Children Use Causality to Guide Question Asking

Elizabeth Lapidow (elapidow@ucsd.edu)
Department of Psychology, University of California, San Diego

Amberley Stein (amberley.stein@gmail.com)
Department of Psychology, University of California, San Diego

Caren M. Walker (carenwalker@ucsd.edu)
Department of Psychology, University of California, San Diego

Abstract

Gathering information via question asking is an essential and effective tool for learning. However, it also requires learners to select from a near infinite space of possible queries. Here, we investigate a potentially powerful guide for question asking in young learners: the relationship between cause and effect. Children (5- and 7-year-olds) read a storybook about an event with an unknown cause and made several choices between two questions to ask about possible candidate causes. Both questions revealed similar information, but only one had the potential to determine whether a candidate was capable of causing the event described. Participants overwhelmingly selected causally relevant over irrelevant questions, with strong performance in both age-groups and for all types of information. These results suggest that young learners employ their prior knowledge of the causal connections between events to identify relevant queries during information search.

Keywords: Question Asking; Causality; Cognitive Development; Information Search

Introduction

Being able to ask good questions is a critical skill for learning. Particularly in childhood, querying knowledgeable sources offers an accessible means of gathering information that may be too costly or impossible to investigate directly. Imagine you want to find out what caused a tree to fall down. You can't directly observe what happened after the fact, and lack the arboreal expertise to make an accurate diagnosis. First-hand investigation is therefore inefficient, especially if a source of second-hand information on the topic is available for questioning.

However, asking questions that are likely to produce desired information is a nontrivial challenge: options for how to focus and phrase even a simple query are vast, and only some combinations are valuable. For example, "What causes trees to fall?" is relevant to the current learning goal, but includes so many possibilities (storms, fungus, earthquakes, termites, nearby construction, too much water, not enough water, and more) that the answer to this question still leaves a great deal of uncertainty. In contrast, the answer to "When was this tree planted?" will be highly specific, but is unlikely to be useful for discovering the cause. Instead, successful question-askers must use what they know about the target of their inquiry to arrive at precise requests for relevant information (e.g., Ronfard et al., 2018; Van der Meij, 1990).

The present study investigates whether young learners can use their understanding of the causal connections between events to evaluate questions during information search. Previous research suggests that, by age five, children have the ability to distinguish more from less effective questions (e.g., Ruggeri et al., 2017). However, this prior work has typically explicitly presented a set of possibilities and asked children to evaluate a question's effectiveness in terms of how many possibilities it targeted, rather than drawing on their conceptual knowledge (see also O'Neill, 2021; Ruggeri & Lombrozo, 2015). Here, we ask children to apply their causal knowledge to evaluate potential questions. Since even young children have expectations about the relationships between causes and effects that guide their inferences (Magid et al., 2015; Tsividis et al., 2015), their existing understanding of the causal world may provide a powerful means of evaluating questions (Schulz, 2012). Below, we briefly review prior work before introducing the current task, which examines whether young learners' apply their causal knowledge to evaluate potential questions during information search.

Selective Question-Asking in Childhood

Even in early childhood, children appreciate that questions differ in their relevance and effectiveness. The robust literature on the development of question-asking suggests that, by the end of preschool, they are already formulating sophisticated and complex questions aimed at gaining specific information (e.g., Callanan & Oakes, 1992; Frazier et al., 2009). Children's questions are appropriate and relevant to the domain in which they are trying to learn (Greif et al., 2016), and they largely avoid asking redundant or ineffective questions (Legare et al., 2013).

