Cognition 143 (2015) 203–216

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Children adapt their questions to achieve efficient search

Azzurra Ruggeri [a,b,*], Tania Lombrozo [b]

[a] Max Planck Institute for Human Development, Berlin, Germany
[b] Department of Psychology, University of California, Berkeley, United States

ARTICLE INFO

Article history:
Received 1 July 2014
Revised 8 July 2015
Accepted 9 July 2015
Available online 17 July 2015

Keywords:
Information search
Causal inference
Hypothesis generation
Hypothesis testing

ABSTRACT

One way to learn about the world is by asking questions. We investigate how younger children (7- to 8-year-olds), older children (9- to 11-year-olds), and young adults (17- to 18-year-olds) ask questions to identify the cause of an event. We find a developmental shift in children's reliance on hypothesis-scanning questions (which test hypotheses directly) versus constraint-seeking questions (which reduce the space of hypotheses), but also that all age groups ask more constraint-seeking questions when hypothesis-scanning questions are least likely to pay off: When the solution is one among equally likely alternatives (Study 1) or when the problem is difficult (Studies 1 and 2). These findings are the first to demonstrate that even young children dynamically adapt their strategies for inquiry to increase the efficiency of information search.

1. Introduction

One way to learn about the world is by asking questions. We metaphorically "ask questions" when we perform experiments or make targeted observations, and we literally ask questions in the form of verbal inquiries to those around us. For children, who are often surrounded by more knowledgeable peers and adults, asking questions is especially important for testing and extending their developing understanding of the world (Gopnik & Meltzoff, 1997; Gopnik, Meltzoff, & Kuhl, 1999; Gopnik & Wellman, 1994; Harris, 2012; Piaget, 1954; see also Graesser & McMahen, 1993; Graesser & Olde, 2003). We know that young children ask domain-appropriate questions (Callanan & Oakes, 1992; Greif, Kemler Nelson, Keil, & Guiterrez, 2006; Hickling & Wellman, 2001), have reasonable expectations about which responses count as answers to their questions (Frazier, Gelman, & Wellman, 2009), and can use the answers they receive to solve problems (Chouinard, 2007; Legare, Mills, Souza, Plummer, & Yasskin, 2013). We also know that children's questions are responsive to the statistics of their environment in that they preferentially question reliable informants (Mills, Legare, Bills, & Mejias, 2010; Mills, Legare, Grant, & Landrum, 2011) and target informative cues (see Nelson, Divjak, Gudmundsdottir, Martignon, & Meder, 2014).

In the present paper, we go beyond previous research on children's strategies for inquiry by investigating whether and how children adapt their questions in response to evidence concerning the "information structure" of the task. Specifically, we investigate whether younger children (7- to 8-year-olds), older children (9- to 11-year-olds), and young adults (17- to 18-year-olds) choose the kinds of questions that they ask in a way that is sensitive to their expected information gain—that is, to how efficiently those questions partition the space of candidate hypotheses. We also examine whether children can modify their strategies "on-line," as they receive feedback in the course of inquiry. Finally, we consider whether the flexibility and efficiency of information search changes in the course of development. Investigating these issues is crucial not only for understanding the specific process of searching for information by asking questions, but also as a way to understand how children seek information and learn from others more generally.

In order to address these issues, we consider information search in the context of a causal attribution task in which children ask questions to uncover the explanation for an anomalous event (e.g., a child arriving late to school). This task is appealing for two reasons: First, it allows us to present children with meaningful information search problems in which candidate solutions are or are not specified in advance, and second, anomalous events are a common real-world trigger for information search. In particular, we know that observations that conflict with prior expectations can trigger a search for explanations (Khemlani & Johnson-Laird, 2011; Legare, Gelman, & Wellman, 2010; Lombrozo, 2006, 2012). Children could spontaneously generate hypotheses (e.g., perhaps the school bus encountered traffic) that can then be "tested" with appropriate questions (e.g., "Was there a lot of traffic?").

* Corresponding author at: Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Lentzeallee 94, Berlin, Germany.
E-mail address: ruggeri@mpib-berlin.mpg.de (A. Ruggeri).

http://dx.doi.org/10.1016/j.cognition.2015.07.004
0010-0277/

1.1. Children's questions in a sequential binary search task

Beginning with Mosher and Hornsby (1966), researchers have studied how children ask questions using variants on the game of "20-questions," in which one player thinks of an object and the second player has to identify that object by asking only yes-or-no questions. In its experimental version, participants are typically presented with a fixed number of objects (e.g., animals) and their task is to identify the object the experimenter has selected by asking as few questions as possible (e.g., Denney & Denney, 1973; Herwig, 1982; Ruggeri & Feufel, 2015; Siegler, 1977). Mosher and Hornsby (1966) pioneered the use of the 20-questions task and developed a useful coding system for classifying children's questions as "hypothesis scanning" or "constraint seeking." Hypothesis-scanning questions are tentative solutions—hypotheses that are tested directly (e.g., "Is it the dog?"). Constraint-seeking questions aim to reduce the space of possible hypotheses by testing higher-order features shared by several different hypotheses (e.g., "Does it have four legs?"). With constraint-seeking questions, children have the potential to rule out multiple hypotheses with each question, typically increasing the efficiency of search (e.g., Ruggeri & Feufel, 2015).

Following Mosher and Hornsby (1966), many researchers have found a reliable developmental trajectory, with the proportion of constraint-seeking questions asked increasing in the course of development, usually accompanied by a decrease in the number of questions required to reach the solution. The observed transition is often explained as a shift away from a perceptual focus on individual stimuli and objects and toward the ability to recognize object-general features that can be used to cluster similar objects into categories (e.g., quadrupeds versus nonquadrupeds). Consistent with this idea, Ruggeri and Feufel (2015) found that describing objects at a basic level (e.g., "dog" as opposed to "Dalmatian") increased the proportion of constraint-seeking questions in children and young adults (see also Herwig, 1982), suggesting that the basic-level representations facilitated children's ability to identify object-general features on which to base their questions.

Mosher and Hornsby (1966) also adapted the 20-questions task to investigate children's inquiry in the context of an open-ended causal inference game. Specifically, children aged 6–11 years were prompted to identify the cause of an event by asking yes-or-no questions. For example, in one task the children were told that "A man is driving down the road in his car; the car goes off the road and hits a tree," and had to reach the solution that "it had been raining and the man's car skidded off the slippery road on a curve." Replicating the findings from 20-questions tasks involving a predefined set of objects, Mosher and Hornsby found that the younger children tended to ask hypothesis-scanning questions (e.g., "did an animal run across the road and the man tried to avoid it?"), whereas older children asked more constraint-seeking questions (e.g., "Does it have something to do with the weather?"). When asked to describe their strategies ("How did you think of the questions asked?"), some of the few younger children who provided an answer mentioned the idea of a "general" question or of narrowing in on the answers. However, almost all the older children were able to articulate a more systematic strategy, with half of them mentioning more or less explicitly the principle of asking broad questions. Mosher and Hornsby (1966) reported this example (p. 99):

> Well, to eliminate big things quickly—like was there anything wrong with the road—was there anything wrong with the weather—was there anything wrong with the car—was there anything wrong with the person—if there's something wrong with the person, you start from the bottom and go to the top. I group like all the things with weather, breaking, then I group them smaller and smaller till I get to the point.

While the 20-questions task may appear quite artificial, Mosher and Hornsby's pioneering work, as well as subsequent efforts (e.g., Herwig, 1982; Nelson et al., 2014; Ruggeri & Feufel, 2015; Ruggeri, Lombrozo, Griffiths, & Xu, in press; Siegler, 1977), demonstrate that it can provide a rich source of data that reflects developmental changes in children's strategies for inquiry. Moreover, the problem of sequential, binary information search is actually a very general one, encountered throughout the lifespan. For instance, a similar process can be used for medical diagnoses: In emergency medicine, resident physicians learn to check for the presence or absence of certain physiological changes to rule out lethal conditions that can be associated with a particular complaint (e.g., Green & Mehr, 1997; Hamilton, Sanders, Strange, & Trott, 2003). Additional real world decision-making, categorization, and causal inference tasks have been modeled with fast and frugal trees that involve sequential, binary branching (see Berretty, Todd, & Blythe, 1997; Berretty, Todd, & Martignon, 1999; Martignon, Katsikopoulos, & Woike, 2008). Thus, studying children's performance on a 20-questions task is a good compromise between experimental tractability and real-world generalizability, and it is the task we adopt in our two studies.

1.2. Assessing the quality of children's questions

How *should* children ask questions? If we assume that the goal of asking questions is to obtain some piece of information in the most efficient way possible (i.e., with the fewest questions), then the best questions are those that are likely to yield the most informative answers—more formally, those with the greatest expected information gain (e.g., Oaksford & Chater, 1994). To make such a computation more tractable, we can imagine a situation like that in the 20-questions task, in which a child can only ask binary questions to arrive at one of a predefined set of possible solutions. In this case, the most informative question will be the one partitioning the space of solutions most evenly, with an answer of "yes" picking out a set of options that is as likely as that picked out by an answer of "no." This example sketches the problem of question asking at a computational level in Marr's sense (Marr, 1982)—i.e., by characterizing the problem an agent is facing in relation to her goals and the structure of the environment (see also Anderson, 1990; Chater & Oaksford, 1999). It also provides a benchmark against which to assess children's performance.

Within this framework, children's questions can be analyzed quantitatively in terms of expected information gain. For example, Nelson et al. (2014) presented 8- to 10-year-old children with variants on the 20-questions task. Children had to identify a person or a number from a set of equally likely alternatives by selecting questions from a list of options. They found that children were reasonably good at selecting questions that evenly partitioned the search space, especially when the statistical structure of the task matched the statistical structure of their real-world experience. For example, in trying to identify a target person from a set that was evenly divided between men and women, children were likely to first ask whether the person was male or female.

This quantitative approach to children's questions differs from that of Mosher and Hornsby (1966), which instead characterizes inquiry in terms of strategies that differ *qualitatively*. The two approaches, however, are not as different as they may appear: Constraint-seeking questions are treated as superior to hypothesis-scanning questions precisely because they usually yield greater information gain (Ruggeri & Feufel, 2015). Hypothesis-scanning questions are effectively a degenerate case of constraint-seeking, one in which the space of possible solutions is partitioned into two sets: One containing a single hypothesis and the other containing *everything else*. While constraint-seeking questions generally dominate hypothesis-scanning questions in terms of their expected information gain, the relative advantage of a constraint-seeking approach is not fixed, and in some cases a hypothesis-scanning question can even dominate constraint-seeking alternatives. For example, with only two equally likely candidate hypotheses, constraint-seeking questions will be no more informative than hypothesis-scanning questions. Moreover, when members within the set of candidate solutions are not all equally likely, a hypothesis-scanning question that targets a single very likely hypothesis (e.g., one that has a 50% probability of being true) can be more informative than a constraint-seeking question that differentiates an even *number* of hypotheses, but where the summed probability of those in one partition is small.

