

Children search for information as efficiently as adults, but seek additional confirmatory evidence

Azzurra Ruggeri (a.ruggeri@berkeley.edu)

Department of Psychology, University of California, Berkeley, USA

Max Planck Institute for Human Development, Berlin, Germany

Tania Lombrozo (lombrozo@berkeley.edu), Thomas L. Griffiths (tom_griffiths@berkeley.edu),

Fei Xu (fei_xu@berkeley.edu)

Department of Psychology, University of California, Berkeley, USA

Abstract

Like scientists, children and adults learn by asking questions and making interventions. How does this ability develop? We investigate how children (7- and 10-year-olds) and adults search for information to learn which kinds of objects share a novel causal property. In particular, we consider whether children ask questions and select interventions that are as informative as those of adults, and whether they recognize when to stop searching for information to provide a solution. We find an anticipated developmental improvement in information search efficiency. We also present a formal analysis that allows us to identify the basis for children's inefficiency. In our 20-questions-style task, children initially ask questions and make interventions no less efficiently than adults do, but continue to search for information past the point at which they have narrowed their hypothesis space to a single option. In other words, the performance change from age seven to adulthood is due largely to a change in implementing a "stopping rule"; when considering only the minimum number of queries participants would have needed to identify the correct hypothesis, age differences disappear.

Keywords: information search, active learning, 20-questions game, cognitive development.

Introduction

How should one seek evidence to test a hypothesis? This question has been discussed extensively within the philosophy of science (e.g., Crupi, 2014), and also describes a basic challenge faced by learners of any age (e.g., Markant & Gureckis, 2012; McKenzie & Mikkelsen, 2000; Oaksford & Chater, 1994): deciding which piece of evidence is most valuable to obtain, be it by questioning knowledgeable informants or directly intervening on the world. Here we explore how children (7- and 10-year-olds) and adults seek information in the context of a hierarchical causal inference task, with the aim of identifying the nature and source of variation in the efficiency of search across development.

Causal inference often requires categorizing objects and determining the level at which a given causal property applies. For example, most exemplars of the basic-level category *lamp* produce light, but not all pieces of furniture, the superordinate-level category, do. Using a 20-question-style task, we consider whether children are able to ask questions and select interventions that are as informative as those of adults, and whether they are able to recognize when to stop searching for information to provide a solution.

Previous research investigating children's information search has used variants of the "20-questions game," where the task is to identify an unknown target object by asking as few yes-or-no questions as possible, either generating the questions from scratch (e.g., Chouniard, 2007; Legare et al., 2013; Mosher & Hornsby, 1966; Ruggeri & Lombrozo, 2015) or selecting them from a list of provided alternatives (Nelson et al., 2013). These studies compared different age groups and found that the ability to ask effective questions undergoes a large developmental change from age 4 to 11.

In the current paper, we go beyond previous work in three ways. First, we adopt a quantitative approach, formalizing the efficiency of information search across development from age 7 to adulthood. This formal approach allows us to disentangle two components, or building blocks (Gigerenzer et al., 1999), required for performance: (1) an information search component, which involves the ability to select the most efficient information search path (e.g., to ask the most informative question at each time point), and (2) a stopping rule, which establishes when enough information has been collected. Our analysis aims at identifying which of these components accounts for developmental differences in information search. One intriguing possibility is that children are just as efficient as adults in their information search, but differ with respect to their stopping rule: they continue to seek information even when it is no longer needed to constrain their search. Beyond developmental comparisons, our formal approach allows us to measure the efficiency of children's information search strategies in absolute terms, by comparing their strategies to optimal or chance models (see Nelson et al., 2013).

Second, we compare two ways in which information can be sought: by asking yes-or-no questions (*question asking*) or by selecting single objects to test sequentially (*intervention*). These two paradigms present different challenges for the learner. The question-asking paradigm involves an explicit, verbal component, but also allows children to target entire categories directly. For instance, a child in the question-asking condition can ask whether "all lamps" have a given property, but a child in the intervention condition would have to test this hypothesis by selecting individual lamps and non-lamps one at a time. These conditions potentially correspond to how information search might unfold when learning from a knowledgeable informant versus directly from the world.

