Journal of Experimental Child Psychology 115 (2013) 517–535

Contents lists available at SciVerse ScienceDirect

Journal of Experimental Child Psychology

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

Make your own kinds of cues: When children make more accurate inferences than adults

Azzurra Ruggeri *, Konstantinos V. Katsikopoulos

Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, 14195 Berlin, Germany

ARTICLE INFO

Article history:
Received 21 May 2012
Revised 14 November 2012
Available online 25 March 2013

Keywords:
Cues
Inferences
Information search
Development
Decision-making
Cue-generation

ABSTRACT

In everyday decision making, we do not always have the luxury of using certain knowledge but often need to rely on cues, that is, pieces of information that can aid reasoning. We ask whether and under what circumstances children can focus on informative cues and make accurate inferences in real-world problems. We tested second-, third-, and fifth-grade children and young adults on two problems: which of two real cars is more expensive and which of two real cities has more inhabitants. We manipulated whether cues were given to the participants or the participants needed to generate their own cues. The main result was that when generating their own cues, younger children matched older children and young adults in accuracy or even outperformed them. On the other hand, when cues were given, children were less accurate than young adults. A possible explanation for this result is that children, on their own, tend to generate "perceptual" cues (e.g., "Which car is longer?") that are informative in the problems we studied. However, children are not able to recognize the most informative cues in a set that is given to them because they are not familiar with all cues (e.g., non-perceptual cues such as which car has more horsepower).

Introduction

Which one of Fiat's two cars, Doblò or Panda, is the more expensive model? If you do not know the answer already, you will need to rely on cues. A cue is a piece of information such as a question, feature, or concept that is useful for making decisions. Children ask questions such as "Which of the two

* Corresponding author.
E-mail address: ruggeri@mpib-berlin.mpg.de (A. Ruggeri).

0022-0965/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jecp.2012.11.007

cars is longer?" whereas adults may ask questions such as "Which of the two cars has greater horsepower?"

In this article, we address the following question: Can children focus on informative cues—cues that often lead to the correct answer—when making inferences in the real world? We investigated how the way cues are obtained by children (younger and older) and young adults affects (a) the informativeness of the cues and (b) the accuracy of the inferences. In particular, we manipulated whether participants needed to generate their own cues from memory (Experiment 1) or were provided with cues that were displayed on an information board (Experiment 2). Note that in all previous studies, information search was investigated by providing cues on an information board (for an exception, see Katz, Bereby-Meyer, Assor, & Danziger, 2010). In addition, virtually none of the developmental studies used real problems in which an objectively correct inference exists.

Previous research on children's cue-based decision making

The previous studies investigating children's cue-based decision making have mainly examined whether children adapt their search strategies in response to different task characteristics. Davidson (1991a) studied how second graders (7- and 8-year-olds), fifth graders (10- and 11-year-olds), and eighth graders (13- and 14-year-olds) search for predecisional information (i.e., cues) when making choices. Her results suggest that second graders are exhaustive in their information search, even though they are less systematic than older children, who were more likely than younger children to base their decision on the same two or three cues (e.g., the color and cost of a bicycle).

In a subsequent study, Davidson (1991b) investigated whether older children search for information more systematically than younger children because they focus on the dimensions they find relevant or important for making the decision. In particular, Davidson manipulated the scenarios, as well as the ways in which the cues were displayed, so that some cues were presented and displayed as more "relevant" or "important" than others. Results show that older children were more likely than younger children to examine the cues defined as relevant in the cover story. In contrast, younger children tended to attend to irrelevant information more often than older children, a finding that is in line with earlier research (e.g., Hagen & Hale, 1973; Pick & Frankel, 1973). Davidson (1996) also found that by highlighting the relevant cues on the information board, second and fifth graders searched more on relevant dimensions and selected the best alternative. However, most of the younger children still searched information exhaustively and without any systematic pattern.

These results are in line with other findings that in the absence of an external motivational incentive, younger children gather more information than older children (Gregan-Paxton & Roedder John, 1995, 1997; Howse, Best, & Stone, 2003). However, younger children search for less information when cues are costly than when they are not (Gregan-Paxton & Roedder John, 1997; Howse et al., 2003).

In a more recent study, Mata, von Helversen, and Rieskamp (2011) investigated the ability of 9- to 12-year-olds to focus their decisions on the most *informative* cues as opposed to the most relevant cues. In particular, Mata and colleagues asked the participants to infer which of two cars would win a race after providing the values of the cars' cues (e.g., horsepower, number of cylinders) displayed on a computerized information board. The informativeness of a cue was defined as the probability of the cue leading to a correct inference in that task. The informativeness of the given cues was learned in a previous training session. Mata and colleagues concluded that less than 30% of the children spontaneously used strategies that employed one cue or just a few cues. The authors suggested that children do not use such frugal strategies because they cannot easily focus on the most informative cues.