These observations are consistent with a much more general point, namely that a given strategy for inquiry cannot be defined as optimal *tout court*; Instead, its efficiency depends on what we will call the "information structure" of the task (Ruggeri, 2012; see also Todd, Gigerenzer, & the ABC Research Group, 2012). Constraint-seeking questions may yield more efficient information search in some cases, but not in all. Moreover, constraint-seeking questions may make greater cognitive demands: To quote Mosher and Hornsby, they offer efficiency "at the expense of cognitive work, the work involved in forming a plan for the strategy and in building the conceptual structure required" (Mosher & Hornsby, 1966, p. 88). While this cognitive work may pay off in a 20-questions task in which prespecified alternatives are all equally likely, it may not for many real-world cases of causal inference and attribution, for which candidate hypotheses are rarely prespecified *or* equally likely.

1.3. Key questions and hypotheses

In the present paper we aim to unite aspects of both qualitative and quantitative approaches to children's strategies for inquiry by considering (i) whether the *types* of questions that children ask (i.e., hypothesis-scanning versus constraint-seeking) change in response to the information structure of the task, (ii) whether question-asking changes dynamically as information is acquired, and (iii) whether sensitivity to information structure changes over the course of development. More concretely, (i) does the proportion of constraint-seeking questions asked increase when their expected information gain is greater relative to that of hypothesis-scanning questions? (ii) Is expected information gain re-assessed in the course of inquiry? And (iii) does the ability to adapt one's questions in response to information structure emerge over the course of development, or is it present even among younger children, who overwhelmingly ask hypothesis-scanning questions (see Mosher & Hornsby, 1966; Ruggeri & Feufel, 2015)? Addressing these questions represents an important step in bridging computational-level analyses of optimal information search with process-level accounts of how children actually go about learning from others.

Across two studies, younger children (7–8), older children (9–11) and young adults (17–18) were presented with an event (e.g., a man arriving late to work) and had to find out why this event occurred by asking as few yes-or-no questions as possible. These age ranges were motivated by prior research suggesting a strong developmental shift in children's strategies for inquiry between the ages of 7 and 10 (see Mosher & Hornsby, 1966; Ruggeri & Feufel, 2015). Critically, we varied the task's information structure by manipulating participants' knowledge of the relative probabilities of candidate hypotheses. When candidate hypotheses are roughly comparable in probability, constraint-seeking questions will be at least as efficient as hypothesis-scanning questions, and typically much more so. But when one candidate hypothesis is very likely, then the relative advantage for constraint-seeking questions is reduced: A hypothesis-scanning question targeting the likely hypothesis could have relatively high information gain, and also offers the potential for a "quick win" by solving the task with a single question.

In Study 1, we presented participants with a scenario (i.e., a man arriving late to work) and a set of ten candidate solutions that varied in relative probability. In one condition, hypotheses were equally likely, such that asking a hypothesis-scanning question would require arbitrarily selecting a hypothesis to test, and would always be dominated by asking a constraint-seeking question. In another condition, some hypotheses were much more likely than others, such that hypothesis-scanning questions could target the most likely options, and come closer to rivaling the efficiency of constraint-seeking questions while also offering the chance of a quick win. We therefore predicted that children would be more likely to ask constraint-seeking questions in the former condition, a result that would demonstrate—for the first time—that children adapt the kinds of questions that they ask in response to the information structure of the task.

In Study 2, we presented participants with a less constrained version of the task, similar to Mosher and Hornsby's (1966) open-ended causal inference game, in which we did not provide a set of candidate solutions. This open-ended task allowed us to investigate whether and how children and young adults modify their strategies for inquiry dynamically, as they receive feedback and learn more about the information structure of the task. We also attempted to manipulate participants' knowledge of the relative probabilities of hypotheses by varying (across participants) whether the scenarios were familiar (e.g., a child arriving late to school) or unfamiliar (e.g., an alien arriving late to a reunion). We reasoned that familiarity would translate into richer prior knowledge about the possible causes of an event and their relative probabilities. For example, one might know that it is much more likely that a child would be late for school due to traffic than because his house was flooded during the night, but be less certain about the corresponding probabilities for an alien. This would make a hypothesis-scanning question about traffic more attractive for the scenario involving the child than for that involving the alien.

In both Study 1 and Study 2, we additionally varied the prior probability of the solution. Solutions were selected such that they would be perceived as having a high prior probability (e.g., the man was late due to traffic) or a low prior probability (e.g., the man was late because he was bitten by a dog in his front yard). This manipulation is important because the efficacy of a given strategy for inquiry in a given case depends not only on whether the strategy maximizes expected information gain, but also on what happens to be true. An optimal strategy will require more questions, for example, when the solution happens to be a hypothesis that was a priori unlikely.

Finally, we considered two developmental predictions. First, in keeping with previous results (e.g., Mosher & Hornsby, 1966), we predicted a linear developmental improvement in participants' performance, with an increase in the proportion of constraint-seeking questions asked and a corresponding decrease in the number of questions required to reach the solution. Second, in line with a related literature suggesting that children's information search in other forms (e.g., revealing values on an information board rather than asking questions) is sensitive to the cost–benefit structure of a decision task (e.g., Davidson, 1991a, 1991b; Gregan-Paxton & Roedder John, 1995), we predicted that even younger children would be sensitive to our manipulation of the information structure of the task, generating a lower proportion of hypothesis-scanning questions when they were least likely to pay off. Such a result would be the first to show that the types of questions children ask depend on the information structure of the task, and to bridge qualitative and quantitative approaches to children's questions.

2. Study 1

In Study 1, children and young adults were presented with an event (a man arriving late to work) and asked to discover why it occurred by asking as few yes-or-no questions as possible. Participants were presented with 10 candidate hypotheses along with their frequencies (i.e., how many times, out of 40, that hypothesis explained a late arrival). For some participants, the hypotheses were all equally frequent. For others, the hypotheses varied in frequency, with some hypotheses being much more frequent than others. We tested the prediction that children would be more likely to ask constraint-seeking questions when all candidate hypotheses were equally frequent compared to conditions in which the candidate hypotheses differed in frequency.

To better understand the basis for this prediction, consider the expected information gain of questions across conditions. With 10 equally likely hypotheses, even the poorest constraint-seeking question (i.e., one that splits the initial set of 10 alternatives into 2 versus 8) would have an expected information gain of .72, whereas a hypothesis-scanning question would have an expected information gain of .15. Moreover, the probability of achieving a quick win (i.e., of guessing the correct hypothesis with an initial hypothesis-scanning question) would be only 1 out of 10. In contrast, with hypotheses that vary in frequency, an initial hypothesis-scanning question that targets a high-frequency hypothesis can rival or even dominate many constraint-seeking questions, and can also have a reasonable probability of yielding a quick win. With the frequencies used in the "mixed distribution" condition of Study 1, an initial hypothesis-scanning question could have an expected information gain of .81, and the chance of a quick win would be 1 out of 4. Although the best constraint-seeking questions would still yield the most efficient search across both of our frequency conditions (uniform or mixed), we anticipated that participants would be more tempted toward hypothesis-scanning questions when they had knowledge that allowed them to identify high-probability hypotheses.

We additionally manipulated whether the solution to the task had a high or low (perceived) prior probability. In so doing we could investigate the actual effectiveness of different question types in environments that were and were not representative of the provided information structure. Specifically, we tested the prediction that participants in the mixed distribution condition would reach the solution with fewer questions when the solution had a high perceived probability than when it had a low perceived probability, in part because hypothesis-scanning questions would pay off with relatively frequent "quick wins."

2.1. Method

2.1.1. Participants

Participants in Study 1 were 58 children in second or third grade (30 female, M_{age} = 7.0 years; SD = .59), 46 children in fifth grade (23 female, M_{age} = 10.2 years; SD = .73), and 70 young adults (39 female, M_{age} = 17.5 years; SD = .79) from two schools in Livorno, Italy. The students were all Italian and belonged to various social classes.

2.1.2. Design and procedure

The experiment consisted of individual interviews. At the beginning of each interview, the experimenter read the participant the task instructions, ensuring that they were completely understood. The instructions were modified from Mosher and Hornsby (1966) and presented in Italian. Below we provide an English translation:

> We're going to play some question-asking games. In these games I will tell you something that happened and your job will be to find out how it happened by asking me questions I can answer with "yes" or "no." If your question isn't clear or I do not know how to answer it, I will say "I can't answer," and then you will have to rephrase or explain your question or ask a different one. The goal of the game is to find the answer in as few questions as possible. However, you can ask as many questions as you need to find the answer.

After being read the instructions, the participants were presented with the following situation: "Yesterday, a man was late for work. Why? The solution is one of the following." The experimenter then took out 10 cards. On each card a different hypothesis was displayed, together with its frequency, expressed both as a label (i.e., "high," "moderate," or "low") and in natural frequencies (10, 4 or 2 out of 40 times). The experimenter read each card aloud, in random order, while putting them down on a table. The cards were left on the table until the end of the session, so that participants could read them again any time, and did not have to recall the presented information. Participants were told that the correct solution was among these 10 and told to begin asking yes-or-no questions to find out which one it was.

Participants were randomly assigned to one of four experimental conditions in a 2×2 design that crossed two independent variables: Hypothesis distribution (uniform, mixed) and solution probability (high probability, low probability). Table 1 presents a complete list of the hypotheses, displayed by condition.

2.1.2.1. Distribution. In the two *uniform distribution* conditions, the alternative hypotheses (i.e., possible solutions) provided to participants were designed to appear equally likely. In these conditions, the experimenter explicitly told participants that "all the alternatives are equally likely to be the correct solution."

In the two *mixed distribution* conditions, the hypotheses were designed such that two would be judged very likely to happen, four moderately likely to happen, and four very unlikely to happen. In these conditions, the experimenter presented the frequency of each hypothesis in a natural frequency format: "Out of 40 times a man is late, 10 times [very likely]/4 times [moderately likely]/1 time [very unlikely] it is because...."

To select hypotheses that would be perceived as equally likely (in the uniform conditions) or very likely/moderately likely/very unlikely (in the mixed conditions), we pretested 20 statements with an independent sample of 25 adults. Participants in the pretest were asked to rate the probability of the 20 described events on a 10-point scale, from 0 (extremely unlikely) to 10 (extremely likely). Using these data, we selected five statements that were judged very likely, five that were judged very unlikely, and two that were judged moderately likely. For each pretested statement, we constructed a "matched" item that was similar but distinct. For example, for the statement "He wasn't feeling well when he woke up," which was judged very unlikely, we constructed a second statement, "He hadn't felt well during the night." This allowed us to increase the total number of statements for the experiment and also ensured that pairs of statements involved common features that could provide a basis for asking constraint-seeking questions (e.g., "Was he feeling bad?").

