JOURNAL OF COGNITION AND DEVELOPMENT
2019, VOL. 20, NO. 4, 555–572
https://doi.org/10.1080/15248372.2019.1631167

Routledge
Taylor & Francis Group

EMPIRICAL ARTICLE

Learning the Relevance of Relevance and the Trouble with Truth: Evaluating Explanatory Relevance across Childhood

Angie M. Johnston, Mark Sheskin, and Frank C. Keil

Yale University

ABSTRACT

In four experiments, we investigate how the ability to detect irrelevant explanations develops. In Experiments 1 and 2, 4- to 8-year-olds and adults rated different types of explanations about "what makes cars go" individually, in the absence of a direct contrast. Each explanation was true and relevant (e.g., "Cars have engines that turn gasoline into power"), true and irrelevant (e.g., "Cars have radios that play music"), or a false statement that would be relevant if it were true (e.g., "Cars have rockets that speed them up"). Participants of all ages spontaneously indicated that false explanations were less helpful than relevant explanations. However, there was a developmental shift for irrelevant explanations: 4-year-olds only detected irrelevant explanations that did not involve internal features of cars (e.g., "Cars have parking lots that they park in"). Crucially, this shift between age 4 and 5 cannot be explained by 4-year-olds' lack of knowledge since 4-year-olds correctly indicated that relevant explanations were more helpful than irrelevant feature explanations when given a direct contrast in Experiment 3. These results are further clarified in Experiment 4, in which we provided a different explanatory goal ("where to find cars") and found that even young children have a nuanced understanding of explanatory relevance that is sensitive to differing explanatory goals. Together, these four experiments suggest an early-emerging ability to understand relevance, but a shift between age 4 and 5 in the ability to spontaneously use this understanding when evaluating individual explanations in isolation.

ARTICLE HISTORY
Received 12 Jul 2017
Accepted 09 Jun 2019
Revised 04 Jun 2019

Children learn about the world not only through direct observation and exploration (e.g., Keil, 2011; Schulz & Bonawitz, 2007), but also through information provided by others (e.g., Gelman, 2009; Koenig & Harris, 2005). Although relying on others for information allows for highly efficient learning (e.g., Csibra & Gergely, 2011), it can also leave the learner susceptible to misinformation. To guard against misinformation, children must evaluate whether information they receive is accurate and helpful. One way children do this is by evaluating the *source* of information. For instance, even preschoolers can compare two sources and understand that certain characteristics (e.g., familiarity, age, niceness) indicate that one source is likely to be more trustworthy than another (e.g., Corriveau & Harris, 2009; Jaswal & Neely, 2006; Johnston, Mills, & Landrum, 2015).

Relatively less is known about how children evaluate isolated pieces of information, independent of the source's characteristics (e.g., Mills, 2013). This topic is important

CONTACT Angie M. Johnston angie.m.johnston@gmail.com Department of Psychology, Yale University, 2 Hillhouse Ave, New Haven, CT 06511

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hjcd.

because much of the information children receive comes with minimal context. Thus far, previous research has shown that children can evaluate information directly by considering how it aligns with their background knowledge. For example, even 2-year-olds will say "no" if an adult mislabels a familiar object (Koenig & Echols, 2003) and will choose to learn information from someone who has accurately labeled familiar objects in the past (Koenig & Woodward, 2010). Indeed, children prioritize accuracy so highly that they favor it over a wide range of other characteristics that they would typically use when evaluating a source, including the source's age (Jaswal & Neely, 2006), familiarity (Corriveau & Harris, 2009), and accent (Corriveau, Kinzler, & Harris, 2013).

If accuracy can so easily override otherwise compelling source characteristics, perhaps children will also focus on accuracy even when it is not the most useful characteristic for evaluating information. Specifically, accuracy does not always guarantee that information will be helpful. In order for information to be helpful it must also be *relevant* (e.g., Grice, 1975; Sperber & Wilson, 1996). For instance, the true statement "cars have radios that play music" might be helpful when trying to explain why there is music blaring out on the street, but not when trying to explain what makes cars move.

In some contexts, children understand the importance of relevance and can detect when information is irrelevant. For instance, when learning novel information about the location of a hidden sticker, 3- to 5-year-olds trust a puppet that provides relevant information (e.g., "The sticker is under the red cup") over one that provides irrelevant information (e.g., "I like cups;" Eskritt, Whalen, & Lee, 2008). Additionally, by age 4 children can detect when information is relevant for their everyday communication, for example by selectively learning information about objects that originate nearby, rather than faraway (Henderson, Sabbagh, & Woodward, 2013). These studies suggest that preschoolers are sometimes able to evaluate whether or not information is relevant.

However, in each of these prior studies, children have been asked to evaluate the relevance of *novel* information. When information is novel, children do not have any way to evaluate its truth based on their background knowledge. In contrast, when information is familiar and aligns with children's background knowledge, they may have more difficulty evaluating whether it is relevant. To evaluate relevance, children must be able to understand that information can be unhelpful even if it is true. Returning to the example presented earlier, children need to look beyond the truth-value of a statement like "cars have radios that play music" in order to evaluate whether or not it is relevant for a particular explanatory goal (e.g., understanding what makes cars move). Given that prior work has shown that children are highly focused on whether or not information fits with their background knowledge (e.g., Corriveau & Harris, 2009; Corriveau et al., 2013; Jaswal & Neely, 2006; Koenig & Echols, 2003), it is possible that they will struggle to look beyond their background knowledge in order to instead determine whether or not it is relevant for a particular explanatory goal.

Alternatively, children might be able to detect irrelevant explanations, even when they contain true information. Children have demonstrated a sophisticated set of criteria for evaluating explanations that often closely parallel the criteria used by adults (e.g., Bonawitz & Lombrozo, 2012; Corriveau & Kurkul, 2014; Frazier, Gelman, & Wellman, 2009, 2016; Johnston, Johnson, Koven, & Keil, 2017; Johnston, Sheskin, Johnson, & Keil, 2018; for a review see Lombrozo, 2016). For instance, when contrasting circular explanations (e.g., "It rains because water falls from the sky and gets us wet") with noncircular explanations (e.g., "It rains because the clouds fill with water and get too heavy"), children as young as 3 prefer noncircular explanations (Corriveau & Kurkul, 2014; see also Baum, Danovitch, & Keil, 2008; Mills, Danovitch, Rowles, & Campbell, 2017). Likewise, when contrasting simple explanations that invoke fewer causes with complex explanations that invoke more causes, children as young as 4 prefer simple explanations (Bonawitz & Lombrozo, 2012). Given this propensity to evaluate explanations based on criteria such as circularity and simplicity, young children might likewise evaluate explanations based on relevance, and apply it to differentially evaluate even explanations they know to contain factually correct information.