All of the studies presented above used an information board procedure or a computerized version of it. An information board presents a matrix of information; the rows represent the choice or decision alternatives (e.g., Car 1 and Car 2), and the columns contain information related to a cue (e.g., 100 horsepower). Participants choose an alternative after having explored the information board, examining as much information as they considered necessary to make the decision in any order they wanted.

We argue that the information board procedure might not be a suitable methodology to test children's ability to focus on the most relevant or informative cues. First, a set of cues to choose among is a necessarily constrained selection that might not include all of the possible information a child could come up with. Second, it requires children to assess and compare the informativeness of exogenously given cues with which they might not be familiar or the meaning of which they might not know (e.g., horsepower). Third, what is relevant or informative is exogenously defined.

An exception to this is a study by Bereby-Meyer, Assor, and Katz (2004). These authors asked children (8- and 9-year-olds and 12- and 13-year-olds) to choose among consumer products, such as bicycles and watches, after being given descriptions of the products' cue values. In this study, each child was asked to personally rate the importance of the cues. The authors found that the few cues that children based their decisions on were not randomly chosen. In fact, children tended to examine the cues that, according *to them*, were the most informative, choosing the alternative that had the highest value on that cue and ignoring all other cues and their values. The Bereby-Meyer and colleagues concentrated solely on choice behavior. A drawback of choice tasks is that the correct answer cannot always be objectively defined.

The current study

We sought to compare two different experimental procedures to investigate whether a different methodology could help children to focus on the most informative cues and, therefore, to be able to search for information more efficiently. In two experiments, second-, third- and fifth-grade children, as well as young adults, were asked to make inferences. An inference (e.g., which of two real cars is more expensive, which of two real cities has more inhabitants) is a special kind of decision for which there is an objective answer (i.e., one car *is* more expensive, one city *is* larger). By using an inference task, we were able to categorize the decisions as right or wrong. In Experiment 1, we introduced a methodology that we believe is more suitable to test children's ability to select informative cues in an inference task. We did not provide participants with cues displayed in an information board matrix but instead prompted them to *generate* their own cues by letting them ask anything they wanted about the objects (e.g., cities) they needed to make inferences about. Participants would be familiar with the cues they themselves had generated. Moreover, they received no instructions or training on which cues should be considered more relevant or informative. However, the informativeness of the generated cues was objectively assessed as the probability of that cue leading to a correct inference in the particular real-world task. In Experiment 2, we replicated the task but with a traditional information board procedure.

The process of searching for information by asking questions from scratch has been thoroughly studied (Graesser & McMahen, 1993; Graesser & Olde, 2003), also as a mechanism for cognitive development and knowledge acquisition (Chouinard, 2007; Vosniadou, 1994). This methodology has also been used to investigate search for information in categorization tasks (see Mosher & Hornsby, 1966; Ruggeri & Feufel, 2013), in causal reasoning tasks (see Mosher & Hornsby, 1966), and in preferential choice tasks (Katz et al., 2010). However, to the best of our knowledge, it has never been used in researching cue-based inferences.

We hypothesized that when generating their own cues, participants would use fewer cues than when provided with cues, both because generating cues is more effortful than selecting them (Katz et al., 2010) and because participants would be more confident that the cues they had generated would be the most informative ones. We also hypothesized that when generating their own cues, each age group would come up with different cues because, being at different developmental phases, they would have access to different cues. Moreover, we explored the hypothesis that, when generating their own cues, children might sometimes come up with cues that were objectively as informative as the cues of the young adults and, thus, would match the accuracy of young adults.

Experiment 1: Free generation of cues

Method

Participants

The experiment involved 66 participants: 17 children in second and third grades (9 female and 8 male, M_{age} = 7.8 years, SD = 0.39) and 27 children in fifth grade (14 female and 13 male,

M_{age} = 9.9 years, *SD* = 0.61) of a primary school[1] in Livorno, Italy, and 22 young adults (12 female and 10 male, M_{age} = 17.9 years, *SD* = 0.75) recruited from a high school[2] in Livorno. The participants all were born in Italy and belonged to various social classes. The results of another 11 participants were excluded because of experimental errors or equipment malfunction.