Table 1
The candidate hypotheses (i.e., possible solutions) provided to participants, displayed by condition. The correct solution for each condition is indicated in bold.

Likelihood	Condition
Uniform distribution, high probability solution	
High	**He forgot something at home and had to return for it**
High	His alarm went off late so he overslept
High	His alarm did not go off so he overslept
High	He forgot something in his car and had to go get it
High	He could not find the keys to the car and had to search for them
High	He could not find his wallet and had to search for it
High	He had an errand to run on the way to work
High	He had a phone call to make before work
High	He wasn't feeling well when he woke up
High	He hadn't felt well during the night
Uniform distribution, low probability solution	
Low	**A dog bit him in his front yard**
Low	The clock was 1 h behind
Low	There were thieves in his house and he had to file a police report
Low	A pipe broke and he had to call the plumber
Low	He got robbed on the street on the way to the bus
Low	He got robbed on the bus
Low	He fell and broke his arm in the shower
Low	He fell on the stairs and broke his leg
Low	The clock at his house was set to the wrong time
Low	A dog on the street chased him
Mixed distribution, high probability solution	
High	**He forgot something at home and had to return for it**
High	He forgot something in his car and had to go get it
Moderate	He had to drop off his children at school that morning
Moderate	He had to drop his wife at work that morning
Moderate	The main street leading to his office was closed
Moderate	The police stopped his car for a random check
Low	He got robbed on the street on the way to the bus
Low	He got robbed on the bus
Low	A dog bit him in his front yard
Low	A dog on the street chased him
Mixed distribution, low probability solution	
High	He forgot something at home and had to return for it
High	He forgot something in his car and had to go get it
Moderate	He had to drop off his children at school that morning
Moderate	He had to drop his wife at work that morning
Moderate	The main street leading to his office was closed
Moderate	The police stopped his car for a random check
Low	He got robbed on the street on the way to the bus
Low	He got robbed on the bus
Low	**A dog bit him in his front yard**
Low	A dog on the street chased him

2.1.2.2. Solution probability. In the two *high probability* conditions, the correct solution had a high perceived probability. In the two *low probability* conditions, the correct solution had a low perceived probability. In order to match the solutions across the *uniform* and *mixed distribution* cases, we constructed four sets of 10 hypotheses as follows. For the *uniform/high probability* condition, all of the candidate solutions had a high perceived probability. For the *mixed/high probability* condition, the candidate solutions varied in perceived probability, and the solution matched that from the former set (i.e., *uniform/high probability*). For the *uniform/low probability* condition, all of the candidate solutions had a low perceived probability. For the *mixed/low probability* condition, the candidate solutions varied in perceived probability and were identical to those in the *mixed/high probability* condition, and the solution matched that from the former set (i.e., *uniform/low probability*).

2.1.2.3. Question coding. We coded questions as either hypothesis scanning or constraint seeking. When the question targeted only one of the candidate hypotheses (i.e., "Is it because a dog in the street chased him?"), the question was coded as hypothesis scanning. When the question involved a higher-order feature common to more than one hypothesis (i.e., "Does it have something to do with a dog?"), the question was coded as constraint seeking. Among the hypothesis-scanning questions, the correct solutions were coded as "hits."

We also introduced a further coding category within the set of hypothesis-scanning questions: *pseudoconstraint-seeking* questions (Mosher & Hornsby, 1966; Ruggeri & Feufel, 2015). Pseudoconstraint-seeking questions are, in terms of efficiency, identical to hypothesis-scanning questions: They are tentative solutions in that they pick out only one of the alternatives in the set. However, they differ from hypothesis-scanning questions because they take the form of constraint-seeking questions by referring to a higher-order feature of the target hypothesis. For example, instead of asking "Did he arrive late because he missed the bus?" (i.e., a hypothesis-scanning question that specifies the content of a particular hypothesis), a pseudoconstraint-seeking question would be "Did he arrive late because of something related to the bus?" when there is only one hypothesis left having something to do with a bus, namely, "He arrived late because he missed the bus."

All questions were coded from audio recordings by the experimenter, an Italian student assistant who did not know the experimental hypotheses, immediately after the session was over. All questions were additionally and independently coded by a second Italian student assistant, who did not know the experimental hypotheses, resulting in total agreement of Kappa = .953 with $p < .001$. In the few cases where the two raters did not agree, a third Italian rater, who did not know the experimental hypotheses or procedure, was consulted.

2.2. Results

The results were analyzed by comparing the three age groups and the four different conditions on two key outcomes: (1) the number of questions needed to reach the solution, and (2) the proportion of constraint-seeking and hypothesis-scanning questions asked. For the mixed distribution conditions, we additionally analyzed (3) whether the candidate solution selected for initial hypothesis-scanning questions was low, moderate, or high frequency. We did not analyze success rate, as all participants succeeded in completing the task and were able to reach the correct solution.

2.2.1. Number of questions

We analyzed the number of questions required to reach the solution as the dependent variable in a univariate ANOVA with age group (3: younger children, older children, young adults), distribution (2: uniform, mixed), and solution probability (2: high probability, low probability) as independent variables. This analysis found no main effects but revealed a significant interaction between distribution and solution probability, $F(2,173) = 12.58$, $p = .001$, $\eta^2 = .07$ (see Fig. 1). For participants in the mixed distribution conditions, those with the low probability solution needed more questions to reach the solution ($M_{\text{mixed\_low}} = 5.91$; $SD = 2.36$) than those with the high probability solution ($M_{\text{mixed\_high}} = 4.16$; $SD = 2.10$), $t(86) = 3.673$, $p < .001$. However, for participants in the uniform distribution conditions, those with the low probability solution ($M_{\text{uniform\_low}} = 4.23$; $SD = 2.03$) did not need more questions than those with the high probability solution ($M_{\text{uniform\_high}} = 5.05$; $SD = 2.85$), $t(84) = 1.541$, $p = .127$. Notably, we did not find interactions between age and other variables.

2.2.2. Types of questions

We performed a univariate ANOVA with the proportion of constraint-seeking questions asked by each participant as the dependent variable and age group (3: younger children, older

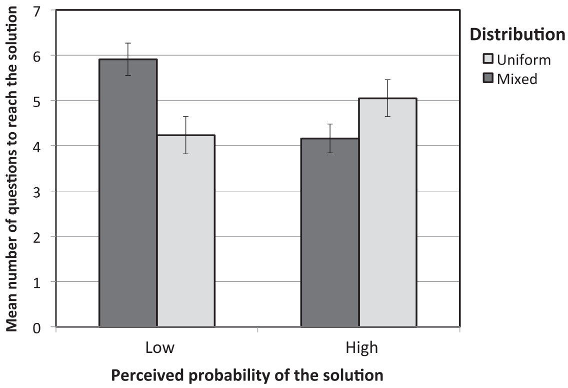

Fig. 1. Mean number of questions participants needed to reach the solution, displayed by distribution and solution probability. Error bars represent one SEM in each direction. These data collapse across age, as there were neither main effects nor interactions between these variables and age groups.

children, young adults), distribution (2: uniform, mixed), and solution probability (2: high probability, low probability) as independent variables. The analysis revealed a main effect of age group, $F(2,173) = 17.13$, $p < .001$, $\eta^2 = .18$. A Bonferroni post hoc analysis found that younger and older children asked a similar proportion of constraint-seeking questions ($M_{young\_child} = .23$; $SD = .37$; $M_{old\_child} = .20$; $SD = .31$, $p = 1.00$), which was lower than the proportion of constraint-seeking questions asked by young adults ($M_{young\_adults} = .51$; $SD = .31$, $p < .001$; see Table 2, top). The analysis also revealed a main effect of distribution, $F(1,173) = 9.29$, $p = .003$, $\eta^2 = .06$. As can be seen in Table 2 (top), participants assigned to the mixed distribution conditions asked a lower proportion of constraint-seeking questions ($M_{mixed} = .26$; $SD = .32$) than the participants assigned to the uniform distribution conditions ($M_{uniform} = .41$; $SD = .39$). Interestingly, this effect did not interact with age ($p = .977$). Even younger children asked a higher proportion of constraint-seeking questions when confronted with a uniform distribution ($M_{uniform} = .32$; $SD = .41$) as opposed to a mixed distribution ($M_{mixed} = .15$; $SD = .31$), $F(1,57) = 3.20$, $p = .079$, $\eta^2 = .05$. We did not find any effect of solution probability.

An analysis restricted to the *first* question asked by each participant similarly revealed a greater probability of asking a constraint-seeking question with age and in the uniform distribution condition compared to the mixed distribution condition, with no effect of solution probability (see Table 2, bottom).[1]

We additionally analyzed the proportion of hypothesis-scanning questions that were pseudoconstraint seeking by performing an equivalent ANOVA. The analysis revealed a main effect of age group, $F(2,151) = 27.08$, $p < .001$, $\eta^2 = .28$. A Bonferroni post hoc analysis found that young adults' hypothesis-scanning questions involved a higher proportion of pseudoconstraint-seeking questions ($M_{young\_adults} = .51$; $SD = .42$) than did those of older ($M_{old\_child} = .20$; $SD = .33$, $p < .001$) or younger children ($M_{young\_child} = .05$; $SD = .14$, $p < .001$).

[1] To analyze the frequency with which first questions were constraint seeking, we conducted a logistic regression with age group, distribution, and solution probability as predictors. A test of the full model against a constant-only model was statistically significant, indicating that the predictors, as a set, reliably distinguished between participants who did and did not initially generate a constraint-seeking question ($\chi^2 = 37.818$, $df = 3$, $p < .001$). Nagelkerke's $R^2 = .262$ indicated a moderately weak relationship between prediction and grouping. Prediction success overall was 71% (70% for constraint-seeking and 72% for non-constraint-seeking questions). The Wald criterion demonstrated that both age group ($p < .001$) and distribution ($p = .011$) made a significant contribution to prediction, whereas solution probability was not a significant predictor. The exp(B) value indicated that the likelihood of generating an initial constraint-seeking question increased with age (by 2.94 times), after the other factors in the model were controlled for. However, younger and older children asked a similar proportion of constraint-seeking questions ($M_{young\_child} = .28$; $SD = .45$; $M_{old\_child} = .29$; $SD = .46$, $p = 1.00$), which was lower than the proportion of constraint-seeking questions asked by young adults ($M_{young\_adults} = .73$; $SD = .45$, $p < .001$). The exp(B) value also indicated that participants assigned to the mixed distribution condition had a decreased probability of generating an initial constraint-seeking question (by .42 times). Participants in the mixed distribution conditions asked a lower proportion of constraint-seeking questions ($M_{mixed} = .38$; $SD = .49$) than the participants assigned to the uniform distribution conditions ($M_{uniform} = .55$; $SD = .50$).