In the current studies, we investigate how the ability to detect true but irrelevant explanations develops. In four experiments, we provided 4-year-olds, 5- to 6-year-olds, 7- to 8-year-olds, and adults with explanations about "what makes cars go" (in Experiments 1–3) or "where to find cars" (Experiment 4) and asked them to evaluate the helpfulness of the potential explanations. In Experiment 1, we presented participants with three types of explanations – relevant causal explanations (e.g., "Cars have engines that turn gasoline into power"), irrelevant feature explanations (e.g., "Cars have radios that make music"), and false explanations that would be relevant if they were true (e.g., "Cars have rockets that speed them up") – and asked participants to rate each explanation one at a time. Based on prior work (e.g., Koenig & Echols, 2003), we anticipated that participants in all age groups would rate false explanations as being less helpful than relevant explanations. However, given that children often focus on accuracy at the expense of other characteristics (e.g., Corriveau & Harris, 2009; Corriveau et al., 2013; Jaswal & Neely, 2006), it was less clear whether children would provide different ratings to true explanations depending on each explanation's relevance.

In Experiments 2 and 3, we modified the task from Experiment 1 to give children and adults the best chance possible of detecting irrelevant explanations. In Experiment 2, we provided participants with the same relevant causal and irrelevant feature explanations from Experiment 1, but we replaced the false explanations with a different type of true but irrelevant explanation: irrelevant extrinsic explanations (e.g., "Cars have parking lots that they park in"). Given that the irrelevant extrinsic explanations we included did not refer to any internal or constitutive components of cars, we anticipated that they would be easier to detect than the irrelevant feature explanations we included.

In Experiment 3, we asked participants to make a forced-choice between relevant causal and irrelevant feature explanations, rather than asking them to rate the explanations one-by-one. Given that prior work in trust in testimony has shown that children are more critical of unreliable sources when they are presented in direct contrast with reliable sources (e.g., Krogh-Jespersen & Echols, 2012; Vanderbilt, Heyman, & Liu, 2014), we anticipated that young children might be more likely to detect irrelevant feature explanations when directly contrasted with relevant causal explanations in Experiment 3.

Finally, in Experiment 4, we asked participants to rate our explanations based on how helpful they were for a different explanatory goal: rather than asking how helpful they were for understanding "what makes cars go" we asked them to rate how helpful they were for understanding "where to find cars." This allowed us to verify that children were sensitive to the relevance of the explanations for the specific explanatory goal they were asked to consider, and that no other feature of our stimuli were driving our results. Specifically, correct performance was the reverse of the previous studies: the explanations that were previously relevant (e.g., "Cars have engines that turn gasoline into power") were now irrelevant, and the previously irrelevant extrinsic explanations (e.g., "Cars have parking lots that they park in") were now relevant. Based on previous work exploring children's evaluation of mechanistic explanations (Lockhart, Chuey, Kerr, & Keil, 2019), we anticipated that children would be sensitive to the explanatory goal and rate the previously irrelevant extrinsic explanations as being more helpful than the previously relevant feature explanations.

Experiment 1

Method

Participants

Participants were 20 4-year-olds (M_{age} = 4;6; SD_{age} = 4 months; 11 females), 20 5- to 6-year-olds (M_{age} = 5;9; SD_{age} = 3 months; 13 females), 19 7- to 8-year-olds (M_{age} = 7;11; SD_{age} = 4 months; 8 females), and 20 adults (M_{age} = 28; SD_{age} = 8.40; 7 females, recruited from Amazon Mechanical Turk).

Procedure

Participants were told that they would be presented with several explanations about cars and asked to indicate how helpful they were for "understanding what makes cars go." Explanations were either true statements that were relevant to the explanatory goal, true statements that were irrelevant to the explanatory goal, or false statements that *would* be relevant to the explanatory goal if they were true. All of the explanations were formatted in a consistent way; specifically, "Cars have ______s that ______." For instance, "Cars have engines that turn gasoline into power" (relevant), "Cars have radios that play music" (irrelevant), and "Cars have rockets that speed them up" (false). There were four explanations in each category (see Appendix A for a complete list). In order to help reduce the memory load for children, each explanation was accompanied by a silhouette image of the car part featured in the explanation (see Appendix A for silhouette images). Additionally, children were reminded of the explanatory goal (i.e., "understanding what makes cars go") for each item.

Explanations were presented individually, and participants were asked to rate how helpful each explanation was for "understanding what makes cars go." Adult participants were asked to indicate their helpfulness ratings for each explanation on a 4-point Likert scale ranging from "Not Helpful at All" to "Very Helpful." A 4-point judgment was elicited from child participants using a series of two forced choice questions (based on the two-step judgment method originally developed by Harter, 1982). Specifically, children were first asked if each explanation was "helpful or not helpful for understanding what makes cars go." If the child responded that the explanation was helpful, then the experimenter followed up by asking if it was "really helpful or a little helpful." In contrast, if the child responded that the explanation was not helpful, the follow-up question was "not helpful at all or a tiny bit helpful." This procedure allowed us to compare helpfulness ratings across age groups without overwhelming the younger children in our sample with a 4-point scale. For adults, the order of explanations was fully randomized using Qualtrics Survey Software, and for children, a single random order was presented to each participant, with half receiving the pre-established order in reverse.

If children are able to spontaneously detect irrelevant feature explanations, then they should rate the irrelevant feature explanations more similarly to false explanations than to relevant causal explanations.

Results and discussion

Our findings in Experiment 1 revealed a shift between ages 6 and 7 in children's ability to spontaneously detect irrelevant feature explanations in the presence of false explanations. Before age 7, children indicated that irrelevant feature explanations were just as helpful as relevant causal explanations, and after age 7, children indicated that irrelevant feature explanations were just as unhelpful as false explanations.