Design and procedure

The experiment was run on a computer and each session consisted of 8 trials. On each trial, participants were presented with two objects and were asked to infer which object had the higher value on a criterion specified in the instructions for that trial. There were two possible criteria: the population of a city and the price of a car. For example, in one trial the participants needed to infer whether Milan or Venice was more populous, and in another trial the participants needed to infer whether a Doblò or a Panda was the more expensive car.

On each trial, participants received a different pair of objects randomly drawn from two environments: the 60 currently most populous Italian cities and 52 cars currently produced by two Italian car manufacturers (Fiat and Alfa Romeo). Our database included for each object its value on the criterion (city population or car price); there were 28 cue values for cities and 20 for cars. The cues were generated before the experiment in a survey with 15 children (8- and 9-year-olds) and 10 adults who did not take part in the experiment. The objects are listed in Table 1, and the cues are listed in Table 2.

In Table 2, we also list cue *success* (Newell, Rakow, Weston, & Shanks, 2004), which we used in analyzing our results. In both experiments, the cues generated (or *selected* in Experiment 2) might be conflicting, meaning that some cues might point to the correct choice, whereas other cues might suggest an incorrect choice. The success of a cue in a task is a measure of the probability, in the particular task, that the cue will lead to a correct inference. To compute the success of a cue, it is assumed that if the cue has a higher value on one of the two objects, this object will be picked, and if the cue has the same value on the two objects, an object will be picked randomly.

Besides the inference criterion (car price or city population), our second independent variable was whether objects were presented with names (i.e., the name of a city, the model of a car) or with generic labels instead of names (i.e., City 1 and City 2, Car 1 and Car 2). Finally, there were 2 trials for each of the four combinations of the two independent variables. These 2 trials constituted a block, and the order of the four blocks was counterbalanced across participants.

The two objects (cities or cars) were displayed on a computer screen. The participants were prompted to generate cues freely by asking questions about the objects. For example, they might ask whether the cities had a university or what the cars' maximum speed was. The only restriction was that cues with subjective values were not allowed; questions such as "Are these cars cool?" and "Do you think I would like to live in one of these cities?" were not answered or considered in the analysis. When an objective cue (e.g., presence of a university in a city) was generated, the experimenter provided the values of the two objects on that cue by using a database stored in the computer. The values of the cues that participants generated were displayed on the screen until the end of the trial. The participants were allowed to ask for as many cues as they wanted, even no cues. If participants generated cues not available in the database, they were told that these cue values were not available. Some cues that we did not expect—and hence did not have in our database—were very original and smart such as number of McDonald's in a city (generated by a 10-year-old) and number of television advertisements for a car (generated by a 17-year-old). Table 3 lists cues that children generated themselves but that were not included in the database; they are not considered in the analysis of frugality because frugality is a measure of the usable cues for making an inference, and in these cases participants did not get any answers they could use to make the inferences.

At the beginning of the experiment, the participants received 60 tokens. For each correct inference, they gained 5 tokens, whereas an incorrect inference left their number of tokens unchanged. In addition, participants needed to pay 1 token for each cue they asked for. The number of tokens was continually updated and appeared in the corner of the screen. The participants were told that, for each age group, the three participants with the highest number of tokens at the end of the experiment would be

[1] Fondazione Sacro Cuore, Ardenza, Livorno, Italy.
[2] Liceo Scientifico "F. Cecioni," Livorno, Italy.

Table 1
Objects and their criterion values (city populations and car prices in euros).