2.2.3. Probability of hypotheses tested in initial questions

Among the hypothesis-scanning questions asked as initial questions in the mixed distribution conditions, few involved candidate hypotheses that were very unlikely (younger children: 26%; older children: 6%; young adults: 0%). Indeed, most were either very likely or moderately likely. However, this trend varied with age: Young adults overwhelmingly tested solutions with high probability first (88%), while younger and older children did so only 46% and 44% of the time, respectively, $\chi^2 = 8.43$, $df = 2$, $p = .015$.

2.3. Discussion of Study 1

As hypothesized, participants of all ages adapted the kinds of questions that they asked to the different probability distributions over hypotheses. When the candidate hypotheses were presented as equally likely, participants tended to ask more constraint-seeking questions. When the candidate hypotheses differed in probability, participants asked more hypothesis-scanning questions. While we found that the proportion of constraint-seeking questions asked increased with age, as predicted, there was no evidence that older participants were more sensitive to changes in the information structure of the task than were younger participants (i.e., there were no interactions between age group and other variables). Despite a general tendency to favor hypothesis-scanning questions, even young children shifted toward constraint-seeking questions when they were most likely to pay off in terms of the efficiency of their search. This is the first demonstration, to our knowledge, that children and adults effectively adapt the *kinds* of questions that they ask to achieve efficient search.

While we failed to find a developmental change in participants' sensitivity to our manipulation of the information structure of the task, we did find an important developmental shift in participants' ability to use distributional information wisely: When asking hypothesis-scanning questions in the condition with a mixed distribution, young adults were more likely to test the hypothesis with the highest probability first; they did so about twice as often as younger and older children. We speculate that children may have relied more heavily on their own experience instead of basing their strategies exclusively on the frequency information provided by the experimenter. Indeed, for 84% of children, the first hypothesis-scanning question asked targeted the hypothesis that the man had to drop off his children at school, which is presumably a familiar occurrence for children. However, it might also be that children's general understanding of frequencies and probability is not as fine-grained as that of young adults, or requires additional scaffolding. For example, it could be that children's ability to use the distributional information could be improved by displaying frequency information visually (e.g., with icon arrays, see Martignon & Krauss, 2009) or with fast and frugal trees (e.g., Martignon & Krauss, 2009; Martignon et al., 2008). It would therefore be valuable to replicate our study with paradigms that are not so dependent on text or numerical representations.

We also found, as hypothesized, an effect of solution probability. For participants in the mixed distribution conditions, those with the low probability solution needed more questions to reach the solution than those with the high probability solution. This effect, again, did not interact with age. Even though children were

Table 2
Percentage and standard deviation of constraint-seeking questions, hypothesis-scanning questions (that were not pseudoconstraint-seeking), and pseudoconstraint-seeking questions asked by participants in Study 1, overall and as the first question, by age group and distribution.

Age group	Distribution	Constraint seeking		Hypothesis scanning		Pseudoconstraint seeking	
		%	*SD*	%	*SD*	%	*SD*
Overall							
Younger children	Uniform	32	41	64	41	4	8
	Mixed	15	31	82	33	3	10
	Total	23	37	74	38	3	9
Older children	Uniform	27	37	65	41	8	16
	Mixed	12	19	70	38	17	26
	Total	20	31	68	39	12	22
Young adults	Uniform	59	32	19	25	22	28
	Mixed	44	30	30	34	26	26
	Total	51	31	25	30	24	27
Total	Uniform	41	39	47	42	12	21
	Mixed	26	32	58	42	16	24
	Total	34	36	52	42	14	23
First question							
Younger children	Uniform	36	41	61	41	4	8
	Mixed	20	31	80	33	0	10
	Total	28	37	71	38	2	9
Older children	Uniform	35	37	57	41	9	16
	Mixed	23	19	68	38	9	26
	Total	29	31	62	39	9	22
Young adults	Uniform	85	32	3	25	12	28
	Mixed	61	30	17	34	22	26
	Total	77	31	10	30	17	27
Total	Uniform	55	39	37	42	8	21
	Mixed	38	32	51	42	11	24
	Total	46	36	44	42	10	23

significantly less likely to initiate their search with the most likely hypotheses, they still tended to test unlikely hypotheses last.

Finally, it is interesting to note that the proportion of pseudoconstraint-seeking questions increased with age, despite the fact that such questions were no more informative than their corresponding hypothesis-scanning alternatives. This suggests that older participants understood the form that questions should typically take to achieve efficient search, even though they sometimes failed to implement the strategy in a way that actually improved efficiency given the actual information structure of the task.

3. Study 2

In Study 1, participants were explicitly given both the full set of candidate solutions and their relative frequencies. Under these conditions, we found that participants in all age groups adapted the kinds of questions that they asked in response to the information structure of the task, asking more constraint-seeking questions when all hypotheses were equally likely to be true. In Study 2, we considered a more realistic context for inquiry: One in which candidate solutions are not specified in advance, and in which relative frequencies are not provided. Under these conditions, we would still expect children's questions to reflect the information structure of the task, but this structure would need to be inferred on the basis of prior knowledge and updated dynamically in response to feedback in the course of asking questions.

In Study 2, we thus investigate how prior knowledge and on-line feedback affect the course of inquiry. To manipulate prior knowledge, we varied the content of the vignette: It involved either a man arriving late to work, a child arriving late to school, or an alien arriving late to a reunion. To investigate how question asking unfolds dynamically, we analyzed whether and how participants changed the kinds of questions asked in response to different kinds of feedback ("yes" or "no"). These manipulations allowed us to test two predictions.

First, we predicted that participants would be more likely to ask constraint-seeking questions when they had weaker prior knowledge, and therefore lacked a firm basis for assigning higher prior probabilities to some hypotheses over others—effectively approaching the "uniform distribution" condition from Study 1. We therefore expected a higher proportion of constraint-seeking questions for the unfamiliar alien vignette relative to the man and boy vignettes.

Second, we predicted that participants would be more likely to ask constraint-seeking questions after receiving "no" feedback than "yes" feedback. Hearing "no" should rule out at least one plausible hypothesis (if the preceding question was hypothesis-scanning), and perhaps an entire set of plausible hypotheses (if it was constraint-seeking). More generally, negative feedback should signal that one's prior beliefs (in the form of a probability distribution over hypotheses) are not a reliable guide to the task, either because those beliefs are inaccurate or because the current solution is unrepresentative. For similar reasons, we would expect a higher proportion of constraint-seeking questions when the solution has an a priori low probability as opposed to a high probability, as arriving at the solution would almost certainly require more questions, and more "no" feedback.

Finally, Study 2 allowed us to revisit our developmental predictions: That the proportion of constraint-seeking questions would increase over development, and that children—like adults—would succeed in adapting the kinds of questions that they asked in response to the information structure of the task.

3.1. Method

3.1.1. Participants

Participants were 96 children in second or third grade (66 female, M_{age} = 7.51 years; SD = .50), 87 children in fifth grade (56 female, M_{age} = 9.83 years; SD = .77), and 90 young adults (35 female, M_{age} = 17.62 years; SD = 1.07) from three schools in Livorno, Italy. The students were Italian and represented a variety of social classes.

3.1.2. Design and procedure

Like Study 1, Study 2 consisted of individual interviews. The instructions were identical to those in Study 1.

After being read the instructions, participants were presented with a short description of an event (e.g., "Yesterday, a boy was late for school") and asked why it occurred ("Why?"), after which they were expected to ask questions. Participants were randomly assigned to one of six experimental conditions, the result of crossing two independent between-subjects variables, scenario (3: man, boy, alien) and solution probability (2: high probability, low probability), which we explain below.

3.1.2.1. Scenario. There were three possible scenarios: (1) "Yesterday, a man was late for work"; (2) "Yesterday, a boy was late for school"; and (3) "Yesterday, an alien was late for the supreme reunion." These scenarios were designed to vary in familiarity, with Scenarios 1 and 2 more familiar than Scenario 3.

3.1.2.2. Solution probability. There were two possible solutions for each scenario: (a) a *high probability* solution, and (b) a *low probability* solution. Each solution was structured into three levels of causal detail. For the high probability solution, the car/spaceship taken by the man/boy/alien to go to work/school/the reunion was caught in a traffic jam (Level 1) due to a car accident (Level 2) that was caused by a driver who ran a red light (Level 3). For the low probability solution, the man/boy's father/alien had to wait for the plumber (Level 1), whom he had called because the house flooded during the night (Level 2) because a pipe had broken (Level 3). Participants were not explicitly told in the instructions that there were three levels of detail, but there was rarely a need to mention this in the course of the task. Instead, when a participant reached an initial level of detail (i.e., whichever of Level 1, Level 2, or Level 3 was reached first) by receiving a "yes" in response to a hypothesis-scanning question, the participant simply continued asking questions; the fact that the game was not over was implicitly communicated by the fact that they were expected to continue asking questions. Only a few participants, after receiving "yes" feedback to a hypothesis-scanning question, asked whether the game was over and were prompted to continue asking questions to reach a more specific solution. Including multiple levels of detail allowed us to analyze how inquiry unfolded in response to both "yes" and "no" feedback (more specifically: inquiry could continue even after receiving "yes" feedback to a hypothesis-scanning question).

The solutions were deemed *high probability* or *low probability* based on intuitions (subsequently confirmed by a post-test) about the perceived probability of the first level of detail—that is, that a traffic jam was a likely cause of tardiness across all scenarios while waiting for a plumber was a less likely cause of tardiness across all scenarios. So, for example, for the high probability solution, we assumed that P(traffic|late) would be high, and for the low probability solution we assumed that P(waiting for plumber|late) would be low. The subsequent levels of detail were selected to roughly match across conditions in having a moderate conditional probability. For example, P(car accident|traffic) and P(running red light|car accident) in the high probability solution, as well as P(flood|waiting for plumber) and P(pipe broken|flood) in the low probability solution, were anticipated to be seen as moderately likely. These assumptions were verified at the end of the experimental session, as described below.

