42–47 months, \(M = 44.6\); \(n = 16\) each) were recruited from a database of parents who had volunteered to participate in studies on child development and came from mixed socio-economic backgrounds. Parents gave their written consent for the participation of their children. Five additional children were tested but excluded from the analysis due to experimenter error (\(n = 2\)), technical error (\(n = 1\)), problems with the procedure due to insufficient language abilities (\(n = 1\)), or failure to wait for the hints by the informants before choosing (\(n = 1\)).

2.2 | Procedure

Children were tested either in quiet separate rooms in their day-care centres or in the child lab. As a warm-up, the child and experimenter played together with toy animals. Then in a short cue-familiarization game, children were familiarized to the types of cues, provided by the informants later in the main study. In the main study itself, children saw six demonstration trials on a tablet computer in which the respective reliability of two informants was established, followed by eight test trials in which children chose which of the two informants to trust about the location of a hidden object by tapping on the respective locations on the touchscreen.

2.2.1 | Cue-familiarization game

The purpose of this game was to make children familiar with the nonconventional ways in which the informants cued the locations in the main study. In this game, the experimenter took toy animals from the former warm-up play and hid them under one of two cups occluded by a barrier. Then both cups were placed in front of the child, and the experimenter placed an iconic marker (a lego figure) on the box where the animal was hidden, before the child was allowed to search for the animal. There were three different types of such search trials. At first, the experimenter said “The [animal] hid under one of these two cups. And I have seen it. This is where it hid” while placing the object. This was to make explicit the communicative purpose of placing the object. Children participated in two such trials, and if they made a mistake searching, the trial was repeated until they searched correctly in two subsequent trials, with a maximum of four such trials altogether. Then there was one trial, in which the experimenter initially placed the object on the wrong location saying the same as above but then immediately correcting himself saying “… no! Nonsense! This is where it hid!”, placing the object to the other cup. This was to open the possibility of wrong hints. Finally, in the last two trials, the experimenter did not say where the target was while placing the object. (Thus, in total, each child participated in five to seven search trials in the cue-familiarization game.)
2.2.2 | Main study

For the main study, a 10″ touchscreen tablet was placed in front of the child. First, children saw the two informants, a duck and a sheep, appear (order counterbalanced), each placing a miniature of itself in front of it and introducing themselves, saying, for example, "Hi, I am the duck. Look, I also have a little duck."

Each subsequent demonstration trial or test trial started with a hiding event: Children first saw two cups at the bottom of the screen, which became occluded by an ascending barrier, then a butterfly appeared from the top of the screen and flew across both sides of the screen until finally disappearing behind the barrier. Then there was a click sound, indicating that the butterfly moved under one of the cups, and the barrier descended revealing the two cups and no visible butterfly. The experimenter commented "oh, look, the butterfly hid in one of the two cups!" (see Figure 1 for the procedure of the hiding event, the demonstration trials, and the test trials).

The demonstration phase consisted of six trials. In each of the first four demonstration trials, children saw one of the informants providing cues to the hiding location of the butterfly. This was supported by verbal testimony to make sure that children understood the meaning. Specifically, after the hiding event, both informants appeared in the middle of the screen and one informant went to a location and placed its miniature on top saying "Hello! ... It's here!" The two informants alternated, with each providing cues on two of the four trials (who started was counterbalanced across participants). One informant constantly provided hints to the correct location, whereas the other one constantly provided incorrect hints. Subsequently the cups automatically accelerated thus revealing the location of the butterfly. The child was asked whether the informant was right and received feedback on her answer. After these first four trials, the child was asked an explicit judgement question ("Who was good at this, the sheep or the duck?") and received feedback on her answer (the experimenter agreed with correct answers and doubted wrong answers).