Of particular importance to the current investigation, young learners also recognize when some questions are better than others, given differences in the information structure of the problem they are trying to solve. For example, Ruggeri et al. (2017) told 5-year-olds a story in which a character is late to school several days in a row for a variety of reasons (e.g., their bike was broken, they could not find their books, they overslept, they could not find their shoes). On the final day, the character is late for an unknown reason, and two other characters try to find out why by asking as few questions as possible. One character asked about a single possible cause (e.g., "Were you late because you overslept?"), while the

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

other queried multiple possibilities (e.g., "Were you late because you could not find something?"). Children correctly identified the latter question as more effective in searching for the unknown cause. Critically, however, when a single cause (e.g., 'oversleeping') led the character to be late on the majority of the previous days, children's preference for which question to ask reversed. These results suggest that, by age 5, learners are sensitive to the relationship between what they are trying to learn and the type of question that is most informative. Understanding this relationship is critical, both to asking effective questions and in constraining the space of possible questions to consider (Chu et al., 2019).

In the present study, we investigate whether young children can still distinguish between more and less effective questions when informativeness is determined by an abstract relationship between cause and effect. Unlike attending to the number of possibilities targeted by a question (Ruggeri et al., 2017), using causal information to constrain question-asking requires a rich conceptual inference from prior knowledge. For example, if the fallen tree has a splintered bend at the base, then the cause of the fall must have been one that delivered a forceful impact or weighty gravitational pull, whereas a clean cut between base and trunk could only have been caused by human intervention. This recognition facilitates more effective question-asking by constraining the space of questions worth asking: "Has there been heavy snowfall?" (splinter), versus, "Could the growth of the tree pose a safety hazard?" (cut). This kind of reasoning can also guide learners away from questions that are topically relevant, but uninformative; for example, "Have there been any earthquakes lately?" targets a likely cause of fallen trees, but one that is irrelevant in both the cases described above.

Conceptual Knowledge in Children's Question Evaluation

To our knowledge, this study is the first to examine whether recognition of the abstract relationships between causes and effects influences children's evaluation of questions. Previous research has shown that even young children have expectations about cause-effect connections that guide their reasoning. Magid et al. (2015) and Tsividis et al., (2015) found that 4- to 6-year-olds use assumptions of shared features (movement dynamics, proportion, etc.,) to select likely causes of observed effects. However, there is little evidence that young children can evaluate questions based on their conceptual relevance. In one study, 7- to 14-year-olds learned about a character who was tasked with making a decision from among several options (e.g., which video game to buy), and were able to gather information about the options by uncovering items on an 'information board.' (Davidson, 1991). This board listed six dimensions of information (*Length of Game*, *Cost of Game, Chances of Winning*, etc.) about each of the character's six options. A dimension was considered relevant if it corresponded to one of the character's desires mentioned in the story (e.g., if the character wanted an "outer space type adventure game," then the '*Adventure Type'* dimension was deemed relevant). Problematically, these low-level matches between the text of the story and the information board largely removed the need for *inferring* relevant questions. Despite this issue, younger children apparently struggled to make connections using conceptual information: 7- to 10-year-olds were equally likely to reveal information from relevant and irrelevant dimensions. However, given that the task design required children to read and keep track of a great deal of information at once, this study likely underestimated the effectiveness of children's information search.

More recently, Jirout and Klahr (2020) investigated 4- to 7-year-olds' ability to distinguish relevant from irrelevant questions. In this study, participants learned about a character who wanted to find out the identity of an animal living in the woods near their home. The experimenter provided children with six questions that the character had asked, along with the answers to each (e.g., "What does the animal eat?… berries."). Children were then asked to sort these question-answer pairs into either "helpful" or "unhelpful" categories. They found that children's ability to correctly categorize irrelevant question as 'unhelpful' improved with age, but all children had a tendency to categorize the majority of questions as 'helpful.' Although these findings also suggest limitations in children's ability to apply their conceptual knowledge in evaluating questions, the knowledge that the character *chose* to ask these questions may have biased their responses.

The Present Study

Despite evidence that children effectively evaluate questions by the early school years, prior work also suggests that this might not extend to instances where they must rely on their real-world knowledge to do so. Here, we designed a simple paradigm to examine whether young learners (aged 5 and 7 years) employ their prior understanding of causal relationships to guide their evaluation of potential questions during information search.