3.1.2.3. Manipulation check. After participants reached the solution, they were asked to estimate prior and conditional probabilities for the scenario they had been presented with (i.e., man, boy, or alien) for both the high probability and low probability solutions. Probabilities were elicited in terms of frequencies. So, for example, participants in the man/likely solution and the man/unlikely solution conditions were asked, (Level 1) "Suppose you had 100 cases of a man being late for work. For how many of those 100 do you think being late would be the result of traffic on the way to work? If you say zero, it would mean none, and if you say 100 it would mean in all cases"; (Level 2) "Suppose you had 100 cases of traffic. Out of these, for how many do you think the traffic would be the result of a car accident? If you say zero, it would mean none, and if you say 100 it would be in all cases"; (Level 3) "Suppose you had 100 cases of a car accident. Out of those, for how many do you think the car accident would be the result of a car running a red light? If you say zero, it would mean none, and if you say 100 it would be in all cases." These questions allowed us to confirm our intuitions about the relative probabilities of our high probability and low probability solutions.

3.1.2.4. Question coding. As in Study 1, questions were coded as either hypothesis scanning or constraint seeking. If a response of "yes" to a question would have meant that the participant had reached at least one level of detail for the actual solution or an alternative of comparable specificity, the question was coded as hypothesis scanning (e.g., "Was the man/boy/alien late because he missed the bus?"). Otherwise, the question was coded as constraint seeking (e.g., "Was the man/boy/alien late because of something related to his means of transportation?"). Among the hypothesis-scanning questions, the correct solutions were coded as "hits," and the coding further specified which level of the solution was reached. Because participants were not given a complete set of candidate hypotheses, we could not introduce in the coding, as we did in Study 1, the designation of pseudoconstraint-seeking questions within the set of hypothesis-scanning questions.

The experimenter, an Italian student assistant who did not know the experimental hypotheses, wrote down all the questions asked during the experiment. In addition, the experimental session was audio recorded. Based on the experimenter's notes and on the recordings, the experimenter coded all questions immediately after the session was over. All questions were additionally and independently coded from the recording following the session by a second Italian student assistant who did not know the experimental hypotheses, resulting in total agreement of Kappa = .993, p < .001. In the few cases where the two raters did not agree, a third Italian rater who did not know the experimental hypotheses and procedure was consulted.

3.2. Results

The results were analyzed by comparing the three age groups on five different measures. Our first three measures track basic indices of performance and success: (1) success rate; (2) number of questions needed to reach the solution; and (3) overall proportion of constraint-seeking versus hypothesis-scanning questions. An additional analysis provides insight into the dynamics of the task: (4) question type asked in response to "yes" or "no" feedback from the previous question. Finally, we report (5) the probability estimates for the manipulation check. In the supplementary materials we additionally report the order in which the details of the solution were reached.

3.2.1. Success rate

Table 3 displays the percentage of participants who succeeded in finding the complete solution. In total, 105 participants did not finish the game: 70 younger children, 27 older children, and 8 young adults.[2]

To analyze success rates, we conducted a logistic regression analysis with age group, scenario, and solution probability as predictors. A test of the full model against a constant-only model was statistically significant, indicating that, as a set, the predictors reliably distinguished between participants who succeeded and those who did not succeed, $\chi^2 = 137.782$, $df = 3$, $p < .001$. Nagelkerke's $R^2 = .533$ indicated a moderate relationship between prediction and grouping. Prediction success overall was 79% (78% for non-success and 82% for success). The Wald criterion demonstrated that only age group ($p < .001$) and solution probability ($p < .001$) made a significant contribution to predicting success, whereas scenario was not a significant predictor. The exp(B) value indicated that, after we controlled for the other factors in the model, younger age groups (younger children compared to older children, and older children compared to young adults) had a *decreased* likelihood of succeeding (by .14 times), and being assigned to the condition with the high probability solution *increased* the likelihood of succeeding by 8.85 times.

In sum, we found a strong developmental change, with successful completion of the task increasing with age, as well as higher success rates when the solution had a higher perceived probability as opposed to a low perceived probability.

3.2.2. Number of questions needed to reach the solution

For this analysis we considered only those participants who reached the complete solution (all three levels). We performed a univariate analysis of variance (ANOVA) with the number of questions needed to reach the complete solution as the dependent variable and age group (3: younger children, older children, young adults), scenario (3: boy, man, alien), and solution probability (2: high probability, low probability) as independent variables. This analysis revealed a main effect of age group, $F(2,167) = 4.22$, $p = .016$, $\eta^2 = .05$. A Bonferroni post hoc analysis revealed that older children asked more questions prior to reaching the complete solution ($M_{\text{old\_child}} = 19.73$, $SD = 14.82$) than did either younger children ($M_{\text{young\_child}} = 11.58$, $SD = 9.11$, $p < .001$) or young adults ($M_{\text{young\_adults}} = 15.74$, $SD = 9.48$, $p = .029$). We found no difference between younger children and young adults ($p = .122$); this surprising result should be interpreted with caution, however, as few younger children completed the game, and this analysis includes only the subset who succeeded ($N = 26$).[3]

We also found a main effect of solution probability, $F(1,167) = 60.65$, $p < .001$, $\eta^2 = .29$: Participants assigned to the version of the game with the low probability solution needed more questions to reach the complete solution than those who were assigned the high probability solution (see Fig. 2, left). No additional main effects or interactions reached significance. Note that the pattern of results from this analysis does not change if we instead consider the number of questions needed to reach only one level of detail (i.e., whichever of Level 1, Level 2 or Level 3 was reached first; see Fig. 2, right).

In sum, we found that older children asked the most questions to reach the solution—likely reflecting a selection effect from younger children who failed to complete the task—as well as an effect of solution probability, with more questions needed when the solution had a low perceived probability.

3.2.3. Type of questions

For participants who reached the complete solution, we calculated the percentage of total questions that were constraint seeking. This percentage was analyzed as the dependent variable in a univariate ANOVA with age group (3: younger children, older children, young adults), scenario (3: boy, man, alien), and solution probability (2: high probability, low probability) as independent variables (see Fig. 3, left).

The analysis revealed a main effect of age, $F(2,167) = 17.08$, $p < .001$, $\eta^2 = .18$. A Bonferroni post hoc analysis confirmed that young adults asked a higher percentage of constraint-seeking questions ($M_{\text{young\_adult}} = 30\%$, $SD = 19\%$) than older children ($M_{\text{old\_child}} = 16\%$, $SD = 15\%$, $p < .001$), who in turn asked a higher percentage of constraint-seeking questions than younger children ($M_{\text{young\_child}} = 5\%$, $SD = 10\%$, $p = .009$).

We also found a main effect of solution probability, $F(1,167) = 19.86$, $p < .001$, $\eta^2 = .12$. Participants asked a higher percentage of constraint-seeking questions in the game with a low probability solution ($M_{\text{low}} = 33\%$, $SD = 19\%$), as compared to the game with a high probability solution ($M_{\text{high}} = 15\%$, $SD = 16\%$). Notably, we found no main effect of scenario on the type of question asked, nor an interaction between age and solution probability. A similar analysis restricted to the initial question—and considering all participants—also revealed an effect of age group, but not of scenario or solution probability.[4]

To investigate whether the proportion of constraint-seeking questions was predictive of the number of questions needed to reach the solution, we ran a linear regression with proportion of constraint seeking questions, solution probability, and the interaction between the two as predictors. The overall model fit was $R^2 = .41$. We found that both solution probability ($Beta = .845$, $p < .001$) and the interaction term ($Beta = -.528$, $p = .021$) were significant predictors, whereas the proportion of constraint-seeking questions alone was not a significant predictor ($p = .103$). To better understand the interaction, we examined Pearson correlations for each solution probability independently: Whereas the correlation between proportion of constraint-seeking questions and number of questions needed to reach the solution was significant in the low probability condition ($r = -.260$, $p = .049$), it was not significant in the high probability condition ($r = -.031$, $p = .751$). In other words, asking constraint-seeking questions improved efficiency when the solution was unlikely, but not when the solution was

[2] Among the participants who dropped out before reaching the complete solution, 81% did not reach any level of detail for the solution, 8% reached only one level of detail, and 11% reached two levels of detail. Participants who dropped out asked on average 16 questions ($SD = 10.80$) before dropping out, and they all indicated that they were tired of asking questions and wanted to give up.

[3] One concern is that, given the small number of younger children who were included in this analysis ($N = 26$), all analyses restricted to participants who succeeded in reaching the solution could misrepresent developmental trends. It is therefore worth noting that all patterns of significance in the corresponding analyses remained the same when excluding this age group from analyses; we can at least be confident that their inclusion is not distorting our comparisons between older children and adults, even if questions about this youngest age group remain.

[4] To analyze the type of question that was asked first—hypothesis scanning versus constraint seeking—we conducted a logistic regression using age group, scenario, and solution probability as predictors. For this analysis we included all participants, even those who did not succeed in finishing the game. A test of the full model against a constant-only model was statistically significant, indicating that the predictors, as a set, reliably distinguished between participants who did and did not ask a constraint-seeking question first ($\chi^2 = 54.704$, $df = 3$, $p < .001$). Nagelkerke's $R^2 = .278$ indicated a moderately weak relationship between prediction and grouping. Prediction success overall was 81% (26% for constraint-seeking and 96% for hypothesis-scanning questions). The Wald criterion demonstrated that only age group ($p < .001$) made a significant contribution to predicting initial question type, whereas scenario and solution probability were not significant predictors. The exp(B) value indicated that older age groups (older children compared to younger children, and young adults compared to older children) had a decreased likelihood of generating an initial hypothesis-scanning question (by .22 times), after controlling for the other factors in the model.

Table 3
Percentage of participants who succeeded in finding the correct solution, displayed by age group, scenario, and solution probability.

Scenario	Solution probability	Age group							
		Younger children		Older children		Young adults		Total	
		Mean (%)	*SD* (%)	Mean (%)	*SD* (%)	Mean (%)	*SD* (%)	Mean (%)	*SD* (%)
Man	High	54	52	87	35	95	22	81	39
	Low	0	0	37	50	93	26	62	50
	Total	23	43	61	50	94	24	61	49
Boy	High	37	50	87	35	100	0	76	43
	Low	6	24	71	47	83	39	48	51
	Total	21	41	79	41	93	25	62	49
Alien	High	69	48	85	38	92	28	81	40
	Low	6	25	50	52	75	45	40	50
	Total	37	49	67	48	84	37	61	49
Total	High	53	50	86	35	96	20	79	41
	Low	4	20	52	51	85	37	43	50
	Total	27	45	69	47	91	29	62	49

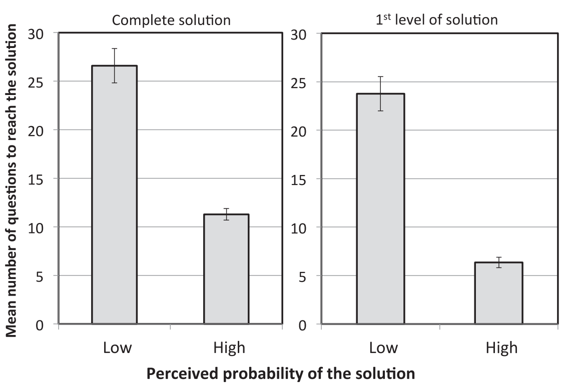

Fig. 2. Mean number of questions participants needed to reach the complete solution (left) and a first level of the solution (right), displayed by solution probability. Error bars represent one SEM in each direction. Note that we collapse across age groups, as there were no interactions between age group and other variables.

likely, presumably because hypothesis-scanning questions were reasonably likely to chance upon the solution.