Object	Criterion value	Object	Criterion value	Object	Criterion value	Object	Criterion value
City population task							
Roma	2,724,347	Taranto	194,021	Sassari	130,306	Andria	99,249
Milano	1,295,705	Brescia	190,844	Siracusa	124,083	Udine	99,071
Napoli	963,661	Reggio Calabria	185,621	Pescara	123,022	Arezzo	98,788
Torino	908,825	Prato	185,091	Monza	121,280	Cesena	95,525
Palermo	659,433	Parma	182,389	Latina	117,149	La Spezia	95,372
Genova	611,171	Modena	181,807	Bergamo	116,677	Lecce	94,775
Bologna	374,944	Reggio Emilia	165,503	Forlì	116,208	Pesaro	94,197
Firenze	365,659	Perugia	165,207	Vicenza	115,012	Barletta	93,869
Bari	320,677	Livorno	161,095	Trento	114,236	Alessandria	93,676
Catania	296,469	Cagliari	157,297	Giuliano di Campania	113,811	Catanzaro	93,519
Venezia	270,098	Ravenna	155,997	Terni	112,021	Pistoia	89,982
Verona	265368	Foggia	153239	Novara	103602	Brindisi	89,691
Messina	243,381	Salerno	140,489	Ancona	102,047	Torre del Greco	87,735
Padova	211,936	Rimini	140,137	Bolzano	101,919	Pisa	87,398
Trieste	205,341	Ferrara	134,464	Piacenza	101,778	Lucca	84,186
Car price task							
Spider	40,851	159 SW Progression	25,711	500 Lounge	19,401	Idea BlackStar	14,801
Ulysse Emotion	37,251	Multipla Emotion	25,151	Panda 4 × 4 Cross	18,701	Punto Evo Fun	14,751
GT Q. Verde	34,601	159 Progression	24,511	500 MJT Lounge	18,601	Grande Punto Actual	13,501
159 Distinctive Q-Tronic	34,151	Croma Active	24,101	Sedici 4 × 2 Dynamic	18,501	Panda 4 × 4	13,351
Croma Emotion	32,601	Multipla Active	23,151	Bravo Dynamic	18,101	QUBO Active	13,051
Spider TBi	31,951	Bravo Dualogic Dyn.	22,251	Doblò Active	18,101	Panda Emotion Eco	12,351
159 Eco Distinctive	31,651	Idea BlackMotion	22,051	Punto Evo Sport	17,901	Punto Evo Active	11,951
Ulysse Active	30,701	500 Rock	21,601	147 Moving	17,481	500 Pop	11,701
Brera TBi	29,951	Sedici Dynamic	21,501	500 by DIESEL	17,351	Grande Punto Actual	11,601
Croma Emotion	28,101	Multipla Dynamic	20,951	Idea BlackLabel GPL	17,151	Punto Classic Active	11,001
GT Progression	26,551	Bravo Dualogic Dyn.	20,501	Grande Punto Actual Natural P.	16,201	Punto Classic	10,301
Sedici Experience	26,501	Giulietta Turbo Progression	20,451	QUBO Dynamic	16,051	Panda Actual Eco	9,001
Giulietta Progression	25,851	Doblò Dynamic	20,051	Doblò 1.4 Actual	15,101	600	7,951

Note. City names are given in Italian.

Table 2
Experiment 1: Cues and their success in the two tasks.

Cue	Success	Cue	Success	Cue	Success	Cue	Success
City population task							
Number of families	.93	University	.76	Airports (overall)	.64	Seismic danger	.53
Number of buildings	.92	Museums	.70	Airports (international)	.64	Regional capital	.52
Primary schools	.87	Universities	.69	Hotels	.62	Area	.52
Preschools	.83	Density of population	.67	Average income	.61	Being a capital city	.52
Secondary schools	.833	Soccer teams	.66	Age index	.58	Climate zone	.52
High schools	.80	Airports (civilian)	.65	Stadiums	.57	Degree days	.51
Hits on Google	.77	Tourism (ranking)	.64	Altitude	.56	Airports (military)	.51
Car price task							
Horsepower	.88	Width	.77	Acceleration	.63	Fuel consumption (mixed)	.60
Mass	.85	Fuel tank capacity	.71	Brand	.63	Number of seats	.58
Capacity	.84	Coachwork	.70	Trunk capacity	.62	Fuel consumption (highway)	.58
Speed	.83	Gears	.68	Fuel consumption (city)	.62	Height	.54
Length	.78	Type of fuel	.67	Revolutions per minute	.60	Doors	.51

Table 3
Experiment 1: Generated cues that were not included in the database.

Task	Participants asking for this cue	Cue
City population	1 younger child	Number of monuments
	1 younger child, 1 young adult	Number of shopping malls
	1 older child	Number of McDonald's
	1 older child	Number of streets
	1 young adult	Existence of a dialect
	2 young adults	First letter
Car price	1 older child	Kind of rims
	2 younger children, 2 older children	Color
	2 younger children, 1 older child, 2 young adults	The most recent
	1 young adult	Number of TV advertisements

rewarded with bookstore vouchers of 45, 25, and 15 euros, respectively. We implemented this particular incentive system because we wanted to better model real decision making, where both the information and the process of acquiring it are often costly.

The experimenter tested each participant individually, and all sessions were audio-recorded. Participants took on average 25 min (ranging from 18 to 35 min) to complete the session, including reading the instructions. The experimenter read aloud the instructions, the two objects and criterion for each trial, and the values of the generated cues; this information was also displayed on the computer screen. To minimize potential effects of computer literacy, only the experimenter operated the computer.

Results

We compared the performances of the three age groups on four outcomes: (a) frugality, (b) accuracy (in terms of percentage of correct inferences), (c) specific cues generated, and (d) success of the generated cues.