Third, we consider the role of *hierarchical structure* in information search. In our 20-questions-style task the objects are hierarchically organized and the possible solutions correspond to the levels of this structure. Using a hierarchical structure allows us to investigate how children search for information in a more complex and realistic environment, as opposed to the traditional scenario where one has to search for a single, arbitrarily-chosen object. This structure also allows us to consider whether children treat more abstract hypotheses preferentially, for instance by using them to initiate their search, and if so, whether they are only able to do so when asking questions (which can target higher levels in the hierarchy directly) versus making interventions. Higher-order hypotheses might be especially important not only in guiding efficient search, but in constraining induction more generally (Goodman, 1955).

Table 1. Materials and scenarios used in Study 1 and 2.

Scenario	Superordinate category	Basic-level category	Subordinate category	
Planet Apres	Animals	Fish	Goldfish Clownfish	
		Birds	Parrots Owls	
		Trees	Apple trees Pine trees	
		Flowers	Tulips Daisies	
	Clothes	Shirts	Long sleeves Short sleeves	
		Shoes	Flip-flops Boots	
		Furniture	Tables	Dining tables Desk
			Chairs	Rocking chair High chair
Vehicles	Cars	Vans Sportcars		
	Planes	Planes Helicopters		
	Fruit	Apples	Yellow apples Green apples	
		Berries	Raspberries Blueberries	
Machine				
Magic-box				

The Hierarchical 20-Questions Task. We designed a new task to investigate the role of hierarchical structure in information search, modeled after 20-questions tasks that have been used with children in prior research (see Mosher & Hornsby, 1966; Ruggeri & Lombrozo, 2015). In each of three trials, participants were presented with 16 objects on an iPad screen (in a random arrangement) and had to find out which set of objects shared a novel causal property. For example, they had to find out what kind of objects would turn on a machine. In Study 1 (question asking), they did so by asking yes-or-no questions, whereas in Study 2 (intervention), they selected and received feedback about individual objects by touching them on the screen. The 16 objects were organized hierarchically, with three levels (see Table 1): there were two sets of eight objects (*superordinate*

level), each containing two sets of four objects (*basic level*), each containing two sets of two objects (*subordinate level*). We manipulated the category level of the objects constituting the solution across the three trials, which were presented to the participants in random order. Each trial was randomly assigned to one of three different scenarios. After each question asked or object touched, participants received feedback from the experimenter with a response of “yes,” “no,” or (in Study 1) “some.” They were prompted to put a red (“no” feedback) or green (“yes” feedback) frame around the object/s to which their question referred, thus reducing memory demands. After receiving feedback, participants could choose whether to ask one more question (or select one more object) or to guess the solution. Participants could ask as many questions (or select as many objects) and guess the solution as many times as they wanted to, but were told to find the solution with as few questions (or selecting as few objects) as possible. At the end of the three trials constituting the experimental session, participants performed a sorting task to determine whether they appreciated the hierarchical structure and were able to verbally label categories at each level.

The Bayesian Framework. Our models and analyses are based on a Bayesian framework for concept learning and generalization (Tenenbaum & Griffiths, 2001). The learner’s hypothesis space is the set of hypotheses about which category of objects has the target causal property (e.g., turning on the machine). In our case, the hypothesis space consists of 14 alternative hypotheses, corresponding to all the object categories at any hierarchical level. We do not consider single-object hypotheses (e.g., only the yellow desk) because participants are explicitly told that the causal property applies to more than one object. Moreover, we do not consider disjunctive hypotheses, i.e., the combinations of objects across categories, such as “a boot or a desk can turn on the machine.” Such hypotheses were never provided by participants as possible solutions. Because participants were told that all categories, at any level, were equally likely to be true, we assume that participants initially expected all hypotheses to be equally likely, regardless of their level in the hierarchy.

To update their beliefs, we assume that after each observation x , participants evaluate all candidate hypotheses (i.e., those that are compatible with all observations X collected until that point) according to Bayes’ rule: computing their *posterior probability* $p(h|X)$, which is proportional to the product of their *prior probabilities* $p(h)$ and *likelihoods* $p(X|h)$:

$$p(h|X) = \frac{p(X|h)p(h)}{\sum_{h'} p(X|h')p(h')}$$

The prior $p(h)$ represents participants’ expectations about how likely the candidate hypotheses are. The likelihood $p(X|h)$ represents how likely it is that X would be observed if h is true. Here we make the simplifying assumption that

for each observation x , $p(x|h)$ is 1 if the observation is compatible with h and 0 otherwise, and that observations are independent conditioned on h , so $p(X|h)$ is just the product of $p(x|h)$ for each observation x . The posterior $p(h|X)$ is thus a function of the observations X and prior knowledge about the likelihood of the candidate hypothesis considered.