Specifically, an omnibus ANOVA showed an interaction between age (4-year-olds, 5- to 6-year-olds, 7- to 8-year-olds, and adults) and explanation type (relevant causal, irrelevant feature, and false) for participants' explanation ratings, $F(6, 150) = 4.89$, $p < .001$, $\eta^2 = .16$. We followed up on this interaction with paired-samples *t*-tests comparing each age group's ratings of the three explanation types. We found that although children in each age group rated relevant causal explanations as more helpful than false explanations (*t*s > 3.45, *p*s < .004, *d*s > .77), it was not until age 7 that children began to distinguish irrelevant feature explanations from relevant causal ones. Specifically, 4-year-olds did not distinguish between relevant causal and irrelevant feature explanations in their helpfulness ratings, $t(19) = .57$, $p = .575$, $d = .13$, and instead indicated that irrelevant feature explanations were more helpful than false explanations, $t(19) = 2.69$, $p = .015$, $d = .60$. The same was true of 5- to 6-year-olds, who gave similar ratings to relevant causal and irrelevant feature explanations, $t(19) = .22$, $p = .828$, $d = .05$, and indicated that irrelevant feature explanations were more helpful than false explanations, $t(19) = 3.26$, $p = .004$, $d = .74$. In contrast, 7- to 8-year-olds indicated that irrelevant feature explanations were *less* helpful than relevant causal explanations, $t(18) = 5.05$, $p < .001$, $d = 1.16$, and did not indicate that irrelevant feature explanations were any more helpful than false explanations, $t(18) = .51$, $p = .615$, $d = .12$. Likewise, adults rated irrelevant feature explanations as less helpful than relevant causal explanations, $t(19) = 5.23$, $p < .001$, $d = 1.17$, and did not distinguish between irrelevant feature explanations and false explanations, $t(19) = 1.01$, $p = .326$, $d = .30$.[1] See Figure 1.[2]

Experiment 2

Although the findings of Experiment 1 clearly demonstrate that children as young as 4 are able to spontaneously detect false explanations, it is less clear whether young children are also able to spontaneously detect when true statements are irrelevant for a particular explanatory goal. In Experiment 1, it was not until age 7 that children were able to indicate that irrelevant feature explanations were less helpful than relevant causal explanations.

[1]For converging results see Allen (2008). We thank Tania Lombrozo for pointing out the relevance of this dissertation to the current work.

[2]We find the same pattern of results when we analyze the data according to participants' first binary choice of "helpful" versus "not helpful" for each item, rather than using the full 4-point helpfulness scale.

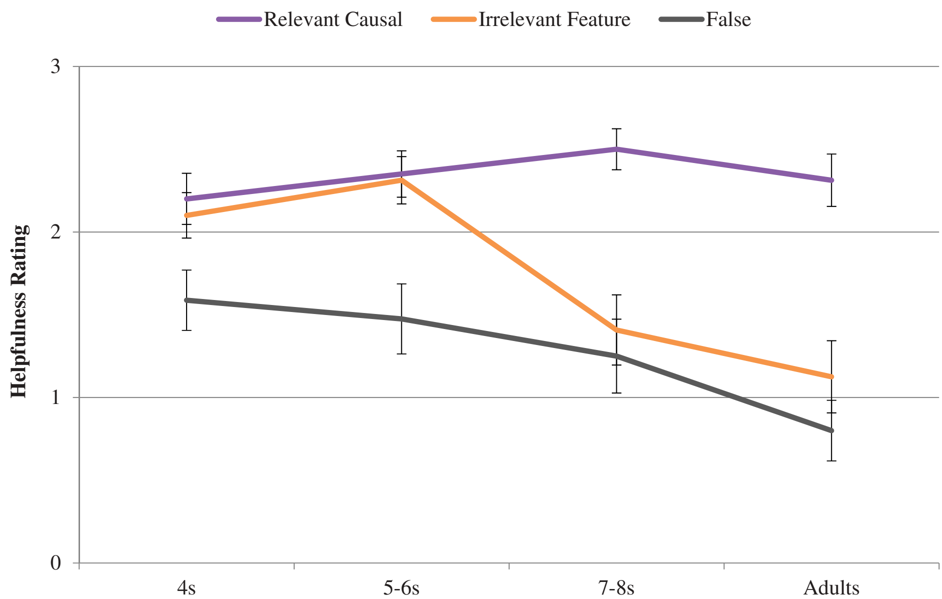


Figure 1. Mean values representing the helpfulness ratings given in Experiment 1, broken down by age group and explanation type. Error bars represent standard errors.

These findings may indicate that children younger than age 7 generally struggle to detect irrelevant explanations when they contain true information.

However, another possibility is that young children's failure to detect irrelevant feature explanations partially stemmed from the salient contrast between true and false explanations in Experiment 1. Specifically, the mere presence of false items in the Experiment 1 rating task might have caused young children to focus more on whether explanations were true or false than whether they were relevant or irrelevant. If so, then young children may have some ability to detect irrelevant explanations, but this ability is fragile and can be overshadowed by the salience of false explanations. To investigate this possibility in Experiment 2, we replaced the false items with items that were irrelevant because they referred to extrinsic factors unrelated to what makes cars go. Specifically, we included irrelevant explanations that referred to extrinsic factors (i.e., locations associated with cars), rather than typical properties of cars (e.g., radios, windshield wipers, etc.). If young children generally struggle to detect irrelevant explanations that contain true information, then they should struggle to detect not only irrelevant feature explanations in Experiment 2, but also irrelevant extrinsic explanations. In contrast, if young children simply have a more fragile ability to detect irrelevant explanations, then they may begin to detect one or both types of irrelevant explanations in Experiment 2.

Method

Participants

Participants were 20 4-year-olds (M_{age} = 4;4; SD_{age} = 3 months; 7 females), 20 5- to 6-year-olds (M_{age} = 5;9; SD_{age} = 6 months; 13 females), 20 7- to 8-year-olds (M_{age} = 7;9; SD_{age} = 6 months; 10 females), and 19 adults (M_{age} = 30; SD_{age} = 5.51; 7 females; recruited from Amazon Mechanical Turk).

Materials and procedure

The explanations used in Experiment 2 were identical to those used in Experiment 1 with one exception: we replaced the false explanations with a new type of true but irrelevant explanation – irrelevant extrinsic explanations (e.g., "Cars have shops that repaint over scratches;" see Appendix A for a complete list of explanations). Crucially, although these new irrelevant extrinsic items referred to locations pertaining to cars (i.e., car shops, carwashes, parking lots, and car dealerships), they all used the same format as explanations in Experiment 1 (i.e., "Cars have ______s that ______") in order to prevent children from rating them differently based on the underlying sentence structure. Other than the change in these explanations, the procedure of Experiment 2 was identical to that of Experiment 1.