In this novel task, children receive a description of an event with an unknown cause. The description includes several details that imply necessary characteristics of the event's cause (e.g., a mess is made on a high surface, suggesting that the agent who caused the mess must have been able to reach it). Following this description, participants are asked to make several choices of which of two questions to ask about candidate causes of the event. Both of these questions targeted the same kind of information, but only one of the two requested information with a potential causal connection to a detail from the story. If children utilize their understanding of causal relationships to guide their question evaluation, we would expect a preference for questions with the potential to determine the cause.

Experiment 1

The experimental design and analyses for this study were preregistered prior to beginning data collection (see: https://aspredicted.org/M8H_P2M).

The study procedure consisted of three phases: *Story*, *Choice*, and *Guessing*. In the *Story* phase, participants listened to a story about a character who wants to find out which one of her three pets made the mess in her kitchen. The story included three details about the mess: (1) it was on top of the fridge, (2) one of four items had been taken, and (3) black fur was left behind. Each of these details implies a characteristic of the cause of the mess: the ability to reach a high surface, a preference for the item taken over the ones left behind, and having black fur, respectively. After listening to the story, participants were told that only one of the three pets could have made the mess, and the goal was to find out which one it was by asking questions about them.

In the *Choice* phase, children were presented with three trials where they had to choose which questions to ask about each of the pets. There were two questions on each trial: both asked for the same kind of information, but only one of the two had the potential to reveal whether a pet had one of the characteristics implied by the story. On the *Ability* trial, for example, participants were given a choice between asking about each pets' ability to climb or to make noise. Both questions target similar and equally unknown information, but an answer to the former might determine whether or not a pet has the ability to reach high surfaces necessary to be a cause of this mess. The primary outcome of interest is whether children preferentially select these questions that target information that is potentially relevant for determining the event's unknown cause.

In the final *Guessing* phase, participants were given the answers to all six questions (three relevant and three irrelevant) for each of the three pets and asked which one made the mess. Of these candidate causes, one was Correct (all three relevant features were consistent with being able to make the mess), one was Incorrect (the relevant features were inconsistent with being able to make the mess), and one was a Distractor (all but one of the relevant features were consistent with being able to make the mess). Performance in this phase served to check whether children had the minimum causal knowledge needed to recognize the connections in the task.

Participants

The final sample will include a total of 96 participants (48 per age group). The sample size was determined using an a priori power analysis. Our effect size ($h = 0.41$) was estimated from the results of Ruggeri et al. (2017), which conducted a similar type of investigation (a forced-choice between a better and worse information seeking question) with a similar age group. This analysis indicated a minimum sample size of 46 to achieve the desired power ($\beta = 0.8$) and significance level ($\alpha = 0.05$). We rounded this minimum sample to 48 to accommodate counterbalancing.

Forty-six children (19 female, $M = 78.7$ mos., $SD = 13.27$ mos., range: 60-95 mos.) have participated in the study so far, including 23 five-year-olds ($M = 66.07$ mos., $SD = 3.42$ mos., range: 60-70 mos.) and 23 seven-year-olds ($M = 91.33$ mos., $SD = 3.87$ mos., range: 84-95 mos.). Participants were recruited through a university database of families collected through in-person participation at local schools and museums, or in response to online advertisements (including Facebook ads and online testing listings). Informed consent was obtained for all participants prior to participation in the study, and families received a $3 Amazon gift card as a gift.