Frugality

The *frugality* of a decision is indicated by the number of cues used to make an inference (Gigerenzer, Todd, & ABC Research Group, 1999). The smaller this number, the more frugal the decision. As a proxy for frugality, we used the number of cues generated by participants. As shown in Table 4 and confirmed by a repeated measures analysis of variance (ANOVA) with the factors label (2 levels: names vs. generic labels) and task (2 levels: cities vs. cars), participants made more frugal decisions when objects had names, $F(1, 66) = 24.97$, $p < .001$, $\eta^2 = .28$, which makes sense because names can carry information. Moreover, the participants generated more cues in the car price task, $F(1, 66) = 10.23$, $p = .002$, $\eta^2 = .14$. All other main or interaction effects on frugality had $p > .10$; in particular, we did not find any effect of age on frugality.

The numbers in Table 4 suggest that one-cue strategies were prevalent. Because participants needed to pay 1 token for each cue they asked for, this result is not too surprising per se. Indeed, in most trials, only one cue was generated. For example, in 64% of the trials (i.e., 22 of 34), younger children generated one cue in the car price task when the objects were presented with generic labels. If we consider only the generic labels condition, to be sure that no prior knowledge was taken into account, we can observe that participants relied on one-cue strategies in more than 50% of the trials. Moreover, younger children seemed to rely on one-cue strategies a bit more than the other two age groups, and only for children was there an effect of task, with children being more willing to ask for only one cue in the city population task.

Accuracy

Table 5 presents the accuracy results for the trials where participants did not ask for any cues. We deemed these inferences made without generating cues in the generic labels condition (in only 3 of 256 trials) to be random guesses. The results of the names condition might indicate inferences based on prior knowledge, but we cannot exclude them to be random guesses as well. We are not interested in further investigating these results.

Table 5 also shows the percentages of correct inferences for those problems where the participants generated at least one cue. A repeated measures ANOVA with the factors label (2 levels: names vs. generic labels) and task (2 levels: cities vs. cars) showed the following. First, there were no effects of labels on accuracy. Second, there was a main effect of age on accuracy, $F(2, 66) = 8.40$, $p = .001$, $\eta^2 = .21$. All post hoc analyses revealed no overall differences in terms of accuracy between younger children and young adults. Looking at Table 5, we can see that in the names condition, younger children in the car price task performed as well as young adults and even slightly better, whereas in the city population task young adults had an advantage. In the generic labels condition, surprisingly, younger children outperformed young adults in both tasks. A possible explanation for our results is that younger children were able to generate more successful cues and so performed better in the generic labels condition. Third, there was a main effect of task on accuracy, $F(1, 66) = 10.79$, $p = .002$, $\eta^2 = .15$; all participants performed better in the car price task.

Table 4

Experiment 1: Frugality in terms of mean number of cues generated by the participants and percentage of trials in which one cue was generated.

Group	City population task		Car price task	
	Names	Generic labels	Names	Generic labels
Mean number of cues generated				
Younger children	0.88 (0.38)	1.24 (0.44)	1.00 (0.79)	1.53 (0.93)
Older children	0.89 (0.45)	1.15 (0.52)	1.33 (0.79)	1.44 (0.76)
Young adults	1.00 (0.65)	1.39 (0.57)	1.18 (0.66)	1.36 (0.41)
Percentage of trials in which one cue was generated				
Younger children	82% (28 of 34)	82% (28 of 34)	42% (14 of 34)	64% (22 of 34)
Older children	76% (38 of 50)	72% (36 of 50)	42% (21 of 50)	52% (26 of 50)
Young adults	57% (25 of 44)	57% (25 of 44)	50% (22 of 44)	55% (24 of 44)

Note. For mean number of cues generated, standard deviations are in parentheses. For percentage of trials, natural frequencies are in parentheses.

Table 5
Experiment 1: Accuracy (as percentage of correct inferences) in problems where no cues were generated and in problems where at least one cue was generated.

Group	City population task		Car price task	
	Names	Generic labels	Names	Generic labels
Problems where no cues were generated				
Younger children	60% (5 of 34, *SD* = 51)	(0 of 34)	91% (11 of 34, *SD* = 48)	(0 of 34)
Older children	50% (10 of 54, *SD* = 42)	100% (1 of 54)	71% (9 of 54, *SD* = 43)	100% (1 of 54)
Young adults	67% (9 of 44, *SD* = 46)	100% (1 of 44)	80% (10 of 44, *SD* = 40)	(0 of 44)
Problems where at least one cue was generated				
Younger children	68% (29 of 34, *SD* = 42)	73% (34 of 34, *SD* = 31)	82% (23 of 34, *SD* = 35)	91% (34 of 34, *SD* = 20)
Older children	56% (44 of 54, *SD* = 40)	55% (53 of 54, *SD* = 40)	62% (45 of 54, *SD* = 38)	65% (53 of 54, *SD* = 40)
Young adults	73% (34 of 44, *SD* = 33)	53% (43 of 44, *SD* = 30)	77% (34 of 44, *SD* = 33)	86% (44 of 44, *SD* = 23)

Note. Number of such problems out of the total number of problems and standard deviations (*SD*) are in parentheses.