Results and discussion

Our findings in Experiment 2 revealed that participants of all ages were able to detect some types of irrelevant explanations (i.e., our new irrelevant extrinsic explanations) when all explanations contained true information. These findings contrast with those of Experiment 1, in which children were unable to detect irrelevant explanations until age 7 when asked to consider false explanations alongside the relevant and irrelevant explanations. However, we still saw a shift in Experiment 2 – this time between age 4 and 5 – in children's ability to spontaneously detect our irrelevant feature explanations. Before age 5, children indicated that irrelevant feature explanations were just as helpful as relevant causal explanations, and after age 5, children indicated that irrelevant feature explanations were less helpful than relevant causal explanations.

Specifically, an omnibus ANOVA showed an interaction between age (4-year-olds, 5- to 6-year-olds, 7- to 8-year-olds, and adults) and explanation type (relevant causal, irrelevant feature, and irrelevant extrinsic) for participants' explanation ratings, $F(6, 150) = 3.56$, $p = .003$, $\eta^2 = .13$. We followed up on this interaction with paired-samples *t*-tests comparing each age group's ratings of the two types of irrelevant explanations (i.e., irrelevant feature and irrelevant extrinsic) to their ratings of the relevant causal explanations. Children in all age groups rated irrelevant extrinsic explanations as being *less* helpful than relevant causal explanations ($ts > 2.55$, $ps < .020$, $ds > .59$), suggesting that children as young as 4 are at least capable of spontaneously detecting irrelevant explanations when they refer to extrinsic features. However, even in the absence of false explanations, it was not until age 5 that children began to distinguish irrelevant feature explanations from the relevant causal ones. Specifically, 4-year-olds did not distinguish between relevant causal and irrelevant feature explanations in their helpfulness ratings, $t(19) = .69$, $p = .497$, $d = .16$, though all other age groups indicated that irrelevant feature explanations were less helpful than relevant causal explanations ($ts > 2.83$, $ps < .011$, $ds > .59$). See Figure 2.[3]

Experiment 3

Experiment 2 demonstrates that even 4-year-olds are somewhat sensitive to relevance when evaluating explanations. Even when all explanations contained true information, 4-

[3]We find the same pattern of results when we analyze the data according to the first binary choice of "helpful" versus "not helpful."

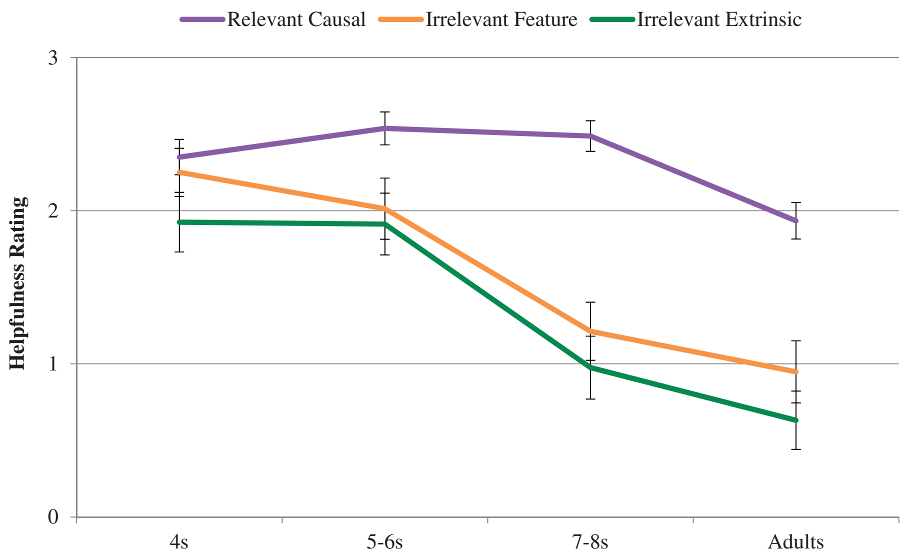


Figure 2. Mean values representing the helpfulness ratings given in Experiment 2, broken down by age group and explanation type. Error bars represent standard errors.

to 8-year-olds rated our irrelevant extrinsic explanations (e.g., pertaining to locations that are simply associated with cars, such as parking lots) as being less helpful than our relevant causal explanations (e.g., pertaining to causal features related to "what makes cars go", such as engines). However, unlike the 5-year-olds, the 4-year-olds in our study still did not detect the irrelevancy of our irrelevant feature explanations (e.g., car radios).[4]

Crucially, though, based solely on the results of Experiments 1 and 2, it is unclear why 4-year-olds struggled to detect irrelevant feature explanations. Perhaps 4-year-olds lacked the background knowledge necessary to evaluate which explanations are relevant to a particular explanatory goal, either for our specific items or more generally. For instance, 4-year-olds might not realize that car radios and windshield wipers are unrelated to "what makes cars go." If this were the case, then 4-year-olds' inability to spontaneously detect irrelevant feature explanations might simply be an artifact of our stimulus set.

To test this possibility, in Experiment 3 we presented children and adults with a forced choice between irrelevant feature explanations and relevant causal explanations. If young children simply lack the background knowledge necessary to evaluate which explanations are relevant, then they should show no preference between relevant causal and irrelevant feature explanations. In contrast, if young children are unable to bring their knowledge of relevance to bear when evaluating irrelevant feature explanations one at a time, then they

[4]As pointed out by an anonymous reviewer, and as highlighted in Experiment 4, some extrinsic explanations are relevant. For example, "cars have gas stations that give them fuel" would be relevant for the explanatory goal of understanding "what makes cars go." Thus, the pattern of results in Experiment 2 should not be interpreted as indicating a generalizable difference in the helpfulness of "feature" vs. "extrinsic" explanations, but rather should be interpreted as indicating that 4-year-olds have a more fragile ability than 5-year-olds to detect irrelevant explanations, and that this difference is captured by the specific stimulus items we chose.

should be able to indicate that irrelevant feature explanations are less helpful than relevant causal explanations when presented with a direct contrast.

Method

Participants

Participants were 20 4-year-olds (M_{age} = 4;6; SD_{age} = 2 months; 8 females), 20 5- to 6-year-olds (M_{age} = 5;5; SD_{age} = 3 months; 9 females), 20 7- to 8-year-olds (M_{age} = 7;4; SD_{age} = 4 months; 10 females), and 21 adults (M_{age} = 30; SD_{age} = 10.38; 7 females; recruited from Amazon Mechanical Turk).