Materials

The study was conducted via online video calling and was presented using a series of still images and animated PowerPoint slides. These materials are all available on osf.org at [blinded]. Table 1 shows the exact wording of these elements and how they correspond. The *Preference* trial was designed to rule out the possibility of children selecting questions by semantic association or matching the description wording to the question text, rather than reasoning about causal connections. Which of the four items (chips, cookies, ball, and rattle) was missing differed across participants, but all four were listed and shown in the kitchen scene. Thus, selecting questions based on association or low-level matching would not result in a preference between 'playing' and 'eating' questions. Also, because we counterbalanced the missing item, each question on this trial served as both a relevant and irrelevant option (i.e., 'playing' was the relevant question for half of the participants and the irrelevant question for the other). Also, while the connection between the story and the question on the *Appearance* trial is certainly causal (piles of black fur are unlikely to have been caused by an animal with fur of any other color), the use of the word 'fur' in both story and question allows for a low-level match. This design choice was intentional; poor performance would indicate children were not grasping the premise of the task. All other trials had no repeated language between item and question.

Procedure

At the beginning of each testing session, the experimenter greeted the child and gave the parent a brief overview of the task. She showed the cover of the storybook to participants and said: "Today, we're going to read a story called 'Who made the mess?' This story is a mystery—that means the story doesn't tell us who made the mess. It's your job to look for clues and figure it out." The experimenter then advanced to the first page of the storybook.

Story Phase The first page of the storybook showed a drawing of a girl introduced as Amy. The storybook then explained that Amy has three pets, and each sleeps in a colorful house. The next page showed three dog houses: red, blue, and yellow. Children were asked to point to each house and name the color, which served as a warm-up to familiarize participants with the mode of responding used later in the task.

The next page presented the description of the mess. Children watched a simple animation of Amy entering an illustration of a kitchen as the experimenter said, "One day, Amy comes home from school to find a *big* mess in her

Table 1. Stimuli descriptions used in the *Story* and *Choice* phases of the task.

Feature Type	Clue from Storybook	Causally Relevant Question	Causally Irrelevant Question
Ability	*"The mess was on top of the fridge."*	Can the pet climb?	Can the pet make noise?
Preference*	*"In the cabinet there were the chips, the cookies, the rattle, but the ball was missing"*	What does the pet like to play with?	What does the pet like to eat?
Appearance	*"There was black fur laying all around…"*	What color fur does the pet have?	What color eyes does the pet have?

**Note.* Which of the four named items was missing in the *Preference* trial was counterbalanced across participants. Therefore, which of the two questions ('eat' or 'play') was relevant was the opposite for half of the data collected.

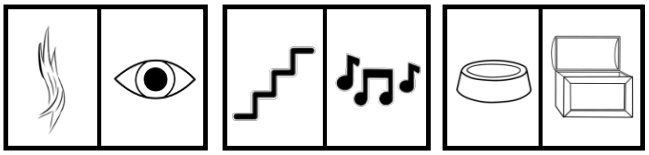
kitchen. She wants to figure out which one of her pets made the mess and looks around for clues." Children were told that Amy noticed that, "The mess was up on top of the fridge, where she keeps the pet toys and treats. There was the rattle, the ball, and the chips, but the cookies were missing. And there was black fur all around." As the experimenter mentioned each detail, the portion of the illustration depicting that detail was briefly highlighted by a glowing yellow outline. The experimenter told children that only one of Amy's pets could have made the mess, and they needed to find out which one by asking Amy questions to learn more about them.

Choice Phase Participants completed three trials (*Ability*, *Preference*, and *Appearance*) in which they had to choose one of two questions to ask about each of the pets. Each pair of questions was presented on a slide using two black-and-white images of equal size, divided by a vertical black line (Fig.1). The experimenter also verbally described the choice. For example, on the *Ability* trial, she said, "For each pet, we can ask 'Can it climb?' or 'Can it make noise?'" as the corresponding images appeared on the screen. The experimenter reiterated the options while highlighting each image again to ask for the child's choice, "Do you want to ask about climbing, or do you want to ask about noise?" If a child did not respond, the experimenter would repeat the question twice before moving on to the next trial. The order of trials and the order and position (on the left or right) of each question was counterbalanced across participants.