Information gain. Within the Bayesian framework it is possible to compute how informative each search option (question or object) is in terms of the expected information gain (e.g., Oaksford & Chater, 1994).

At each step of the search process, the participant evaluates all the possible actions in terms of their *information gain*, computed by subtracting the *predicted posterior entropy* from the *prior entropy*:

$$I = H_{\text{prior}} - H_{\text{posterior}}$$

The entropy H embodies the uncertainty as to which, among the candidate hypotheses, is true. Its computation is based on the probabilities of each of the candidate hypotheses:

$$H_{\text{prior}} = - \sum_h p(h) \log_2 p(h)$$

The prior entropy H_{prior} defines the status of uncertainty preceding every action. The predictive posterior entropy $H_{\text{posterior}}$ refers to the predicted status of uncertainty after the action is chosen and the correspondent feedback is observed. The predictive posterior entropy is measured as the sum of the entropies corresponding to each possible future scenario weighted according to the probability of that scenario:

$$H_{\text{posterior}} = p(x_1|X)H(x_1) + \dots + p(x_n|X)H(x_n)$$

where x_i is a possible observation, $p(x_i|X)$ is the probability of that observation resulting from taking the candidate action given all the information from previous observations, and $H(x_i)$ is the entropy of the posterior distribution over hypotheses after observing x_i . More formally,

$$p(x_i|X) = \sum_h p(x_i|h)p(h|X)$$

$$H(x_i) = - \sum_h p(h|X, x_i) \log_2 p(h|X, x_i)$$

Study 1: Question-asking

Participants. Participants were 24 children in second grade (10 female, $M_{\text{age}} = 90.5$ months; $SD = 5.56$ months), and 23 children in fifth grade (8 female, $M_{\text{age}} = 119.4$ months; $SD = 12.7$ months), recruited from a primary school and a local children’s museum, as well as 23 university students (15 female, $M_{\text{age}} = 21.1$ years; $SD = 2.6$ years).

Results

Results were analyzed by running repeated-measures ANOVAs with age group (3 levels: 7-year-olds, 10-year-olds, adults) as a between-subjects variable and trial number (3 levels: 1, 2, 3), solution condition (3 levels: subordinate-level, basic-level, superordinate-level) or scenario (3 levels: Magic box, Machine, Planet) as within-subjects variables. All main effects and interactions were tested; we report only significant effects.

Descriptive analysis. We analyzed the questions asked prior to reaching the correct solution (which we refer to as the “complete path”). We found a main effect of age group on the *number of questions asked* prior to giving the solution, $F(2, 67) = 5.29, p = .007, \eta^2 = .136$. A Bonferroni corrected multiple comparisons analysis confirmed that 7-year-olds ($M_{7\text{-year-olds}} = 4.92, SE = .34$) asked more questions than adults ($M_{\text{adults}} = 3.36, SE = .35, p = .006$). We did not find any difference between the number of questions asked between 7- and 10-year-olds ($M_{10\text{-year-olds}} = 4.38, SE = .35, p = .807$), or between 10-year-olds and adults ($p = .126$). We did not find any within-subject effect of scenario or trial number, but we did find an effect of condition, $F(2, 67) = 20.02, p < .001, \eta^2 = .320$. A Bonferroni corrected multiple comparisons analysis showed that participants needed fewer questions in the superordinate condition ($M_{\text{superordinate}} = 3.37, SE = .26$) than in the basic condition ($M_{\text{basic}} = 4.08, SE = .24, p = .038$), and in the basic condition than in the subordinate condition ($M_{\text{subordinate}} = 5.21, SE = .31, p < .001$).

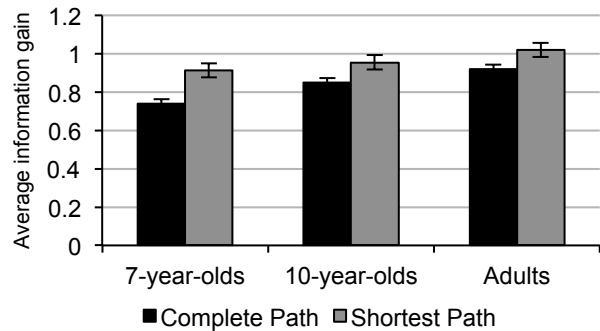


Figure 1. Study 1: Average information gain of the questions asked before giving the solution (complete path) or before having narrowed down the hypothesis space to one hypothesis (shortest path). Error bars represent one SEM in each direction.