Materials and procedure

The procedure of Experiment 3 was similar to that of Experiments 1 and 2. However, in Experiment 3, children and adults were presented with a forced choice between two explanations and asked to indicate which explanation was most helpful for "understanding what makes cars go." One explanation in each explanation pair was a relevant causal explanation (e.g., "Cars have engines that turn gasoline into power") and the other was an irrelevant feature explanation (e.g., "Cars have radios that play music"), such that there were four total explanation pairs. The explanations used in Experiment 3 were identical to those used in Experiments 1 and 2 (see Appendix A). Two randomized orders of the explanations were developed using a list-randomizing program. Both the order of the explanations and the explanation pairs differed between the two versions to ensure that any significant results were not driven by any specific comparisons.

Results and discussion

When presented with a forced choice, children in all age groups were able to indicate that relevant causal explanations were more helpful than irrelevant feature explanations, though this tendency increased with age. Specifically, a one-way ANOVA examining the number of times (ranging from 0 to 4) children said the relevant causal explanation was more helpful showed a significant effect of age (i.e., 4-year-olds, 5- and 6-year-olds, 7- and 8-year-olds, and adults), $F(3, 80) = 6.43$, $p = .001$, $\eta^2 = .20$. Post-hoc Tukey tests revealed this effect of age was primarily driven by differences between 4-year-olds and the 6/7-year-olds, $p = .021$, and the 4-year-olds and adults, $p < .001$. See Figure 3. Follow-up single-sample t-tests comparing each age group to chance revealed that children in all age groups preferred relevant causal explanations significantly more often than chance (all ts > 3.34; ps < .004; ds > .77). Thus, even 4-year-olds have the background knowledge necessary to indicate that relevant causal explanations are more helpful than irrelevant feature explanations.

Experiment 4

Experiments 1–3 demonstrate that young children differentiate between relevant and irrelevant explanations, even when all explanations describe true features of the world. However, these studies always used the same four causal explanations as the relevant

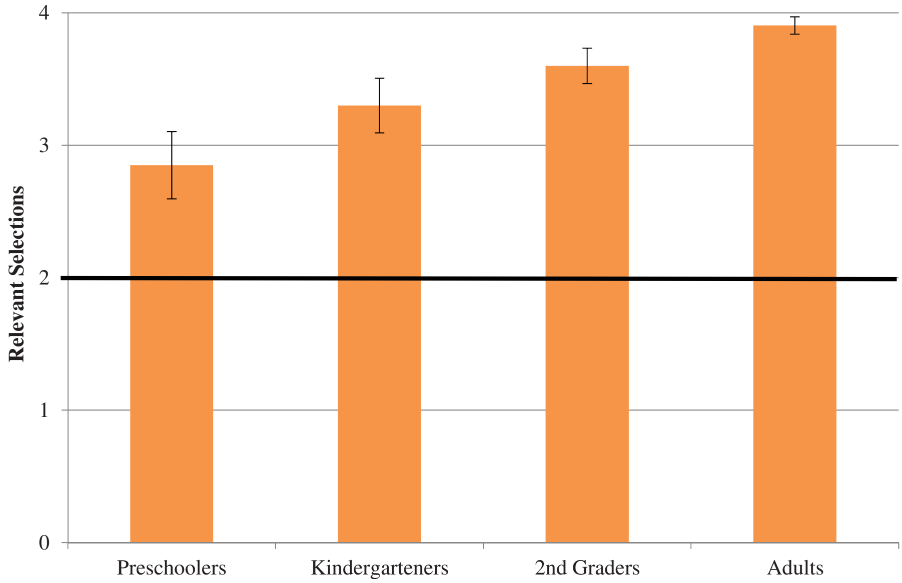


Figure 3. Mean values representing the number of times participants selected the relevant explanation as being more helpful than the irrelevant feature explanation in Experiment 3. Error bars represent standard errors.

explanations, leaving open the possibility that the results of these studies were driven by particular features of these explanations. For example, "Cars have engines that turn gasoline into power" might have been highly rated by children simply because they know the word "engine" is associated with cars but without having any insight into how an engine is relevant for a car or even what an engine is. Furthermore, specific features of the other explanations might have led them to be rated lower, such as the phrasing of the extrinsic items (e.g., "Cars have parking lots that they park in"), which were constructed to be parallel to the other explanations, but might come across as syntactically awkward.

To address these issues, in Experiment 4 we switched the explanatory goal from understanding "what makes cars go" to understanding "where to find cars." This allowed us to reverse which explanations were relevant, such that the explanations that were relevant in the previous experiments (e.g., "Cars have engines that turn gasoline into power") were irrelevant for the explanatory goal in Experiment 4, and explanations that were irrelevant in the previous experiments (e.g., "Cars have parking lots that they park in") were relevant for the explanatory goal in Experiment 4. If the new sample of children in Experiment 4 show a different pattern of results than did the children in the previous experiments, specifically by rating the extrinsic explanations higher than the other explanations, it would suggest that children are truly tracking which explanations are relevant to different explanatory goals, and not responding to any other feature (e.g., familiarity, awkwardness, etc.) of the stimuli.

Method

Participants

Participants were 20 4-year-olds (M_{age} = 4;7; SD_{age} = 4 months; 14 females), 20 5- to 6-year-olds (M_{age} = 5;10; SD_{age} = 6 months; 9 females), 20 7- to 8-year-olds (M_{age} = 7;10; SD_{age} = 8 months; 9 females), and 21 adults (M_{age} = 31; SD_{age} = 8.52; 11 females; recruited from Amazon Mechanical Turk). The child participants were run over a video chat platform developed for collecting developmental data, and that has been previously validated by replicating standard results in the literature (Sheskin & Keil, 2018). The images in Appendix A were presented via a shared PowerPoint presentation (rather than as printed out and laminated cards), but the experience was comparable (e.g., the images appeared on screen as the experimenter talked about them, rather than being placed on a table as the experimenter talked about them).

Materials and procedure

The same pictures and explanations were used as in Experiments 1 and 2, but whereas in Experiments 1 and 2 children were asked how helpful the explanations were for "understanding what makes cars go," in Experiment 4 children were asked how helpful the explanations were for "understanding where to find cars."