In the last two demonstration trials, two aspects from the test trials were already introduced for reasons of practice: Children were allowed to search for the butterflies themselves by touching the intended location, and they were given an incentive to search correctly (i.e., a sticker of the butterfly was given as a reward if they searched successfully). They saw one more trial of the reliable informant cueing the correct location and one more trial with the unreliable informant cueing the incorrect location (in the same counterbalanced order as above). When children touched a location, the cup automatically lifted and revealed either the butterfly or nothing. If the search was successful, children received the sticker of the butterfly and put it into their album and the experimenter opened the second location to show that it is empty. If the search was unsuccessful, the sticker was not given to the child but put into a box and the experimenter opened the second location to reveal the butterfly. Children no longer received explicit feedback from the experimenter.

The test phase consisted of eight trials. Children again first saw the hiding event (same as above) before both informants appeared one after the other saying "hello" and cueing conflicting locations with their replicas while providing no verbal information. The children were then encouraged to search the location where they expected the butterfly to be via touch. If they found a butterfly, they received the according sticker. If the search was unsuccessful, the sticker was put into the box and the experimenter opened the second location to reveal the butterfly. After the last test trial, the box was opened and children received additional stickers.

2.3 | Counterbalancing

The duck always appeared in the left centre, and the sheep always appeared in the right centre, yet the identity of the reliable and the unreliable informant was counterbalanced across participants. The order in which the two informants provided their hints and the butterfly's hiding location was counterbalanced across trials, for each the demonstration and test trials. Thus, in half of the trials, the informants moved across the screen and indicated the location opposite to the side where they initially appeared.
2.4 | Coding

For each test trial, “1” was coded if children searched successfully at the location cued by the reliable informant and “0” was coded if children searched unsuccessfully at the location cued by the unreliable informant. For further analysis, the mean proportion of successful searches across the eight trials was calculated. Five children searched before both informants had provided their hints in one trial each. These trials were excluded from analysis; therefore, for these children, the mean proportion of successful trials was based on seven trials only.

3 | RESULTS

We were interested in children's search patterns when confronted with conflicting information from a reliable and an unreliable informant. Children searched for the target at the locations cued by the previously reliable informant at above chance levels (one-sample Wilcoxon signed-rank test on the mean proportion of successful trials, Z = 3.0,
Their rates of following the reliable model's cues did not improve across the eight test trials (Friedman test, $\chi^2(7) = 11.34, p = .13$) and did not differ between age groups (Kruskal–Wallis H test, $\chi^2(2) = 3.08, p = .21$).

Because we were specifically interested in whether this preference for the previously reliable model really held for all age groups, we performed planned analyses for each age group separately. Three-year-olds selectively searched at the location cued by the previously reliable informant (younger 3-year-olds [36–41 months]: $Z = 2.51, p < .05, r = 0.63$; older 3-year-olds [42–47 months]: $Z = 2.14, p < .05, r = 0.54$), whereas 2-year-olds searched at chance level (30–35 months: $Z = 0.21, p = .84$). Figure 2 shows children's preference for the location cued by the previously reliable informant, separated by age groups.

After the first four demonstration trials, children were asked an explicit judgment question ("Who was good at this, the sheep or the duck?"). Table 1 shows children's answers to this question. We replicated the analyses with the subsample of those 34 children who had answered the explicit judgment question correctly and found the same pattern of results as in the whole sample: 3-year-olds selectively searched at the location cued by the previously reliable informant (younger 3-year-olds [36–41 months]: $Z = 2.33, p < .05, r = 0.67$; older 3-year-olds [42–47 months]: $Z = 2.11, p < .05, r = 0.61$). Crucially, even those ten 2-year-olds who had answered the explicit judgment question correctly still searched at the locations cued by the reliable and the unreliable informants exactly at chance level ($Z = 0, p = 1$).

To assess whether children generally understood the use of iconic markers as cues for target locations, we measured whether they searched at the indicated locations in the last two trials of the cue familiarization game (conducted before the main study started), in which the experimenter said nothing while placing the marker. The data

![Figure 2](image_url) - Children's choices separated by age groups. Children's mean proportion of searches at the locations cued by the previously reliable informant, separated by age groups. Asterisks indicate deviation from chance level (Wilcoxon signed-rank tests, *$p < .05$). Error bars show standard errors.