The cues generated

In the city population task, both older children and young adults generated more diverse cues than younger children (14, 21, and 6 cues, respectively). Even if we consider only the cues asked by at least two participants, older children and young adults still generated more diverse cues than younger children (10, 14, and 4 cues, respectively). In the car price task, the participants generated a similar number of diverse cues (younger children: 16; older children: 21; young adults: 14). The numbers of cues generated by at least two older children, young adults, and younger children were 9, 12, and 9, respectively. Fig. 1 shows the percentages of participants by age group who generated a certain cue in the city population task and in the car price task, taking into account only the cues generated by at least 10% of the participants of one age group.

In the city population task (Fig. 1A), more than 40% of all the participants ($M_{\text{younger children}}$ = 41%, $M_{\text{older children}}$ = 45%, and $M_{\text{young adults}}$ = 43%) generated the cue "area." Nearly half of the younger children (48%) generated the cue "number of buildings," a cue generated by only 13% of the older children and 4% of the young adults. The cues "tourism" and "density of population," with the former generated by 11% of the older children and the latter generated by 10% of the young adults, were both generated by only a few of the younger children.

In the car price task (Fig. 1B), two cues predominated for the younger children, with 31% of the children generating "width" and 22% generating "length"; 31% of the young adults generated the cue "horsepower," and 22% generated "capacity." The other cues were generated by relatively few of the young adult participants. Older children, on the other hand, seemed not to have a strong preference for any one cue even though many of them generated "length" (17%), "speed" (18%), "width" (12%), and "horsepower" (10%).

Success of generated cues

Table 6 displays the average success of the cues generated by the participants. A repeated measures ANOVA with the factors label (2 levels: names vs. generic labels) and task (2 levels: cities vs. cars) showed a main effect of labels, $F(1, 66) = 6.32$, $p = .015$, $\eta^2 = .09$; overall, when the objects were presented by using their names, participants generated more successful cues.

We found a main effect of age, $F(1, 66) = 7.74$, $p = .001$, $\eta^2 = .20$; all post hoc analyses revealed overall no difference between younger children and adults in terms of success of the generated cues, whereas older children generated less successful cues.

We also found a main effect of task, $F(1, 66) = 17.22$, $p < .001$, $\eta^2 = .22$, and an interaction between age group and task, $F(2, 66) = 4.86$, $p = .011$, $\eta^2 = .13$; in general, participants generated more success-