In sum, we found that the proportion of constraint-seeking questions asked increased sharply with age, and was also higher when the solution probability was low as opposed to high. We also found that constraint-seeking questions "paid off," in terms of efficiency (i.e., number of questions needed to reach the solution), when the solution probability was low, but not when it was high.

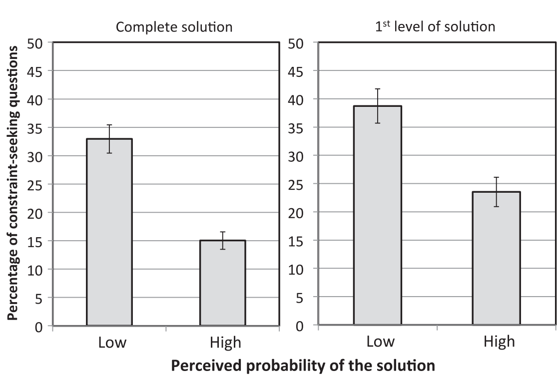

Fig. 3. Percentage of questions that were constraint seeking of those required to reach the complete solution (left) or a first level of the solution (right), displayed by solution probability. Error bars represent one SEM in each direction. Note that we collapse across age groups; although there was a linear increase with age in the percentage of constraint-seeking questions asked, there were no interactions between age group and other variables.

3.2.4. Response to feedback

To better understand the dynamics of the task, we analyzed how the type of question asked changed in response to "yes" versus "no" feedback. We first considered all questions that immediately followed a constraint-seeking question. We ran a mixed ANOVA with age group (3: younger children, older children, young adults), scenario (3: man, boy, alien) and solution probability (2: high probability, low probability) as between subject factors, and two within-subjects factors: Feedback on the preceding constraint-seeking question (2: yes, no) and the proportion of constraint-seeking (versus hypothesis-scanning) questions asked in that position (i.e., immediately following a constraint-seeking question that received "yes" feedback or "no" feedback) as the dependent variable.[5]

This analysis revealed a main effect of question feedback, $F(1,103) = 26.27$, $p < .001$, $\eta^2 = .20$: Participants asked a higher proportion of constraint-seeking questions after receiving "no" feedback on the previous question ($M_{no} = 57\%$, $SD = 37\%$) than after receiving "yes" feedback ($M_{yes} = 23\%$, $SD = 31\%$). There were no additional significant effects.

We ran an equivalent analysis for questions following hypothesis-scanning questions that received "yes" feedback or "no" feedback.[6] Note that this analysis was possible because solutions had multiple levels of detail, and so a response of "yes" to a hypothesis-scanning question did not entail that the solution had been reached—in many cases, inquiry continued. This analysis again revealed a main effect of feedback, $F(1,170) = 40.30$, $p < .001$, $\eta^2 = .19$: Participants asked a higher proportion of constraint-seeking questions after receiving "no" feedback ($M_{no} = 20\%$, $SD = 26\%$) than after receiving "yes" feedback ($M_{yes} = 6\%$, $SD = 19\%$).

There was also an interaction between feedback and solution probability, $F(1,170) = 15.97$, $p < .001$, $\eta^2 = .09$: The effect of feedback on the proportion of constraint-seeking questions asked after

[5] Note that this analysis was restricted to the 109 participants who asked at least one constraint-seeking question that received "yes" feedback and at least one constraint-seeking question that received "no" feedback. This excluded 73% of the younger children, 66% of the older children, and 41% of the young adults, $\chi^2 = 21.17$, $p < .001$.

[6] Note that this analysis was restricted to the 176 participants who asked at least one hypothesis-scanning question that received "yes" feedback and at least one hypothesis-scanning question that received "no" feedback. This excluded 59% of the younger children, 31% of the older children, and 13% of the young adults, $\chi^2 = 42.07$, $p < .001$.

a hypothesis-scanning question was greater in the low-probability solution condition (M_{yes} = 3%, SD = 12%; M_{no} = 29%, SD = 26%) than in the high-probability solution condition (M_{yes} = 9%, SD = 22%; M_{no} = 14%, SD = 24%).

Finally, we found a main effect of age group, $F(2,170) = 13.21$, $p < .001$, $\eta^2 = .13$. A Bonferroni post hoc analysis confirmed that younger children asked a lower proportion of constraint-scanning questions (M_{young_child} = 5%, SD = 12%) than older children (M_{old_child} = 11%, SD = 24%, $p = .032$), who in turn asked a lower proportion of constraint-seeking questions than young adults (M_{young_adult} = 20%, SD = 29%, $p = .024$).

In sum, we found reliable effects of question feedback on the type of question subsequently asked: A response of "no" to either a constraint-seeking question or a hypothesis-scanning question increased the probability of asking a constraint-seeking question. This effect did not interact with age, although we did find that the probability of asking a constraint-seeking question immediately after a hypothesis-scanning question increased with age, and also that the effect of feedback to a hypothesis-scanning question was more pronounced when the solution probability was low.

3.2.5. Manipulation check

Table SI1 of the Supporting Information presents the prior and conditional probabilities that participants estimated for both scenarios, elicited in terms of frequencies (cases out of 100). As expected, participants estimated the frequency of occurrence of the first level of detail of the high probability solution, P(traffic|late), as higher (M = 49.80, SD = 24.63) than the frequency of occurrence of the first level of detail of the low probability solution, P(waiting for plumber|late) (M = 21.63, SD = 26.25), paired t test, $t(268) = 12.01$, $p < .001$, with the conditional probabilities for subsequent levels of detail receiving moderate ratings, as expected (see Supplementary Information B for full statistical analyses). Interestingly, we found that the frequency estimates for the first level of detail generated by young adults (M = 29.70, SD = 14.70) were significantly lower than those of younger children (M = 40.44, SD = 23.06, $p < .001$) and older children (M = 35.92, SD = 30.02, $p = .031$), which is consistent with a lower tendency to ask hypothesis-scanning questions.

3.3. Discussion of Study 2

The results of Study 2 corroborate and extend the findings from Study 1. First, as in Study 1, we found a developmental increase in the proportion of constraint-seeking questions asked, and that children were just as responsive to information structure as were adults (that is, age did not interact with any of our other variables).

Second, participants of all ages asked a higher proportion of constraint-seeking questions in the low probability solution condition than in the high probability solution condition, as was the case in the "mixed distribution" condition from Study 1. As a result, the efficiency of participants' strategies for asking questions arguably *improved* in the face of a low probability solution even though their performance deteriorated (both in terms of success rate and number of questions needed to reach the solution). We also found that the correlation between the proportion of constraint-seeking questions asked and the number of questions needed to reach the solution was significant only in the low probability condition, but not in the high probability condition, suggesting that constraint-seeking questions only "paid off" under these conditions.

Going beyond Study 1, we found that both children and adults changed the kinds of questions that they asked in response to feedback, with a higher proportion of constraint-seeking questions in response to hearing "no." Importantly, this was true whether the preceding question was hypothesis-scanning or constraint-seeking, suggesting that participants were not simply inclined to abandon a particular *question type* because it generated a negative response. Rather, participants moved away from hypothesis-scanning questions as they received evidence that the best candidates for hypothesis-scanning questions were not solutions, and perhaps more generally that their prior beliefs were not an effective guide in the task. It is worth recalling, however, that analyses involving feedback were restricted to those participants who asked questions of each type and that received "yes" and "no" feedback.

Two additional predictions were not confirmed. First, we failed to find an effect of vignette on performance, with a higher proportion of constraint-seeking questions for less familiar scenarios (i.e., the alien attending the reunion). It is likely that this case was sufficiently analogous to familiar situations to provide both children and adults with the prior knowledge necessary to ask effective hypothesis-scanning questions. Indeed, scenario was not a significant factor in any of our analyses—participants required no more questions to reach the solution for the alien than the boy or the man.

Second, we did not find a linear decrease with age in the *number* of questions required to reach the solution. This could be in part because a large proportion of children gave up before reaching the solution. Had these children continued, they would likely have needed even more questions to reach the solution, and this could have resulted in an effect of age on the number of questions required to reach the solution. Nevertheless, some previous studies have found that an increase in the proportion of constraint-seeking questions does not always lead to fewer questions to reach the solution (Denney, 1972; Denney, Denney, & Ziobrowski, 1973; Laughlin, Moss, & Miller, 1969). For one thing, participants' constraint-seeking questions were not always maximally informative. Moreover, in our task, a hypothesis-scanning approach could sometimes lead to a quick win, especially when the solution was a priori likely.

4. General discussion

In two novel experiments, we found that children and young adults could adaptively change the types of questions that they asked in response to the information structure of the task, whether the information structure was specified in advance (Study 1) or inferred in response to feedback (Study 2). In particular, Study 1 showed that when hypothesis-scanning questions were least likely to pay off (i.e., when candidate hypotheses were all equally likely), participants in all age groups increased the frequency with which they asked constraint-seeking questions. This finding suggests that when prior knowledge (strongly) favors some hypotheses over others, participants in the three age groups studied are more likely to ask hypothesis-scanning questions, possibly in the hope of achieving a quick win. Study 2 additionally found that participants change the type of question that they ask *dynamically*, increasing the rate of constraint-seeking questions in response to negative feedback. Notably, these findings extended to those children in our youngest age group (at least, to those we could include in our analyses), even though they overwhelmingly asked hypothesis-scanning questions.

Perhaps surprisingly, neither study found a consistent boost in overall performance with age (i.e., fewer questions required to reach the solution). This implies that young adults' more frequent use of constraint-seeking questions did not yield a (statistically significant) advantage over younger groups. Suggestively, however, their performance was (non-significantly) better than that of younger groups in those conditions for which hypothesis-scanning questions would least pay off: When the solution was low probability (in Studies 1 and 2) and when the hypotheses provided were all equally likely (in Study 1). Asking constraint-seeking questions in other conditions required giving up on the non-negligible chance of obtaining a quick win by correctly guessing the solution. These findings therefore reinforce the point that constraint-seeking questions are not always the best approach.

More generally, the two kinds of questions that we consider might be seen as instances of "exploiting" versus "exploring" (Cohen, McClure, & Yu, 2007; Hills, Todd, & Goldstone, 2008, 2010). When prior knowledge is available and strongly favors some hypotheses over other, a hypothesis-scanning strategy arguably *exploits* this knowledge in the hope of achieving a quick win. In contrast, constraint-seeking questions more efficiently *explore* the broader hypothesis space. Understanding the current findings in these terms suggests promising directions for new research and also suggests a broader framework within which the current proposal can be understood.