Guessing Phase After completing all three forced-choice questions, children moved on to the last phase of the task. The experimenter showed the kitchen scene again and reminded children of the description. The next slide showed the three pet houses from earlier in the task. Inside of each house were images representing the answers to each of the six questions. These were the same images from the *Choice* trials, modified to convey the features of each pet (e.g., the staircase image overlaid with a red 'X' represented 'cannot climb' and wrapped candy superimposed on the food bowl represented 'likes to eat sweet things'). For the Correct pet, children were told: "The pet in the [color] house has green eyes, black fur, it can climb, it cannot make noise, it likes to eat sweet things[1], and it likes to play with noisy things." The Incorrect pet "has black eyes, white fur, cannot climb, cannot make noise, likes to eat salty things and likes to play with bouncy things." The Distractor pet, "has blue eyes, black fur, it cannot climb, can make noise, likes to eat sweet things, likes to play with noisy things." The image of the feature also flashed briefly when mentioned. The presentation order for the Correct, Incorrect, and Distractor pets was counterbalanced across participants.

After listing the features of all three pets, the experimenter asked, "Which pet do you think made the mess in the kitchen? Can you point to it?" Following their response, all children saw a final page with an image of a black cat carrying the stolen object next to their chosen house. The experimenter congratulated the child on succeeding and ended the call.

Figure 1. The presentation of question choice on each of the three *Choice* phase trials.



Appearance *Ability* *Preference*

Results

Prior to testing the hypothesis that children preferentially select relevant questions, it was necessary to determine whether other factors led to differences in children's choice behavior. We conducted a mixed-effects binomial logistic regression with age group (5 years, 7 years), question target (ability, preference, appearance), and story type (missing toy, missing treat) as fixed effects and participant included as a random effect. The total explanatory power of the fitted

[1] To account for counterbalancing the Preference trial, there two versions of each pet (e.g., preference for "sweet things" was changed to "salty things" when the chips were missing).

model was substantial (conditional R2 = 0.39). None of the fixed effects included in the model were significant (all *p's* > 0.1). That children's behavior was consistent regardless of surface-level content is not necessarily surprising, and consistent with our hypothesis that learners employ a general understanding of causal relevance to guide their choice of queries. Therefore, we collapse over these factors in the subsequent analysis.

Consistent with our hypothesis, children showed a clear preference for asking relevant over irrelevant questions (Fig.2). Overall, 85.51% of the questions selected during the Choice phase were relevant, significantly greater than expected by chance ($p < 0.0001$, 95% CI [0.78, 0.91], two-tailed binomial). This preference for relevant questions was seen in both age groups (5-year-olds, 81.16%; 7-year-olds, 89.86%) and across all three question targets (*Ability,* 84.78%; *Preference,* 84.78%; *Appearance,* 86.96%).

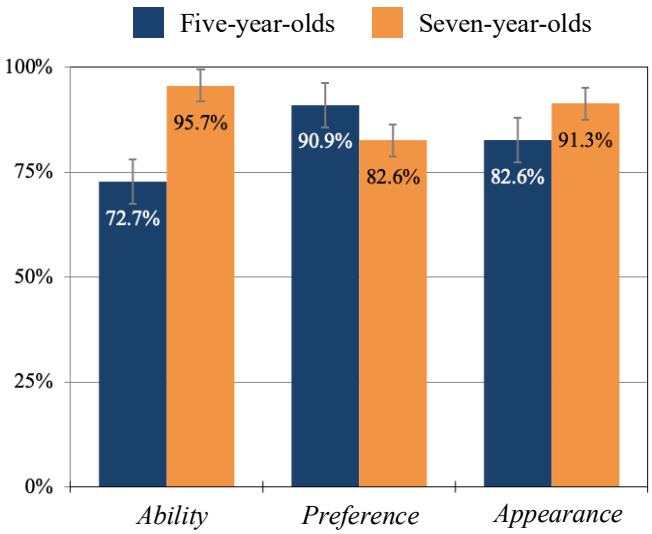
We also examined children's ability to identify the cause that met the required criteria in the Guessing phase. When all the information (both relevant and irrelevant) was made available to them, a significant majority of participants (84.78%) selected the candidate with the correct features to be the cause ($p < 0.0001$, 95% CI [0.71, 0.94], two-tailed binomial). There was also a significant relationship between performance on the *Choice* and *Guessing* phases of the task (Fisher Exact, $p = 0.03$), reflecting that the vast majority of children who selected only relevant questions also correctly identified the candidate that met all requirements to be the cause of the event described in the story (Table 2).