Results and discussion

Our findings in Experiment 4 revealed that, by age 5, children are sensitive to the explanatory goal at hand when evaluating explanations. With the explanatory goal "understanding where to find cars," children as young as 5 – and to some degree as young as 4 – indicated that an explanation such as "Cars have parking lots that they park in" was more helpful than an explanation such as "Cars have engines that turn gasoline into power." This is a reversal of the pattern shown by children in the previous experiments, who were asked about the explanatory goal "understanding what makes cars go."

Specifically, an omnibus ANOVA showed an interaction between age (4-year-olds, 5- to 6-year-olds, 7- to 8-year-olds, and adults) and explanation type (irrelevant causal, irrelevant feature, and relevant extrinsic) for participants' explanation ratings, $F(6, 152) = 2.78$, $p = .014$, $\eta^2 = .01$. We followed up on this interaction with paired-samples *t*-tests comparing each age group's ratings of the two types of irrelevant explanations (i.e., irrelevant causal and irrelevant feature) to their ratings of the relevant extrinsic explanations. Although 4-year-olds did not differentiate between the relevant extrinsic explanations and either of the irrelevant explanation types (all $ts < 1.71$; $ps > .104$; $ds < .27$), all of the other age groups rated the relevant extrinsic explanations as being more helpful than either the irrelevant feature (all $ts > 2.14$; $ps < .046$; $ds > .43$) or irrelevant causal explanations (all $ts > 2.85$; $ps < .011$; $ds > .56$) See Figure 4.

As in the prior experiments, we also examined children's responses to the first binary question for each item (i.e., "Is this helpful or not helpful for understanding where to find cars?"), separately from the 4-point scale that resulted from asking the binary follow-up questions (e.g., "Is it really helpful or a little helpful?"). For this analysis, children could have selected each explanation type as being "helpful" (as opposed to "not helpful") a maximum of 4 times for each explanation type. Although this analysis revealed the

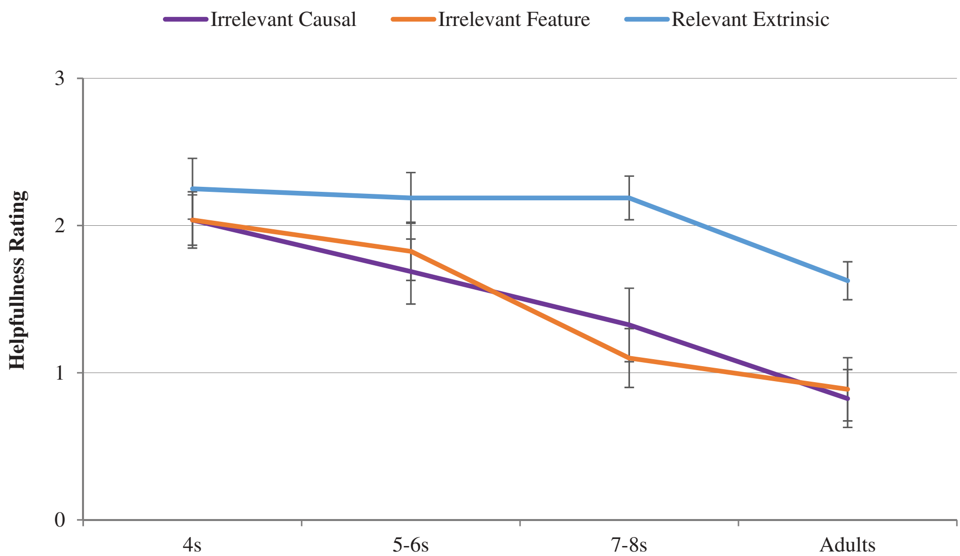


Figure 4. Mean values representing the helpfulness ratings given in Experiment 4, broken down by age group and explanation type. Error bars represent standard errors.

same pattern of results as the 4-point scale in Experiments 1 and 2, there was a difference in the results for 4-year-olds in Experiment 4. Specifically, when looking only at the first binary question regarding whether the explanations were "helpful" or "not helpful" in Experiment 4, there was some evidence that 4-year-olds did see the relevant extrinsic explanations (M = 3.25 out of 4, SD = 1.16) as being more helpful than either the irrelevant feature explanations (M = 2.70, SD = 1.49), $t(19)$ = 2.15, p = .045, d = .42, or the irrelevant causal explanations (M = 2.80, SD = 1.40), $t(19)$ = 2.02, p = .058, d = .35. These findings suggest that 4-year-olds may have some sensitivity to explanatory relevance, even when all of the explanations presented are true. However, 4-year-olds' understanding of explanatory relevance does not appear to be as robust as it is for older children.

General discussion

Young children can use relevance to distinguish between statements that are equally true, but that vary in their relevance for an explanatory goal. Most notably, when given an explanatory goal (e.g., "what makes cars go") and provided with a choice between a relevant explanation for that explanatory goal (e.g., "Cars have engines that turn gasoline into power") and an irrelevant explanation for that explanatory goal (e.g., "Cars have radios that play music"), children as young as 4 years old select the relevant explanation as more helpful (Experiment 3).

Given that children often focus on truth at the expense of other important characteristics (e.g., accent: Corriveau et al., 2013; age: Jaswal & Neely, 2006; familiarity: Corriveau & Harris, 2009), it is impressive that children as young as 4 years old are able to look beyond surface-level truth and evaluate explanations based on relevance. Although children in prior work have been able to recognize the relevance of novel information (e.g., novel words: Henderson et al., 2013; novel hiding locations: Eskritt et al., 2008), children in the current experiments went beyond simply evaluating novel information, and instead were sensitive to the relevance each explanation had for a particular explanatory goal, even when both explanations contained correct information. These findings build on past work showing that children as young as 2 will spontaneously probe for more information when their *why* and *how* questions (e.g., "Why won't this open?") are met with non-explanatory answers (e.g., "It looks like it has some buttons on the outside;" Frazier et al., 2009, 2016).