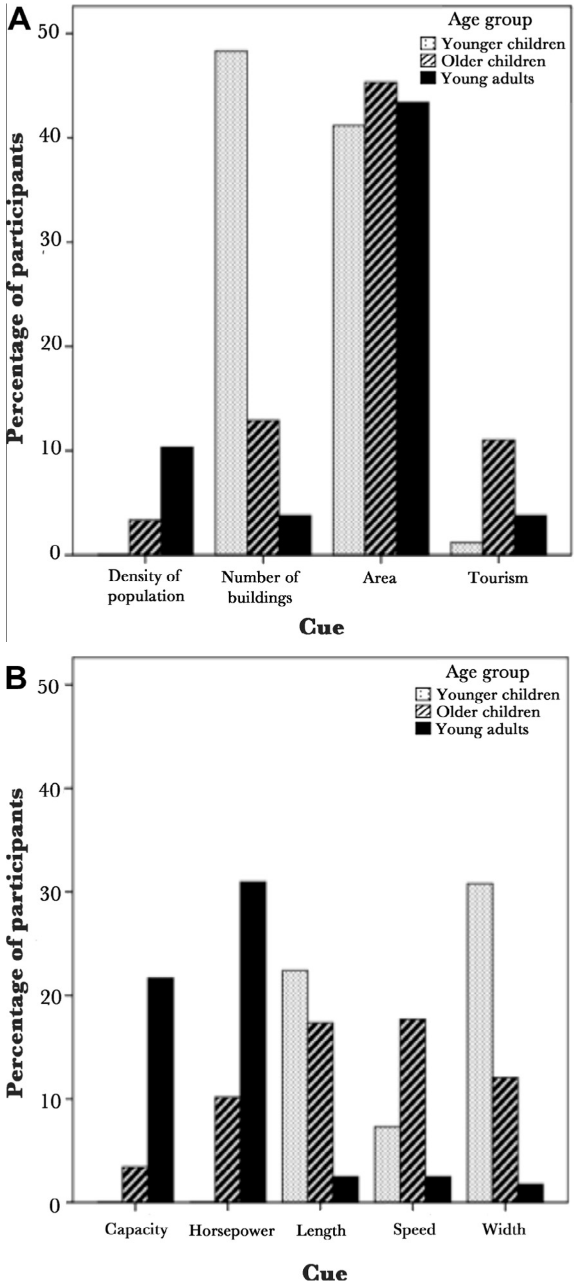

Fig. 1. Experiment 1: Percentages of participants by group who generated a certain cue in the city population task (A) and the car price task (B).

ful cues in the car price task. However, in the generic labels condition, younger children generated more successful cues in the city population task. This fits with our finding of higher accuracy of younger children in the city population task. But as we saw, younger children performed better in the car price task as well. One reason for their better performances is that younger children always

Table 6
Experiment 1: Average success of the cues generated by the participants.

Group	City population task		Car price task	
	Names	Generic labels	Names	Generic labels
Younger children	.72 (.18)	.75 (.19)	.75 (.13)	.70 (.12)
Older children	.63 (.15)	.63 (.15)	.76 (.11)	.74 (.10)
Young adults	.69 (.17)	.63 (.16)	.78 (.12)	.80 (.11)

Note. Standard deviations are in parentheses.

interpreted the cue in the right direction (i.e., the longer it is, the more expensive it should be), and they happened to apply a cue only in those problems where it led to a correct inference. On the contrary, six different adults in 6 trials did not correctly interpret the direction where the cue pointed; after generating only one cue, in 1 trial they chose the slower car, in 3 trials they chose the car with the lower power, and in 2 trials they chose the car with less horsepower.[3]

Summary of results of Experiment 1

In Experiment 1, we tested children and young adults on inference problems about real objects by prompting them to generate their own cues. We focused on the frugality of the decisions (i.e., the number of cues asked), the accuracy of the participants, the specific cues generated, and their success (which is a measure of how often the cue points to the correct inference).

We found that (a) all participants, and especially younger children, mostly generated only one cue; (b) children and young adults generated different cues; (c) for problems where no names were given—only generic labels such as Car 1 and Car 2—younger children outperformed the other two age groups in making cue-based inferences in both tasks; (d) participants belonging to the same age group generated the same one or two cues more often; and (e) younger children overall generated cues that were as successful as those generated by the young adults, with their cues being more successful than the cues generated by the young adults in the city population task.

In Experiment 2, we tested whether the results obtained in Experiment 1 were sensitive to the experimental design adopted as hypothesized.

Experiment 2: Fixed set of cues

Method

Participants

The experiment involved 75 participants: 24 children in second and third grades (11 female and 13 male, M_{age} = 8.4 years, SD = 0.72) and 15 children in fifth grade (7 female and 8 male, M_{age} = 10.3 years, SD = 0.46) of a primary school[4] in Livorno, Italy, and 36 young adults (7 female and 29 male, M_{age} = 17.4 years, SD = 0.65) recruited from a high school[5] in Livorno. The participants all were born in Italy and belonged to various social classes.

Design and procedure

In this second experiment, the design and objects used were the same as in Experiment 1 (see Table 1) with one crucial difference: Participants were not allowed to generate their own cues but rather needed to select them from a fixed set of five cues by clicking on corresponding buttons on the screen. There were two different sets of cues for each of the two inference tasks. One set comprised the five cues generated the most by the participants in Experiment 1 (as in Fig. 1, with the addition of the cue

[3] We could not resist the temptation to ask some of the students we interviewed what horsepower is. One of them innocently replied, "Horsepower? I really have no clue. But I suppose it is something I should ask if I have to pick one of two cars, no?"

[4] Istituto Sacro Cuore, Livorno, Italy.

[5] Liceo Scientifico "F. Enriques," Livorno, Italy.

Table 7
Experiment 2: Cues and their success in the two tasks for the new cues and old cues manipulations.