The analysis also showed a main effect of age group on the *average information gain of the questions asked* prior to giving the solution, $F(2, 67) = 5.27, p = .007, \eta^2 = .136$ (see Figure 1). A Bonferroni corrected multiple comparisons analysis confirmed that the average information gain of the questions asked by 7-year-olds ($M_{7\text{-year-olds}} = .74, SE = .04$) was lower than the average information gain of the questions asked by adults ($M_{\text{adults}} = .92, SE = .04, p = .006$). There were no differences between 7- and 10-year-olds

($M_{10\text{-year-olds}} = .85$, $SE = .04$, $p = .185$) nor between 10-year-olds and adults ($p = .123$).

Level of the first question asked. Across the three trials, adults asked a larger number of first questions at the superordinate level ($M_{\text{adults}} = 1.87$, $SE = .16$) than did older children ($M_{10\text{-year-olds}} = 1.13$, $SE = .20$), $t(162) = -3.33$, $p = .001$, who in turn asked a larger number of such questions than did younger children ($M_{7\text{-year-olds}} = .41$, $SE = .16$), $t(145) = -3.02$, $p = .003$. Symmetrically, adults asked fewer first questions at the subordinate level ($M_{\text{adults}} = .13$, $SE = .06$) than did older children ($M_{10\text{-year-olds}} = .29$, $SE = .10$), $t(162) = -3.14$, $p = .002$, who in turn asked fewer such questions than did younger children ($M_{7\text{-year-olds}} = .69$, $SE = .17$), $t(145) = -2.14$, $p = .034$. The number of initial questions at the basic level did not vary significantly across age groups ($ps > .05$): $M_{7\text{-year-olds}} = .79$, $SE = .18$, $M_{10\text{-year-olds}} = .92$, $SE = .17$, $M_{\text{adults}} = .55$, $SE = .12$.

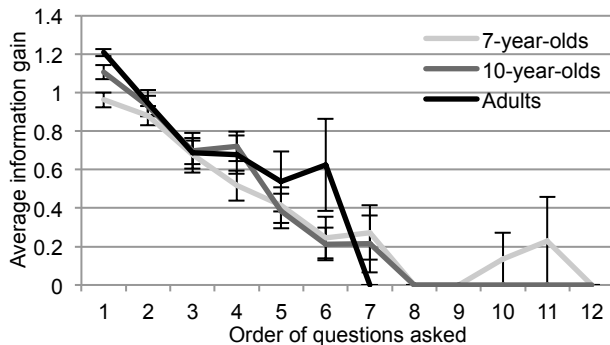


Figure 2. Study 1: Average information gain, displayed by order of questions asked and age group (minimum number of participants per data point displayed = 2). Error bars represent one SEM in each direction.

Analysis of the shortest path. To disentangle participants' information search from their stopping rules, we considered the number and efficiency of the questions asked prior to obtaining enough information to identify the solution, whether or not the participant went on to ask additional (uninformative) questions. In other words, we considered participants' information search had they stopped asking questions and stated the solution the moment a single hypothesis remained. We refer to this as the "shortest path." Surprisingly, we did not find a main effect of age group on the number of questions asked in the shortest path ($p = .122$). We did not find any within-subject effect of scenario or trial, but we did find an effect of the level of the solution, $F(2, 67) = 28.31$, $p < .001$, $\eta^2 = .300$. A Bonferroni corrected multiple comparisons analysis showed that participants needed fewer questions in the superordinate condition ($M_{\text{superordinate}} = 2.64$, $SE = .16$) than in the basic condition ($M_{\text{basic}} = 3.61$, $SE = .15$, $p < .001$) or in the subordinate condition ($M_{\text{subordinate}} = 4.09$, $SE = .19$, $p < .001$) prior to obtaining enough information to isolate a single hypothesis. There was no significant difference in the number of questions across the basic and subordinate

conditions. Most interestingly, the analysis did not show a main effect of age group on the average information gain of the questions asked in the shortest path ($p = .134$; see Figure 1). Thus, despite the developmental trend, the average information gain of the questions asked in the shortest path by 7-year-olds ($M_{7\text{-year-olds}} = .91$, $SE = .04$), 10-year-olds ($M_{10\text{-year-olds}} = .96$, $SE = .04$) and adults ($M_{\text{adults}} = 1.02$, $SE = .04$) did not differ, suggesting that children were just as efficient as adults in narrowing down the hypothesis space.