Importantly, children's ability to detect irrelevant explanations was not equally robust across all ages. In Experiment 1, when children had to rate relevant and irrelevant explanations alongside false explanations, only children over the age of 7 rated the relevant explanations as being more helpful than the irrelevant explanations. When we removed the false items in Experiment 2 and replaced them with additional irrelevant (extrinsic) explanations, 5- and 6-year-olds began to spontaneously detect the irrelevant feature explanations, but 4-year-olds continued to rate irrelevant feature explanations as being equally helpful as relevant explanations. Yet, 4-year-olds did show some ability to spontaneously detect irrelevant explanations in Experiment 4 when the explanatory goal was focused on "where to find cars." However, this ability to spontaneously detect irrelevant explanations in Experiment 4 was tenuous, as 4-year-olds only differentiated between the relevant and irrelevant explanations when we looked at their initial responses to the first binary choice question ("Is this explanation helpful or not helpful for understanding where to find cars?"), not in their responses to the full 4-point scale. It was only in Experiment 3, in which children were provided with a forced choice between relevant and irrelevant explanations, that 4-year-olds reliably used relevance to evaluate explanation helpfulness.

Thus, the way in which children spontaneously evaluate irrelevant explanations changes across development. By age 4, children are already able to choose a relevant over an irrelevant explanation in a forced-choice (Experiment 3), and sometimes rate explanations differentially as well (e.g., 4-year-olds in Experiment 2 rated the extrinsic feature explanations lower than the relevant causal explanations). By age 5, children spontaneously detect irrelevant explanations more reliably, but performance is still fragile enough that it can be disrupted by (e.g.,) the presence of false explanations drawing attention away from the distinction between "true and relevant" and "true but irrelevant" (as in Experiment 1). Finally, by age 7, children have a robust ability to evaluate the relevance of true explanations for a given explanatory goal, even when the explanations are not placed in direct comparison with each other and false items are included in the task.

In many ways, the developmental pathway that characterizes children's increasing sensitivity to explanatory relevance mirrors the development of children's selective trust. Previous work investigating children's selective trust shows that young children sometimes demonstrate understanding of a particular characteristic (e.g., source reliability) when given a direct contrast between two sources, but fail to spontaneously use this characteristic when evaluating sources one-by-one (Krogh-Jespersen & Echols, 2012; Vanderbilt et al., 2014). Our findings extend this line of research by demonstrating that this discrepancy between the characteristics children *understand* and the characteristics they spontaneously *use* extends to children's evaluations of explanations, independent of source characteristics. Future work should further explore how children evaluate explanations in the absence of a direct contrast. Many of the largest and most informative developmental differences may arise when exploring how children spontaneously evaluate explanations (for a recent example, see Mills et al., 2017).

Along these lines, future work could also investigate children's understanding of explanatory relevance across a wider range of contexts. In particular, it will be informative to probe children's understanding of explanatory relevance in unfamiliar or novel contexts. In the current studies, we specifically chose a domain (i.e., cars) that children were familiar with so that we could investigate how children would evaluate explanations that they knew to be true based on their background knowledge. However, one could also explore how children evaluate explanatory relevance in the absence of any prior background knowledge. Children's preference for the relevant explanations in the current studies cannot simply be explained by a greater familiarity with the relevant items, as the relevant items varied across experiments based on the shifting explanatory goals (i.e., "understanding what makes cars go" in Experiments 1–3 and "understanding where to find cars" in Experiment 4). That said, older children may have had some previous exposure to explanations about "what makes cars go," and/or "where to find cars." Although we believe it is unlikely that our findings can be explained away by a familiarity effect (as that would require that children had heard explanations about *both* explanatory goals to such a degree that they could associate our explanations to the explanatory goal), future work might investigate this more directly by investigating children's understanding of explanatory relevance in novel contexts.

Finally, future research should further investigate which features of explanations children pay attention to when evaluating whether a claim is true, and whether it is relevant for a particular explanatory goal. For example, when faced with a potential explanation, do children separately evaluate the truthfulness of the explanation as an individual claim (e.g., whether cars have rockets that speed them up) and the relevance of the explanation to the explanatory goal (e.g., whether rockets speeding a car up is relevant to what makes it go), or do these evaluations influence each other? Our findings in Experiment 1 suggest that children focus first on an explanation's accuracy as a statement and then evaluate whether it is relevant to a particular explanatory goal. Specifically, children in all age groups rated the false explanations in Experiment 1 as less helpful than the true/relevant explanations, even though the false explanations were designed to be relevant to understanding what makes cars go, had they been true (e.g., "Cars have rockets that speed them up."). To further investigate how children evaluate the relevance of *false* statements, future work should compare explanations that are false and relevant (similar to our false explanations in Experiment 1) to explanations that are false and irrelevant (e.g., "Cars have ovens that heat food."). Although the current set of studies clearly demonstrates that children as young as 4 are able to distinguish between relevant and irrelevant explanations in at least some contexts, future work investigating how they evaluate the relevance of false statements would provide more insight into the way in which children are integrating statement accuracy with explanatory relevance.

In sum, as young as 4 years old, children are already beginning to understand the importance of explanatory relevance. When presented with a direct contrast, 4-year-olds reject irrelevant explanations in favor of relevant explanations and do so even when the irrelevant explanations contain true information. Thus, explanatory relevance is added to the growing arsenal of explanatory features that children use to evaluate explanations (e.g., simplicity: Bonawitz & Lombrozo, 2012; circularity: Corriveau & Kurkul, 2014; Mills et al., 2017; scope: Johnston et al., 2017, 2018). However, this arsenal may remain somewhat fragile until later childhood, as it is not until age 5 or even later that children begin to spontaneously detect irrelevant explanations outside the context of a forced choice. Thus, when evaluating explanations in real time – as is often the case in real life – children have to learn the relevance of relevance.

Acknowledgments

We would like to thank the members of the Cognition and Development Lab, particularly Mariel Goddu, for helpful feedback in designing the studies. Additionally, we would like to thank Candice Mills and Amanda Royka for helpful comments on an earlier version of the manuscript. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1122492 and by NSF Grant DRL 1561143 to F. Keil.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Science Foundation [DGE-1122492,DRL 1561143].