Our results suggest that the critical variable influencing information search is the probability distribution over alternative hypotheses, not the general familiarity of the scenario. We had hypothesized, in Study 2, that familiarity would translate into the ability to generate a more complete set of alternative hypotheses, for which the probability distribution would typically be more skewed. We therefore expected that participants would ask a higher proportion of hypothesis-scanning question for more familiar domains, as they did in the mixed distribution condition in Study 1. However, the manipulation of familiarity in Study 2 did not specify or guarantee a particular probability distribution. In fact, a scenario that is highly familiar can involve candidate hypotheses with equal probability (e.g., which number would come up when a child rolls a die). We therefore regard Study 1 as the cleaner test of our initial hypothesis, but find it nonetheless notable that our manipulation of familiarity failed to have reliable effects on information search.

4.1. Relationships to prior research

Prior research has shown that several task features can influence children's and adults' reliance on constraint-seeking questions in a 20-questions task. For example, Siegler (1977) found that 13- and 14-year-old adolescents were influenced by the order in which two isomorphic 20-questions problems were presented (see also Nelson et al., 2014). Also, as mentioned in the introduction, Ruggeri and Feufel (2015) found that describing objects at a basic level increased the proportion of constraint-seeking questions that children asked. These findings, like ours, suggest that younger children have the capacity to ask constraint-seeking questions, even though they fall short of mastery (Denney, 1972, 1975; Denney & Denney, 1974; Denney et al., 1973; Laughlin et al., 1969) and may need scaffolding in some forms (see Ruggeri & Feufel, 2015). However, our results differ importantly from these demonstrations that contextual factors can influence children's questions: Our design manipulated the *effectiveness* of different question types, not the cognitive demands required to enact them.

It is interesting to note that most previous work using the 20-questions task mimics the structure of the uniform distribution condition from Study 1: Participants were presented with a set of prespecified alternatives of equal probability (see Mosher & Hornsby, 1966; Ruggeri & Feufel, 2015; Siegler, 1977). We found that this condition boosted the proportion of constraint-seeking questions among young children, while prior work has found that such children predominantly ask hypothesis-scanning questions. However, any apparent inconsistency would be misleading. We, too, found that younger children predominantly asked hypothesis-scanning questions, and would hesitate to make comparisons in absolute proportions given the many ways in which implementations of the 20-questions task can differ, including the number of items, the specific instructions, and so on. The appropriate conclusion to draw from our results is that even young children appropriately *increase* the frequency with which they ask constraint-seeking questions under these conditions (as compared with a mixed distribution), not that such questions are in fact frequent.

Our finding that the proportion of constraint-seeking questions increased with age replicates prior work (see Herwig, 1982; Mosher & Hornsby, 1966; Ruggeri & Feufel, 2015; Siegler, 1977), and is additionally consistent with research on children's decision-making. For example, Davidson (1991a) presented children with an information board from which they could choose which information to reveal prior to making a decision. She found that 2nd grade children were generally exhaustive and inefficient in their search, whereas 5th and 8th grade children implemented more demanding but more efficient strategies (see also Davidson & Hudson, 1988; Gregan-Paxton & Roedder John, 1995; Wellman, 1985; Woody-Ramsey & Miller, 1988). Although our task differed considerably in its form (i.e., asking questions) and in the particular manipulations of information structure (e.g., the frequency distribution over hypotheses), these findings are in line with the idea that with age, children are better able to identify and implement efficient strategies for information search (see also DeMarie-Dreblow & Miller, 1988).

It is potentially surprising, then, that children's increasing sophistication over development did not result in greater sensitivity to variations in information structure across conditions. Specifically, we did not find that older participants were more responsive to our experimental manipulations than were younger children (bearing in mind that some analyses were restricted to a smaller subset of younger children). This suggests that the developmental factors that result in a greater proportion of constraint-seeking questions *across* age groups may not be the same as those that we successfully manipulated *within* age groups. In particular, the shift in constraint-seeking questions across age groups could reflect more general developmental changes, such as those in the ability to organize, cluster, and represent information at different hierarchical levels (Hollister Sandberg, Huttenlocher, & Newcombe, 1996), or in cognitive flexibility (Anderson, 2002; Legare et al., 2013). It could also relate to more general developmental changes in language and working memory. Our experiments, like the vast majority of previous studies of children's strategies for inquiry, did not include independent measure of these factors; doing so would be a valuable direction for future work.

4.2. A heuristic model of question asking

What underlies children's strategies for inquiry in the real world and how they change in response to context? One possibility is that children explicitly assess which type of question is likely to be most efficient in a given context. This seems unlikely, however, given the difficulty of making such an assessment in a real world situation, as well as the fact that cues pointing to the effectiveness of a given question are at best indirect. We suggest instead that for complex, real world problems, children (and adults) might rely on a heuristic procedure that is likely to yield flexible, dynamic, and efficient behavior in a range of contexts. Specifically, we propose that people simply ask themselves a single question: Is one of the given/self-generated alternative hypotheses sufficiently *more* likely than the others? If so, test it directly (with a hypothesis-scanning question); otherwise, collect information to reduce the number of alternative hypotheses (using a constraint-seeking question).

This heuristic involves three crucial components: (a) a set of alternative hypotheses to be considered; (b) an absolute or relative probability assignment for each hypothesis included in the set; and (c) a threshold defining how much more likely than alternatives (or how likely in absolute terms) a hypothesis must be to trigger a hypothesis-scanning question. While the first two components can be either self-defined or externally given, the third component (the threshold) will typically be self-defined and can depend on many factors, including the underlying motivation of the agent (e.g., avoiding mistakes versus striving for speed) and the agent's attitude toward risk. This threshold might also be implicitly defined by information gain: A hypothesis-scanning question could be triggered when it is about as efficient as the best constraint-seeking question one could generate. At every step of this cyclic process, the three components can be redefined. For example, some hypotheses will be eliminated as new information becomes available, new hypotheses may come to mind and change the balance of probabilities within the preexisting set, and a person's motivation might change (e.g., because efficiency might become more important as time goes on).

Our results are in line with this heuristic model. In Study 1, participants asked a higher proportion of hypothesis-scanning questions in the mixed distribution condition, where some hypotheses were more likely than others. Moreover, in Study 2, participants asked a higher proportion of constraint-seeking questions after negative feedback, suggesting that with the elimination of especially likely hypotheses, the resulting distribution was more uniform. Indeed, the hypothesis spaces generated by children and adults in a similar causal inference task (i.e., "Why was John yesterday late to work?") resemble our mixed information structure task, with a small number of very likely hypotheses and a large number of lower-probability hypotheses, yielding a relatively uniform distribution once the top few hypotheses are eliminated (Ruggeri, Abbott, Lombrozo, & Griffiths, 2015). Additionally, negative feedback might have increased the threshold, triggering a hypothesis-scanning question by suggesting that "exploring" would be more valuable than "exploiting."

This heuristic is plausible from a developmental perspective, as it requires relatively low cognitive effort and takes into account only a few pieces of information at a time. To apply the heuristic, the agent does not need to generate an exhaustive set of hypotheses or order them all in terms of probability. The agent could in theory generate only one hypothesis at a time or assess only relative rather than absolute probabilities when comparing more than one hypothesis. While people may be able to use explicit probabilistic information when it is available, as was the case in our Study 1, real-world cases are more likely to involve an approximation of this information using previously-identified processes, such as fluency (Hertwig, Herzog, Schooler, & Reimer, 2008) and availability (Tversky & Kahneman, 1973). In fact, if hypotheses tend to come to mind in a way that reflects their objective probabilities, relying on a single hypothesis might have minimal costs for performance.

This rough heuristic model helps identify possible sources of developmental change. Specifically, it could be that an important developmental difference lies in the initial hypothesis-generation phase, with young children simply generating fewer hypotheses than young adults, or in the testing phase, with children adopting a lower threshold for pursuing a hypothesis-scanning question. It could also be that children tend to ask hypothesis-scanning question because they are poor at generating efficient constraint-seeking questions. With poor constraint-seeking questions, the information gain for the two types of questions might be similar, and the possibility of obtaining a quick-win with a hypothesis-scanning question could make it more attractive. We would also expect developmental differences if children and adults make different assumptions about the absolute or relative probabilities of candidate hypotheses, and our data provide some evidence that this is the case: In the manipulation check for Study 2, we found that young adults judged the solutions less likely than did older or younger children, consistent with asking a lower proportion of hypothesis-scanning questions.

5. Conclusion

The process of asking questions plays a crucial role in learning and development. In addition to solving everyday, practical problems, children must also acquire concepts and knowledge structures to help them understand the world, including physical, psychological, and biological phenomena. However, little is known about the qualitative and quantitative features of children's questions in a systematic search for information. Here we documented the adaptive and dynamic flexibility of children's questions and proposed a tentative heuristic model. This paper represents a first attempt to investigate what might be called "ecological learning," that is, how children develop the ability to adapt their learning to different information structures and environments.

The tradition of studying children's question-asking behavior by using the 20-questions game emerged with Mosher and Hornsby (1966), who presented two versions of the task: one involving the selection of an object from a pre-defined set of alternatives, and the other an open-ended causal attribution task. However, almost all subsequent work has focused on the former version of the task instead of the causal attribution version on which we focus here. While the formal considerations we articulate, as well as the heuristic model we provide, can in principle apply to all cases, it is an interesting and open question whether the domain of causal attribution might have unique characteristics.

We hope that these initial steps help pave the way for additional research involving a broader range of question types and tasks, and incorporating strategies for information search beyond asking questions, such as direct observation and experimentation.

Acknowledgments

We would like to thank all the students and teachers from Fondazione Sacro Cuore, Fondazione San Carlo Borromeo, Istituto Sacro Cuore, Scuola Antonio Benci, and Liceo Scientifico F. Enriques of Livorno, who participated in our study. We would also like to thank Marianna Sgherri, Selica Vicidomini, Aurora Fontanelli, Giorgio Zalum, Chiara Chiellini, Riccardo Ruggeri, and Claudia Mazzeranghi for helping us run the experiments, as well as Anita Todd and Nicolai Bodemer for their support and helpful comments. The first author was partially supported by a Marie Curie International Outgoing Fellowship within the 7th European Community Framework Programme. The second author was partially supported by grant DRL-1056712 from the National Science Foundation and a McDonnell Foundation Scholar Award in Understanding Human Cognition.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2015.07.004.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.
Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology: A Journal on Normal and Abnormal*

Development in Childhood and Adolescence, 8(2), 71–82. http://dx.doi.org/10.1076/chin.8.2.71.8724.