Discussion

This study investigates whether young learners can use their understanding of the connections between causes and effects to guide question-asking during information search. Past work has shown that 5-year-olds are adept at selecting more over less effective questions to distinguish between competing hypotheses (Ruggeri et al., 2017). However, this work has typcially presented explicit information about the probability or prevalence of the information targeted by each question. Here, we instread offered a choice between questions that differed in their conceptual relevance. We found that children still strongly prefer questions that are more likely to provide relevant information for determining an unknown cause. Both 5- and 7-year-olds were able to draw on their prior real-world knowledge to identify questions that targeted relevant information (e.g., an ability to climb is necessary to cause a mess on a high surface, while an ability to make noise has no clear causal relationship).

The current task also goes beyond previous research on question evaluation by conceptual relevance (e.g., Jirout & Klahr, 2020) in that both questions could (in principle) be relevant for gathering information about the identity of the pet. Unlike in Jirout and Klahr (2020), the difference in relevance in the current study must be *inferred* from the connection between an element of the story and a query about the agent. Children also had considerable uncertainty about potential information value of the questions offered. There is

Figure 2. Percentage of relevant questions selected in the *Choice* phase of the task.



Note: Error bars represent standard error.

no guarantee that the answers to any of the questions will (or will not) be informative (e.g., you might learn that all three pets can climb). This is particularly true of the Preference trial: learning what a pet likes to eat may, for example, be unhelpful in determining whether it prefers chips or cookies (e.g., 'crunchy things' applies to both, 'sour things' doesn't apply to either). Conversely, learning what the pet likes to play with might help by ruling out *both* of the toys left behind. Since children cannot be certain about the expected information gain of either question, they must make an inference about the likelihood that each will generate useful information. This presents a more realistic barrier to asking good questions than those probed in prior work.

Children's success on the current task is striking, but it is also consistent with recent research showing that young children are adept at selecting informative inquiries during self-directed learning (e.g., Lapidow & Walker, 2020; Lapidow et al., 2022; Wang et al., 2021). For example, Wang and colleagues (2021) found that preschoolers are more likely to choose to reveal uncertain outcomes for events from domains in which they still hold competing theories rather than mature knowledge. These studies suggest that young learners are selective in the information they ask for and have sophisticated expectations about which queries are likely to support learning and inference.

In addition to our main question of interest, we also examined whether children could use responses to the questions to determine the correct cause. When presented with a description of both relevant and irrelevant features for the three candidate causes, a significant majority of children selected the correct one. While not the focus of the current research, children's successful identification of which pet was capable of causing the even from these details is notable. First, as intended, it confirms that children had the minimum

background knowledge needed to recognize the causal connections in the task (e.g., knowing that climbing provides access to high surfaces is a prerequisite for determining that it is more useful to ask climbing about than to ask about making noise). Furthermore, even if children had this causal knowledge, combining and evaluating evidence to arrive at the correct conclusion is often challenging for young learners (Zimmerman, 2007). Children also succeeded despite the high volume of verbal and visual information and the presence of distractors (irrelevant information in each description and the almost-correct Distractor cause). This success, and the significant relationship between the *Choice* and *Guessing* phases (children who chose more relevant questions were also more likely to identify the correct cause), suggest that children drew on their causal understanding both to guide their queries and to draw inferences from the information generated in response.