In sum, these analyses suggest that the developmental differences that we observed in overall efficiency were driven largely by children's tendency to ask questions beyond the point at which a single hypothesis remained (see Figure 2). Eighty-three percent of the 7-year-olds ($n = 20$), 70% of the 10-year-olds ($n = 16$), and only 52% of the adults ($n = 12$) asked, in at least one trial, more questions than strictly necessary to identify a single hypothesis. A repeated-measures ANOVA showed an age group effect on the number of questions asked beyond the point at which a single hypothesis remained, $F(2, 67) = 4.50$, $p = .015$, $\eta^2 = .118$. A Bonferroni corrected multiple comparisons analysis confirmed that the 7- and 10-year-olds asked on average more additional questions ($M_{7\text{-year-olds}} = 1.28$, $SE = .22$; $M_{10\text{-year-olds}} = .81$, $SE = .23$) than adults ($M_{\text{adults}} = .32$, $SE = .23$, $p = .011$).

Difference between an optimal model, a random model, and participants' information search. We compared participants' information search against an optimal model and a random model. The optimal model follows the best possible information search path – that is, it selects at each step the question that has the highest information gain, considering the current hypothesis space. The random model selects an option at random. This random selection is repeated ten times at each step, with replacement, and we consider the average information gain of the ten randomly selected options. A repeated-measures ANOVA showed that participants' average information gain ($M_{\text{participants}} = .83$, $SE = .02$) was higher than the information gain resulting from a random selection ($M_{\text{random}} = .50$, $SE = .01$), but lower than the one resulting from the optimal model ($M_{\text{optimal}} = 1.06$, $SE = .02$), $F(2, 134) = 697.97$, $p < .001$, $\eta^2 = .91$. The analysis revealed no main effect of age group nor interactions.

Study 2: Intervention

Participants. Participants were 22 children in second grade (7 female, $M_{\text{age}} = 90.0$ months; $SD = 6.2$ months) and 23 children in fifth grade (11 female, $M_{\text{age}} = 119.6$ months; $SD = 11.7$ months), recruited from a primary school and a local children's museum, as well as 22 university students (16 female, $M_{\text{age}} = 23.8$ years; $SD = 5.7$ years).

Results

Descriptive analysis. Did the efficiency of information search vary across age groups or solution types when children selected objects, as opposed to asking questions?

We found a main effect of age group, $F(2, 64) = 21.16, p < .001, \eta^2 = .398$, with fewer objects tested with increasing age. A Bonferroni corrected multiple comparisons analysis confirmed that 7-year-olds selected more objects ($M_{7\text{-year-olds}} = 7.79, SE = .44$) than 10-year-olds ($M_{10\text{-year-olds}} = 5.96, SE = .43, p = .013$), and 10-year-olds more than adults ($M_{\text{adults}} = 4.11, SE = .44, p = .012$). We did not find any within-subject effect of condition, scenario, or trial. A parallel analysis also revealed a main effect of age group on the average information gain of the objects selected prior to stating the solution, $F(2, 64) = 11.91, p < .001, \eta^2 = .274$ (see Figure 3). A Bonferroni corrected multiple comparisons analysis confirmed that the average information gain of the objects selected by 7-year-olds ($M_{7\text{-year-olds}} = .49, SE = .027$) was lower than the average information gain of the objects selected by 10-year-olds ($M_{10\text{-year-olds}} = .60, SE = .027, p = .022$) and adults ($M_{\text{adults}} = .68, SE = .027, p < .001$). However, the difference between 10-year-olds and adults was not significant ($p = .123$). We did not find any within-subject effect of condition, scenario, or trial.