Berretty, P. M., Todd, P. M., & Blythe, P. W. (1997). Categorization by elimination: A fast and frugal approach to categorization. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the nineteenth annual conference of the cognitive science society* (pp. 43–48). Mahwah, NJ: Lawrence Erlbaum Associates.

Berretty, P. M., Todd, P. M., & Martignon, L. (1999). Categorization by elimination: Using few cues to choose. In G. Gigerenzer, P. M. Todd, & the ABC research group (Eds.), *Simple heuristics that make us smart* (pp. 235–254). New York: Oxford University Press.

Callanan, M. A., & Oakes, L. M. (1992). Preschoolers' questions and parents' explanations: Causal thinking in everyday activity. *Cognitive Development, 7*, 213–233. http://dx.doi.org/10.1016/0885-2014(92)90012-G.

Chater, N., & Oaksford, M. (1999). Ten years of the rational analysis of cognition. *Trends in Cognitive Sciences, 3*(2), 57–65.

Chouinard, M. M. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development, 72*, vii–ix. http://dx.doi.org/10.1111/j.1540-5834.2007.00412.x.

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 362*(1481), 933–942. http://dx.doi.org/10.1098/rstb.2007.2098.

Davidson, D. (1991a). Children's decision making examined with an information-board procedure. *Cognitive Development, 6*, 77–90.

Davidson, D. (1991b). Developmental differences in children's search of predecisional information. *Journal of Experimental Child Psychology, 52*, 239–255. http://dx.doi.org/10.1016/0022-0965(91)90061-V.

Davidson, D., & Hudson, J. (1988). The effects of decision reversibility and decision importance on children's decision making. *Journal of Experimental Child Psychology, 46*, 35–40. http://dx.doi.org/10.1016/0022-0965(88)90021-5.

DeMarie-Dreblow, D., & Miller, P. H. (1988). The development of children's strategies for selective attention: Evidence for a transitional period. *Child Development*, 1504–1513.

Denney, D. R. (1972). Modeling and eliciting effects upon conceptual strategies. *Child Development, 43*, 810–823. http://dx.doi.org/10.1111/j.1467-8624.1972.tb02036.x.

Denney, D. R. (1975). The effects of exemplary and cognitive models and self-rehearsal on children's interrogative strategies. *Journal of Experimental Child Psychology, 19*, 476–488. http://dx.doi.org/10.1016/0022-0965(75)90077-6.

Denney, D. R., & Denney, N. W. (1973). The use of classification for problem solving: A comparison of middle and old age. *Developmental Psychology, 9*, 275–278. http://dx.doi.org/10.1037/h0035092.

Denney, N. W., & Denney, D. R. (1974). Modeling effects on the questioning strategies of the elderly. *Developmental Psychology, 10*, 458. http://dx.doi.org/10.1037/h0036444.

Denney, D. R., Denney, N. W., & Ziobrowski, M. J. (1973). Alterations in the information-processing strategies of young children following observation of adult models. *Developmental Psychology, 8*, 202–208. http://dx.doi.org/10.1037/h0034144.

Frazier, B. N., Gelman, S. A., & Wellman, H. M. (2009). Preschoolers' search for explanatory information within adult-child conversation. *Child Development, 80*, 1592–1611. http://dx.doi.org/10.1111/j.1467-8624.2009.01356.x.

Gopnik, A., & Meltzoff, A. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (1999). *The scientist in the crib: Minds, brains, and how children learn*. Cambridge, MA: MIT Press.

Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257–293). New York: Cambridge University Press.

Graesser, A. C., & McMahen, C. L. (1993). Anomalous information triggers questions when adults solve quantitative problems and comprehend stories. *Journal of Educational Psychology, 85*(1), 136–151.

Graesser, A. C., & Olde, B. A. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology, 95*(3), 524–536.

Green, L., & Mehr, D. R. (1997). What alters physicians' decisions to admit to the coronary care unit? *Journal of Family Practice, 45*(3), 219–226.

Gregan-Paxton, J., & Roedder John, D. (1995). Are young children adaptive decision makers? A study of age differences in information search behavior. *Journal of Consumer Research, 21*, 567–580.

Greif, M. L., Kemler Nelson, D. G., Keil, F. C., & Gutierrez, F. (2006). What do children want to know about animals and artifacts? Domain-specific requests for information. *Psychological Science, 17*, 455–459. http://dx.doi.org/10.1111/j.1467-9280.2006.01727.x.

Hamilton, G. C., Sanders, A., Strange, G. S., & Trott, A. T. (2003). *Emergency medicine: An approach to clinical problem solving* (2nd ed.). Philadelphia: Saunders Co.

Harris, P. L. (2012). *Trusting what you're told: How children learn from others*. Cambridge, MA: Belknap Press/Harvard University Press.

Hertwig, R., Herzog, S. M., Schooler, L. J., & Reimer, T. (2008). Fluency heuristic: A model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1191–1206. http://dx.doi.org/10.1037/a0013025.

Herwig, J. E. (1982). Effects of age, stimuli, and category recognition factors in children's inquiry behavior. *Journal of Experimental Child Psychology, 33*, 196–206. http://dx.doi.org/10.1016/0022-0965(82)90015-7.

Hickling, A. K., & Wellman, H. M. (2001). The emergence of children's causal explanations and theories: Evidence from everyday conversation. *Developmental Psychology, 37*, 668–683. http://dx.doi.org/10.1037/70012-1649.37.5.668.

Hills, T., Todd, P. M., & Goldstone, R. L. (2008). Search in external and internal spaces: Evidence for generalized cognitive search processes. *Psychological Science, 19*, 676–682.

Hills, T. T., Todd, P. M., & Goldstone, R. L. (2010). The central executive as a search process: Priming exploration and exploitation across domains. *Journal of Experimental Psychology: General, 139*(4), 590–609. http://dx.doi.org/10.1037/a0020666.

Hollister Sandberg, E. H., Huttenlocher, J., & Newcombe, N. (1996). The development of hierarchical representation of two-dimensional space. *Child Development, 67*(3), 721–739. http://dx.doi.org/10.1111/j.1467-8624.1996.tb01761.x.

Khemlani, S. S., & Johnson-Laird, P. N. (2011). The need to explain. *Quarterly Journal of Experimental Psychology, 64*(11), 2276–2288. http://dx.doi.org/10.1080/17470218.2011.592593.

Laughlin, P. R., Moss, I. L., & Miller, S. M. (1969). Information-processing in children as a function of adult model, stimulus display, school grade, and sex. *Journal of Educational Psychology, 60*, 188–193. http://dx.doi.org/10.1037/h0027556.

Legare, C. H., Gelman, S. a., & Wellman, H. M. (2010). Inconsistency with prior knowledge triggers children's causal explanatory reasoning. *Child Development, 81*(3), 929–944. http://dx.doi.org/10.1111/j.1467-8624.2010.01443.x.

Legare, C. H., Mills, C. M., Souza, A. L., Plummer, L. E., & Yasskin, R. (2013). The use of questions as problem-solving strategies during early childhood. *Journal of Experimental Child Psychology, 114*, 63–76. http://dx.doi.org/10.1016/j.jecp.2012.07.002.

Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences, 10*, 464–470. http://dx.doi.org/10.1016/j.tics.2006.08.004.

Lombrozo, T. (2012). Explanation and abductive inference. In K. J. Holyoak & R. G. Morrison (Eds.), *Oxford Handbook of Thinking and Reasoning* (pp. 260–276). Oxford, UK: Oxford University Press.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York, NY: Henry Holt and Co., Inc.. 2.

Martignon, L., Katsikopoulos, K. V., & Woike, J. K. (2008). Categorization with limited resources: A family of simple heuristics. *Journal of Mathematical Psychology, 52*, 352–361. http://dx.doi.org/10.1016/j.jmp.2008.04.003.

Martignon, L., & Krauss, S. (2009). Hands-on activities for fourth graders: A tool box for decision-making and reckoning with risk. *International Electronic Journal of Mathematics Education, 4*, 227–258.

Mills, C. M., Legare, C. H., Bills, M., & Mejias, C. (2010). Preschoolers use questions as a tool to acquire knowledge from different sources. *Journal of Cognition and Development, 11*(4), 533–560. http://dx.doi.org/10.1080/15248372.2010.516419.

Mills, C. M., Legare, C. H., Grant, M. G., & Landrum, A. R. (2011). Determining who to question, what to ask, and how much information to ask for: The development of inquiry in young children. *Journal of Experimental Child Psychology, 110*(4), 539–560. http://dx.doi.org/10.1016/j.jecp.2011.06.003.

Mosher, F. A., & Hornsby, J. R. (1966). On asking questions. *Studies in cognitive growth*. New York, NY: Wiley.

Nelson, J. D., Divjak, B., Gudmundsdottir, G., Martignon, L. F., & Meder, B. (2014). Children's sequential information search is sensitive to environmental probabilities. *Cognition, 130*(1), 74–80. http://dx.doi.org/10.1016/j.cognition.2013.09.007.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101*, 608–631.

Piaget, J. (1954). *The construction of reality in the child*. New York: Ballantine Books.

Ruggeri, A. (2012). *Opening up the cuebox: A developmental perspective*. Doctoral dissertation, Humboldt-Universität zu Berlin, Germany.

Ruggeri, A., & Feufel, M. A. (2015). How basic-level objects facilitate question-asking in a categorization task. *Frontiers in Psychology, 6*, 1–13. http://dx.doi.org/10.3389/fpsyg.2015.00918.

Ruggeri, A., Abbott, J., Lombrozo, T., & Griffiths, T. L. (2015). *[Hypothesis-generation]*. Unpublished raw data.

Ruggeri, A., Lombrozo, T., Griffiths, T. L., & Xu, F. (2015). Children search for information as efficiently as adults, but seek additional confirmatory evidence. In *Proceedings of the 37th annual conference of the Cognitive Science Society*, in press.

Siegler, R. S. (1977). The twenty questions game as a form of problem solving. *Child Development, 48*, 395–403. http://dx.doi.org/10.1111/j.1467-8624.1977.tb01176.x.

Todd, P. M., Gigerenzer, G., & the ABC Research Group (2012). *Ecological rationality: Intelligence in the world*. New York, NY: Oxford University Press.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*(2), 207–232. http://dx.doi.org/10.1016/0010-0285(73)90033-9.

Wellman, H. M. (1985). *Children's searching: The development of search skill and spatial representation*. Hillsdale, NJ: Erlbaum.

Woody-Ramsey, J., & Miller, P. H. (1988). The facilitation of selective attention in preschoolers. *Child Development, 59*, 1497–1503. http://dx.doi.org/10.1111/j.1467-8624.1988.tb03677.x.