We believe it is unlikely that children's strong performance is due to low-level associations: Fridges are not, for instance, typically defined as surfaces for climbing, nor is there any perceptual match between the fridge and the image used to represent climbing (stairs) in the *Choice* phase. The exception to this rule is in the *Appearance* trial, which intentionally included a low-level match to check participants' understanding of the task design. To our surprise, however, children's performance on the *Appearance* trial did not differ from the other two trial types. Further, if children had relied on non-causal associations between story details and questions, we would expect chance performance on the *Preference* trial, since the description included both playing-related toys and eating-related treats.

Table 2: Comparison of performance on the *Choice* and *Guessing* phases of the task.

Number of relevant questions chosen in *Choice*	Candidate chosen in *Guessing*		
	Correct	*Incorrect*	*Distractor*
None Relevant	1	0	0
One Relevant	2	0	1
Two Relevant	7	3	1
All Relevant	29	0	2

It is also unlikely that the current results are due to the specific content of the stimuli; however, if our hypothesis is correct, learners' performance should be consistent even if different events, causes, and questions are used. In ongoing work, we are therefore investigating children's performance on a different version of the current task. Although the structure of the task remains the same, children are presented with a different scenario in which an event has an unknown cause. If it is the case that the current results reflect a general ability to draw on causal knowledge to guide information search, then we should expect children's performance to be the same across distinct versions of the task.

Having determined children's ability to use conceptual knowledge in question evaluation, the natural next step is to ask whether this is also true of their question *generation*. Past work has shown that young children's ability to generate relevant questions in laboratory tasks lags behind their actual understanding of relevance for a variety of developmental reasons (see Jones et al., 2020). On the other hand, recent work suggests that children's attention to the abstract and structural features of problems provides a critical constraint on the otherwise intractable problem of generating relevant solutions (Chu & Schulz, 2020; Dechter et al., 2013; Schulz, 2012; Ullman et al., 2012). Our findings that children can select questions based on their causal relevance suggests that even young learners may be representing the abstract characteristics of events (e.g., a cleanly severed tree trunk or a mess on a high surface) as criteria for its cause (e.g., an action by a human agent, a creature that can climb) in a way that constrains the near infinite space of possibilities to those ideas that are worth considering.

Taken together, the current findings extend previous research into children's question-asking (e.g., Ruggeri et al., 2017; Jirout & Klahr, 2020) by demonstrating learners' ability to call on real world-causal knowledge to evaluate relevance. These results are consistent with prior empirical and theoretical work (e.g., Lapidow & Walker, 2022; Magid et al., 2014) suggesting the importance of causal intuitions in guiding learners' inferences in childhood. We find that young children employ expected connections between causes and effects as indications of what kind of information is most likely to be relevant, and are adept at selecting questions that target this information during search. These results offer novel insight into the strategies underlying early question-asking, as well as and children's developing abilities to choose from the vast space of investigations and queries available during learning.

References

Callanan, M. A., & Oakes, L. M. (1992). Preschoolers' questions and parents' explanations: Causal thinking in everyday activity. *Cognitive Development*, *7*(2), 213–233. https://doi.org/10.1016/0885-2014(92)90012-G

Chu, J., Gauthier, J., Levy, R., Tenenbaum, J., & Schulz, L. (2019). Query-guided visual search. *Proceedings of the 41st Annual Conference of the Cognitive Science Society.*

Chu, J., & Schulz, L. E. (2020). Play, Curiosity, and Cognition. *Annual Review of Developmental Psychology*. https://doi.org/10.1146/annurev-devpsych-070120-014806

Davidson, D. (1991). Developmental differences in children's search of predecisional information. *Journal of Experimental Child Psychology*, *52*(2), 239–255. https://doi.org/10.1016/0022-0965(91)90061-v

Dechter, E., Malmaud, J., Adams, R. P., & Tenenbaum, J. B. (2013). Bootstrap learning via modular concept discovery. *Twenty-third International Joint Conference*

on Artificial Intelligence, 1302–1309. https://www.ijcai.org/Proceedings/13/Papers/196.pdf