<table>
<thead>
<tr>
<th>Age group</th>
<th>Reliable informant</th>
<th>Unreliable informant</th>
<th>No answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>30–35 months</td>
<td>10</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>36–41 months</td>
<td>12</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>42–47 months</td>
<td>12</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Sum</td>
<td>34</td>
<td>10</td>
<td>4</td>
</tr>
</tbody>
</table>

*Table 1* Answers to the explicit judgment question by age group.
of two children could not be analysed because the video capture did not start before the main study. Due to an experimenter error, 10 children received only one (instead of two) trials in which the experimenter said nothing while placing the marker. The proportion of correct answers was above chance level for both 2-year-olds (87%, $Z = 3.05$, $p < .01$, $r = 0.79$) and 3-year-olds (younger 3-year-olds: 91%, $Z = 3.36$, $p < .01$, $r = 0.84$; older 3-year-olds: 90%, $Z = 3.21$, $p < .01$, $r = 0.83$).

Finally, we investigated whether children simply followed the first advice they saw—regardless of whether it was provided by the previously reliable or the previously unreliable informant. Three-year-olds did not simply follow the first cue provided at levels above chance (younger 3-year-olds: 51%, $Z = 0.10$, $p = .92$; older 3-year-olds: 48%, $Z = 0.68$, $p = .50$). Two-year-olds, however, did follow the first cue more often than expected by chance (59%, $Z = 2.18$, $p < .05$, $r = 0.55$).

4 | DISCUSSION

The current study explored 2- and 3-year-olds’ capacity to selectively choose between reliable versus unreliable informants in their active search behaviour in a novel test with reduced task demands (simple procedure with reduced inhibitory and verbal demands). The main result was that 3-year-olds, but not 2-year-olds, selectively searched for targets at places cued by a previously reliable informant (see also Ganea, Koenig, & Millett, 2011, for a similar set of findings). Instead, 2-year-olds followed the first cue even if it was provided by the unreliable informant.

The positive findings concerning 3-year-olds are informative in light of inconsistent previous findings and suggest that children this age have the competence in principle to choose selectively between informants. One possible reason for these positive findings is that in the present study, the task demands were suitably reduced. Most importantly, children expressed their answers via direct search behaviour rather than via verbal responses to questions. Moreover, the method of cueing locations with iconic markers might have reduced inhibitory demands. Additionally, in the present study, children had an incentive to choose the correct location. Yet these positive findings (similar as in much early research on selective trust) do not clearly tell us whether children really attributed knowledge to the reliable, but not the unreliable informant. Instead, more basically, children may have chosen on the basis of an undifferentiated, positive global impression of the reliable informant (Hermes, Behne, & Rakoczy, 2018), or they could have associated the reliable informant with success or avoided the somewhat strange (unreliable) informant and chose the reliable one by exclusion. But the present results do tell us that from age 3, children form some sort of selective associations on the basis of the individual informants’ past behaviour and are able to use these as a basis for their future decisions. Further research is needed to identify the cognitive foundations of such early selectivity more clearly.

But what about the negative findings regarding the 2-year-olds? One possibility is that these results reflect a true competence deficit in the sense that 2-year-olds are incapable of actively selecting between conflicting sources of information. If that is the case, the early indicators for an understanding of differentially reliable models shown in previous studies with single informants (and in between-subject comparisons across conditions) may have tapped different and more basic cognitive capacities than selective trust. Rather, the results that children often followed the information provided by an unreliable informant less than that provided by a reliable informant may have arisen out of differences in rather diffuse feelings of (un)certainty and (in)consistency: Children in the reliable condition were repeatedly confronted with accurate and reliable information and thus presumably felt certain and learnt and imitated at rates they usually do. Children in the unreliable condition, in contrast, after being repeatedly exposed to inaccurate information or bizarre behaviour, may have felt uncertain and lost and may have been inhibited to act as they usually do. More systematic comparisons of the different methods involving two competing informants versus one single informant are needed to investigate such possibilities.