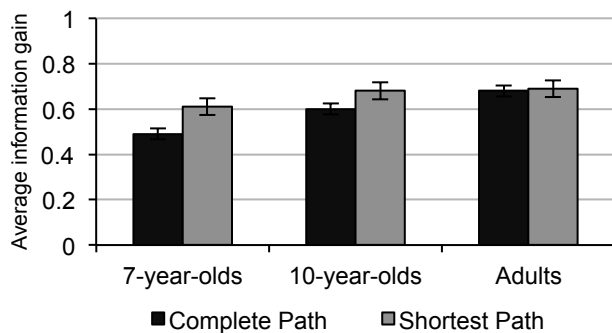


Figure 3. Study 2: Average information gain of the objects selected before giving the solution (complete path) or before having narrowed down the hypothesis space to one hypothesis (shortest path). Error bars represent one SEM in each direction.

Analysis of the shortest path. We analyzed participants' performance for the "shortest path": the number of objects selected, and their associated information gain, had they given the solution the moment they had enough information to isolate a single hypothesis. The analysis revealed a main effect of age group on the number of objects selected prior to narrowing down the hypothesis space to one hypothesis, $F(2, 64) = 9.03, p < .001, \eta^2 = .220$. A Bonferroni corrected multiple comparisons analysis confirmed that 7-year-olds selected more objects ($M_{7\text{-year-olds}} = 5.73, SE = .29$) than adults ($M_{\text{adults}} = 4.02, SE = .29, p < .001$). However, we did not find any differences between the number of objects selected by 7-year-olds and 10-year-olds ($M_{10\text{-year-olds}} = 4.77, SE = .28, p = .058$), or 10-year-olds and adults ($p = .192$). Note that these age group effects were weaker than those found when analyzing all objects selected prior to giving the solution. We did not find any within-subject effect of condition, scenario or trial.

Most interestingly, the analysis did not show a main effect of age group on the average information gain of the objects selected prior to narrowing the hypothesis-space down to one hypothesis ($p = .060$; see Figure 2), although there was a developmental trend from 7-year-olds ($M_{7\text{-year-olds}} = .61, SE = .02$) to 10-year-olds ($M_{10\text{-year-olds}} = .68, SE = .02$) to adults ($M_{\text{adults}} = .69, SE = .02$). Again, we did not find any within-subject effect of condition, scenario, or trial.

As in Study 1, this suggests that what changes over development is not (only) the ability to select an efficient information path, but the stopping rule (see Figure 4). Eighty-six percent of the 7-year-olds ($n = 19$), 87% of the 10-year-olds ($n = 20$), and only 48% of the adults ($n = 10$) selected, in at least one trial, more objects prior to giving the solution than they would have needed. A repeated-measures ANOVA showed an age group effect on the number of objects selected beyond the point at which a single hypothesis remained, $F(2, 64) = 10.61, p < .001, \eta^2 = .249$. A Bonferroni corrected multiple comparisons analysis confirmed that the 7- and 10-year-olds tested on average more additional objects ($M_{7\text{-year-olds}} = 2.23, SE = .33; M_{10\text{-year-olds}} = 1.17, SE = .32$) than the adults did ($M_{\text{adults}} = .09, SE = .33, p < .001$).

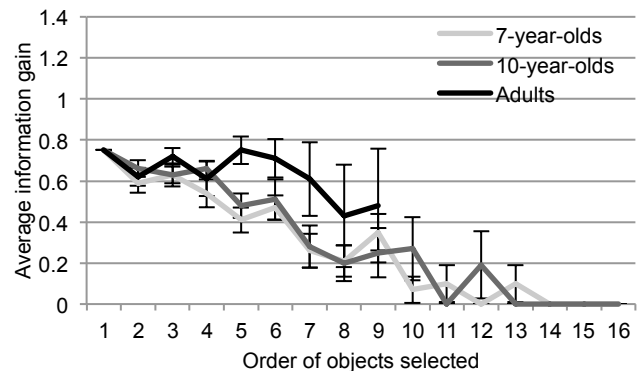


Figure 4. Study 2: Average information gain, displayed by selection number and age group (minimum number of participants per data point displayed = 2). Error bars represent one SEM in each direction.

Difference between an optimal model, a random model, and participants' information search. The analysis showed that participants' average information gain ($M_{\text{participants}} = .60, SE = .02$) was higher than the information gain resulting from a random selection ($M_{\text{random}} = .45, SE = .01$), but lower than that resulting from the optimal model ($M_{\text{optimal}} = .76, SE = .01$), $F(2, 128) = 429.49, p < .001, \eta^2 = .87$. The analysis revealed a main effect of age group, $F(2, 64) = 12.74, p < .001, \eta^2 = .29$, but no interactions.