Frazier, B. N., Gelman, S. A., & Wellman, H. M. (2009). Preschoolers' search for explanatory information within adult-child conversations. *Child Development*, *80*(6), 1592–1611. https://doi.org/10.1111/j.1467-8624.2009.01356.x

Jirout, J., & Klahr, D. (2020). Questions–and some answers–about young children's questions. *Journal of Cognition and Development*, *21*(5), 729-753. https://doi.org/10.1080/15248372.2020.1832492

Jones, A., Swaboda, N., and Ruggeri, A. (2020). "Developmental changes in question asking," in The Questioning Child: Insights from Psychology & Education, eds L. Butler, S. Ronfard, and K. Coriveau (Cambridge: Cambridge University Press).

Lapidow, E., & Walker, C. M. (2020). Informative experimentation in intuitive science: Children select and learn from their own causal interventions. *Cognition*, *201*, 104315. https://doi.org/10.1016/j.cognition.2020.104315

Lapidow, E., Killeen, I., & Walker, C. M. (2022). Learning to recognize uncertainty vs. recognizing uncertainty to learn: Confidence judgments and exploration decisions in preschoolers. *Developmental Science*, *25*(2), e13178. https://doi.org/10.1111/desc.13178

Lapidow, E., & Walker, C. M. (2022). Rethinking the "gap": Self-directed learning in cognitive development and scientific reasoning. *Wiley Interdisciplinary Reviews: Cognitive Science*, *13(2),* e1580.

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*(1), 63–76. https://doi.org/10.1016/j.jecp.2012.07.002

Magid, R. W., Sheskin, M., & Schulz, L. E. (2015). Imagination and the generation of new ideas. *Cognitive Development*, *34*, 99–110. https://doi.org/10.1016/j.cogdev.2014.12.008

O'Neill, B. (2021). Optimal guessing in 'Guess Who'. *Plos one*, *16*(3), e0247361. https://doi.org/10.1371/journal.pone.0247361

Ronfard, S., Zambrana, I. M., Hermansen, T. K., & Kelemen, D. (2018). Question-asking in childhood: A review of the literature and a framework for understanding its development. *Developmental Review*, *49*, 101–120. https://doi.org/10.1016/j.dr.2018.05.002

Ruggeri, A., & Lombrozo, T. (2015). Children adapt their questions to achieve efficient search. *Cognition*, *143*, 203-216. https://doi.org/10.1016/j.cognition.2015.07.004

Ruggeri, A., Sim, Z. L., & Xu, F. (2017). "Why is Toma late to school again?" Preschoolers identify the most informative questions. *Developmental Psychology*, *53*(9), 1620–1632. https://doi.org/10.1037/dev0000340

Schulz, L. E. (2012). Finding New Facts; Thinking New Thoughts. *Advances in Child Development and Behavior*. https://doi.org/10.1016/B978-0-12-397919-3.00010-1

Tsividis, P., Tenenbaum, J. B., & Schulz, L. (2015). Constraints on hypothesis selection in causal learning. *Annual Meeting of the Cognitive Science Society (CogSci)*, 2434–2439. https://cogsci.mindmodeling.org/2015/papers/0418/paper0418.pdf

Ullman, T. D., Goodman, N. D., & Tenenbaum, J. B. (2012). Theory learning as stochastic search in the language of thought. *Cognitive Development*. https://doi.org/10.1016/j.cogdev.2012.07.005

Van der Meij, H. (1990). Question asking: To know that you do not know is not enough. *Journal of Educational Psychology*, 82(3), 505–512. https://doi.org/10.1037/0022-0663.82.3.505

Wang, J., Yang, Y., Macias, C., & Bonawitz, E. (2021). Children with more uncertainty in their intuitive theories seek domain-relevant information. *Psychological Science*, *32*(7), 1147–1156. https://doi.org/10.1177/0956797621994230

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, *27*(2), 172–223. https://doi.org/10.1016/j.dr.2006.12.001