References

Allen, E. A. (2008). *Constructing and judging answers to why-questions: What can children tell us?* (Doctoral dissertation). Retrieved from ProQuest Dissertations Publishing. (3320766)

Baum, L. A., Danovitch, J. H., & Keil, F. C. (2008). Children's sensitivity to circular explanations. *Journal of Experimental Child Psychology*, *100*(2), 146–155. doi:10.1016/j.jecp.2007.10.007

Bonawitz, E. B., & Lombrozo, T. (2012). Occam's rattle: Children's use of simplicity and probability to constrain inference. *Developmental Psychology*, *48*(4), 1156–1164. doi:10.1037/a0026471

Corriveau, K. H., & Harris, P. L. (2009). Choosing your informant: Weighing familiarity and recent accuracy. *Developmental Science*, *12*(3), 426–437. doi:10.1111/j.1467-7687.2008.00792.x

Corriveau, K. H., Kinzler, K. D., & Harris, P. (2013). Accuracy trumps accent in children's endorsement of object labels. *Developmental Psychology*, *49*(3), 470–479. doi:10.1037/a0030604

Corriveau, K. H., & Kurkul, K. E. (2014). "Why does rain fall?": Children prefer to learn from an informant who uses noncircular explanations. *Child Development*, *85*(5), 1827–1835. doi:10.1111/cdev.12240

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1149–1157. doi:10.1098/rstb.2010.0319

Eskritt, M., Whalen, J., & Lee, K. (2008). Preschoolers can recognize violations of the Gricean maxims. *British Journal of Developmental Psychology*, *26*(3), 435–443. doi:10.1348/026151007X253260

Frazier, B. N., Gelman, S. A., & Wellman, H. M. (2009). Preschoolers search for explanatory information within adult–Child conversation. *Child Development*, *80*(6), 1592–1611. doi:10.1111/j.1467-8624.2009.01370.x

Frazier, B. N., Gelman, S. A., & Wellman, H. M. (2016). Young children prefer and remember satisfying explanations. *Journal of Cognition and Development*, *17*, 718–736. doi:10.1080/15248372.2015.1098649

Gelman, S. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology*, *60*(1), 115–140. doi:10.1146/annurev.psych.59.103006.093659

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Speech acts Vol. 3* (pp. 41–58). New York, NY: Academic Press.

Harter, S. (1982). The perceived competence scale for children. *Child Development*, *53*, 87–97. doi:10.2307/1129640

Henderson, A. M., Sabbagh, M. A., & Woodward, A. L. (2013). Preschoolers' selective learning is guided by the principle of relevance. *Cognition*, *126*(2), 246–257. doi:10.1016/j.cognition.2012.10.006

Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science*, *17*(9), 757–768. doi:10.1111/j.1467-9280.2006.01778.x

Johnston, A. M., Johnson, S. G., Koven, M. L., & Keil, F. C. (2017). Little Bayesians or little Einsteins? Probability and explanatory virtue in children's inferences. *Developmental Science*, *20*(6), e12483. doi:10.1111/desc.12483

Johnston, A. M., Mills, C. M., & Landrum, A. R. (2015). How do children weigh competence and benevolence when deciding whom to trust? *Cognition*, *144*, 76–90. doi:10.1016/j.cognition.2015.07.015

Johnston, A. M., Sheskin, M., Johnson, S. G. B., & Keil, F. C. (2018). Preferences for explanation generality develop early in biology, but not physics. *Child Development*, *89*(4), 1110–1119. doi:10.1111/cdev.12804

Keil, F. C. (2011). Science starts early. *Science*, *331*(6020), 1022–1023. doi:10.1126/science.1195221

Koenig, M. A., & Echols, C. H. (2003). Infants' understanding of false labeling events: The referential roles of words and the speakers that use them. *Cognition*, *87*(3), 179–208. doi:10.1016/S0010-0277(03)00002-7

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, *76*(6), 1261–1277. doi:10.1111/j.1467-8624.2005.00849.x

Koenig, M. A., & Woodward, A. L. (2010). Sensitivity of 24-month-olds to the prior inaccuracy of the source: Possible mechanisms. *Developmental Psychology*, *46*(4), 815–826. doi:10.1037/a0019664

Krogh-Jespersen, S., & Echols, C. H. (2012). The influence of speaker reliability on first versus second labeling. *Child Development*, *83*(2), 581–590. doi:10.1111/j.1467-8624.2011.01713.x

Lockhart, K. L., Chuey, A., Kerr, S., & Keil, F. C. (2019). The privileged status of knowing mechanistic information: An early epistemic bias. *Child Development*. doi:10.1111/cdev.13246

Lombrozo, T. (2016). Explanatory preferences shape learning and inference. *Trends in Cognitive Sciences*, *20*(10), 748–759. doi:10.1016/j.tics.2016.08.001

Mills, C. M. (2013). Knowing when to doubt: Developing a critical stance when learning information from others. *Developmental Psychology*, *49*(3), 404–418. doi:10.1037/a0029500

Mills, C. M., Danovitch, J. H., Rowles, S. P., & Campbell, I. L. (2017). Children's success at detecting circular explanations and their interest in future learning. *Psychonomic Bulletin and Review*, *24*, 1465–1477. doi:10.3758/s13423-016-1195-2

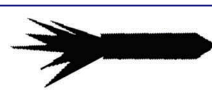
Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, *43*(4), 1045–1050. doi:10.1037/0012-1649.43.4.1045

Sheskin, M., & Keil, F. (2018). *TheChildLab.com A video chat platform for developmental research*. Retrieved from psyarxiv.com/rn7w5

Sperber, D., & Wilson, D. (1996). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.

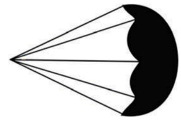
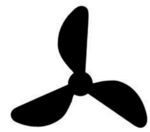
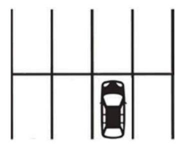
Vanderbilt, K. E., Heyman, G. D., & Liu, D. (2014). In the absence of conflicting testimony children trust inaccurate informants. *Developmental Science*, *17*(3), 443–451. doi:10.1111/desc.12134

Appendix A.

	Causal Explanations (Experiments 1–4)
Cars have engines that turn gasoline into power.	
Cars have batteries that start the motor.	
Cars have steering wheels that turn the tires.	
Cars have brakes that stop the wheels.	
	Feature Explanations (Experiments 1–4)
Cars have radios that play music.	
Cars have headlights that light up the road.	
Cars have trunks that hold extra bags.	
Cars have windshield wipers that clean the windows.	
	False Explanations (Experiment 1)
Cars have rockets that speed them up.	
Cars have skis that slide them side to side.	

(*Continued*)

(Continued).

	Causal Explanations (Experiments 1–4)
Cars have parachutes that slow them down.	
Cars have propellers that push them forward.	
	Extrinsic Explanations (Experiments 2 & 4)
Cars have carwashes that clean them.	
Cars have stores that sell them.	
Cars have parking lots that they park in.	
Cars have shops that repaint over scratches.