Discussion

Across two studies involving different kinds of information search (asking questions versus testing objects), we investigated the efficiency of information search across development. Our task allowed us to address several related questions.

First, we adopted a quantitative approach to consider the role of hierarchical structure in two distinct forms of search: asking questions versus testing objects (interventions). We found that performance in the question-asking task was better than in the intervention task (for all age groups), as predicted by the optimal model. We also showed, for the first time, that the information search strategies of 7- and 10-year-olds are more efficient than random strategies, both in a question-asking and in an intervention paradigm. We were also able to analyze whether children and adults are able to exploit hierarchical structure when searching for information, by approaching the task top-down. We found that this was the case when asking questions, as reflected in a more efficient solution path when the solution involves a higher-level category; for the intervention task, however, the advantage for higher levels disappeared.

Second, our formal analysis allowed us to home in on the sources of developmental differences in information search. We found developmental trends in efficiency, with older participants taking fewer and more efficient steps in their search. These results replicate prior research (see Davidson, 1991a; 1991b; Mosher & Hornsby, 1966), which suggest that children are less efficient. This inefficiency is usually explained in terms of immature strategic abilities or inability to focus on the most relevant pieces of information. Instead, we find that children's inefficiency stems largely from a tendency to ask questions or test objects beyond the point at which only one hypothesis remains. Specifically, our analysis allows us to disentangle the role of children's information search from their stopping rule, suggesting that children's initial search is no less efficient than adults'. However, whereas adults stop searching once they obtain enough information to solve the task, children continue.

There are two plausible interpretations for these results, not mutually exclusive. First, children might entertain more hypotheses than those considered in our model's hypothesis space. For example, they might consider disjunctive hypotheses (e.g., a desk OR a high chair will produce the effect). This possibility deserves further study, but it is notable that children never spontaneously offered such hypotheses. Second, children's stopping rule itself might differ from that of adults. In particular, they may seek confirming evidence even when it's not strictly "informative," according to our analysis. This interpretation is supported by some children's comments accompanying the selection of additional objects (e.g., "I think I know, but let me try just one more question, to be sure"). Although these results are surprising in light of previous research showing that children of this age tend to be *overconfident* (e.g., Finn & Metcalfe, 2013; Salles et al., 2015), they are also consistent with research on children's decision making, which finds that younger children tend to be more exhaustive in their search than older children (Davidson, 1991a; 1991b).

Looking for confirming evidence is also a strategy that could make sense if there's uncertainty about the hypothesis space, the feedback one has received, or the stability of what

is being learned. As novice learners in a noisy world, children might do well to err on the side of obtaining extra feedback. Many questions remain open, but our task and analyses provide first steps in a more formal approach to understanding testing and confirmation throughout the lifespan.

Acknowledgments

We thank the Madera elementary School and the Lawrence Hall of Science, and Celina Vicuna, Minh-Thy Nguyen, Samika Kumar, Meg Bishop and Simone Riley for assistance in data collection and coding. This research was supported by a Marie Curie International Outgoing Fellowship within the 7th European Community Framework Programme to Azzurra Ruggeri.

References

- Chouinard, M. M. (2007). Children's Questions: A Mechanism for Cognitive Development. *Monographs of the Society for Research in Child Development*, 72(1), vii–ix.
- Crupi, V. (2014). Confirmation. In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition)
- 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.
- Finn, B., & Metcalfe, J. (2014). Overconfidence in children's multi-trial judgments of learning. *Learning and Instruction*, 32, 1–9.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Goodman, N. (1955). *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard UP.
- 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.
- Markant, D.B. and Gureckis, T.M. (2012) Does the utility of information influence sampling behavior? In N. Miyake, Peebles, D. & Cooper, R.P. (Eds.) *Proceedings of the 34th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- 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.
- Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101(4), 608-631.
- Ruggeri, A., & Lombrozo, T. (2015) *Children Adapt Their Question-Asking Strategies to Achieve Efficient Search*. Manuscript under review.
- Salles, A., Semelman, M., Sigman, M., & Calero, C. (2015). *The Metacognitive Abilities of Children and Adults*. Manuscript under review.
- Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and brain sciences*, 24(04), 629-640.