Alternatively, the unconventional way of cueing locations (by using iconic markers) may have been problematic. These cues were used in order to reduce inhibitory demands, but this may have posed other demands in the sense
that the cues were more difficult to understand in the first place. It seems unlikely, however, that children had difficulty understanding the communicative act of using the markers as such, in light of previous research and in light of the present findings. On the one hand, previous research has shown that from 30 months or earlier, children understand markers as communicative gestures in similar tasks (Tomasello, Call, & Gluckman, 1997; Zlatev et al., 2013). And in the present study, children's above-chance successful searches in the cue-familiarization game (when only one cue was provided per trial) indicate that children understood the iconic markers as communicative gestures. On the other hand, the cue familiarization game was done interactively, whereas in the test trials, children needed to understand the iconic markers in a computer-based setting. This transfer from a real-life context to a computer context might have been more difficult for 2-year-olds than for 3-year-olds (see also Barr, 2010).

Yet another potential limiting factor of the present task relates to the types of mistakes made by the unreliable informant. These mistakes consisted in giving inaccurate episodic information about the current location of a target and children were provided no information about whether the informants had any access to the correct information. Therefore, the unreliable informant might have been perceived merely as a poor guesser rather than as someone who disregards potential individual or common knowledge. Interestingly, a recent study has shown that 3- and 4-year-olds excuse episodic errors more readily than semantic errors (Stephens & Koenig, 2015). In that study, children saw two informants who were (in)correct either about locations (episodic condition) or labels (semantic condition) of objects. In later test trials, children could ask for and endorse information from these informants concerning the locations or labels of objects. Children in the semantic condition showed a much clearer preference for the reliable informant than did those in the episodic condition. Interestingly, the 3- and 4-year-olds in the episodic condition did not endorse information about the location of novel objects selectively from the reliable model (although they selectively asked her), in contrast to the 3-year-olds' selectivity in the present study where task demands were reduced. It is thus possible that the 2-year-olds would show more selectivity in response to semantic rather than episodic (un)reliability.

More generally, the demands on children's working memory when being presented with two informants simultaneously might generally be too high for 2-year-olds, which is probably the reason why such designs have rarely been applied for 2-year-olds. Possibly, young children succeeded in identifying the more reliable informant in the demonstration phase. But when they had to choose between conflicting advice in the test trials, they may have failed to use this information due to limited working memory and instead fell back on a simpler strategy and just endorsed the first cue that was provided.

All in all, the present study shows that the competence to actively select between reliable and unreliable informants is present at age 3. Future research, though, will need to test more systematically, and possibly with even more simplified tasks, between competence and performance explanations of the negative findings with the 2-year-olds.

ACKNOWLEDGEMENTS
Thank you very much to Marlen Kaufmann and Konstanze Schirmer for help with recruiting children and organizing the study and to Raphaela Wandt for help with conducting the study. We also like to thank all children who participated as well as their parents and preschool teachers. This work was supported by the German Science Foundation (DFG) (Grant Ra-2155/1-2).

ENDNOTE
1 Note that one overall finding from studies using this design is that children generally tend to trust novel information, even from unreliable sources and even after multiple instances of misleading information (e.g., Jaswal, Carrington Croft, Setia, & Cole, 2010; Krogh-Jespersen & Echols, 2012; Vanderbilt, Heyman, & Liu, 2014). Yet, when reliable and unreliable conditions are compared between subjects, the rates often differ.
REFERENCES


How to cite this article: Hermes J, Rakoczy H, Behne T. Making sense of conflicting information: A touchscreen paradigm to measure young children’s selective trust. Inf Child Dev. 2019;28:e2119. https://doi.org/10.1002/icd.2119