City population task		Car price task	
Cue	Success	Cue	Success
New cues manipulation			
Number of buildings	.92	Horsepower	.88
Density of population	.67	Capacity	.84
Tourism	.64	Speed	.83
Altitude	.56	Length	.78
Area	.52	Width	.77
Old cues manipulation			
University	.76	Horsepower	.88
Train stations	.74	Capacity	.84
Museums	.70	Type of fuel	.67
Soccer teams	.66	Fuel consumption	.62
Regional capital	.52	Number of seats	.58

"altitude" in the city population task, not included in the figure because it was generated by less than 10% of the participants). We refer to this set of cues as the "new cues." The other set comprised, for the cities, five of the nine cues in the original city task of Gigerenzer and colleagues (1999): number of universities, number of museums, number of soccer teams playing in the first league, number of train stations, and whether or not the city is a regional capital (instead of the original state capital cue). We excluded four cues from the original task: three that were not applicable in Italy (the license plate, industrial belt, and East Germany cues) and one that was the least successful of the original nine (national capital cue). Moreover, we modified some of the other cues ("Is the city on the Intercity line?" became "number of train stations," and "Was the city once an exposition site?" became "number of museums") mainly to have cues with continuous values instead of binary cues. For the cars, the second cue set comprised the first five cues shown in car descriptions on a popular Italian website advertising and selling used cars.[6] We refer to the set of cues taken from Gigerenzer and colleagues (1999) and the used car website as the "old cues" to indicate that these cues come from prior research stimuli. Some of the old cues coincide with the new cues generated by the participants in Experiment 1. Because of the rationale that led us to the selection of the cues in both cue sets, we saw no reason to exclude any of them. All cues and their relative success rates are presented in Table 7.

Experiment 2 differed in that participants faced in total 40 trials divided into two rounds of 20 trials. In each round, we gave the participants either the old cues set or the new cues set, so that the cue set manipulation (old vs. new cues) is considered a within-participant variable. Each round consisted of 5 trials for each of the four blocks constituting the 2×2 matrix design, having as independent variables the task (city population or car price) and the objects' presentation (by using their names or by using generic labels). The order of the rounds and the four blocks within the rounds, as well as the order of the manipulations and the position of the cue buttons on the screen, was randomized.

Results

We compared the performances of the three age groups according to the same four criteria we used for analyzing the results in Experiment 1: (a) frugality, (b) accuracy (in terms of percentage of correct inferences), (c) specific cues selected, and (d) success of the selected cues.

Frugality

Regarding the average number of cues the participants consulted (Table 8), a repeated measures ANOVA with the factors label (2 levels: names vs. generic labels), task (2 levels: cities vs. cars), and

[6] http://www.autosupermarket.it.

Table 8
Experiment 2: Frugality in terms of mean number of cues selected by the participants and percentage of trials in which one cue was selected for the new cues and old cues manipulations.

Group	City population task		Car price task	
	Names	Generic labels	Names	Generic labels
Mean number of cues selected				
New cues manipulation				
Younger children	1.9 (1.6)	2.2 (1.5)	2.3 (1.5)	2.2 (1.3)
Older children	0.9 (1.0)	1.3 (1.4)	0.9 (0.8)	1.0 (0.9)
Young adults	0.9 (0.6)	1.7 (0.7)	1.3 (0.9)	1.6 (0.5)
Old cues manipulation				
Younger children	1.9 (1.6)	1.9 (1.2)	1.8 (1.2)	1.7 (1.0)
Older children	0.5 (0.8)	1.0 (0.8)	0.7 (0.6)	0.9 (0.6)
Young adults	0.6 (0.5)	1.7 (0.6)	1.2 (0.5)	1.5 (0.6)
Percentage of trials in which one cue was selected				
New cues manipulation				
Younger children	35% (42 of 120)	41% (49 of 120)	29% (35 of 120)	31% (37 of 120)
Older children	37% (28 of 75)	45% (34 of 75)	39% (29 of 75)	52% (39 of 75)
Young adults	29% (52 of 180)	32% (58 of 180)	31% (56 of 180)	44% (79 of 180)
Old cues manipulation				
Younger children	30% (36 of 120)	43% (52 of 120)	40% (48 of 120)	57% (68 of 120)
Older children	25% (19 of 75)	49% (37 of 75)	45% (34 of 75)	55% (41 of 75)
Young adults	29% (52 of 180)	45% (81 of 180)	44% (80 of 180)	61% (109 of 180)

Note. For mean number of cues selected, standard deviations are in parentheses. For percentage of trials, natural frequencies are in parentheses.

cue set (2 levels: old vs. new cues) revealed a main effect of the cue set manipulation, $F(1, 75) = 12.82$, $p = .001$, $\eta^2 = .15$; participants selected more cues in the new cues manipulation. We found no main effect of task, but we did find a main effect of age on the number of cues participants looked up, $F(1, 75) = 9.40$, $p < .001$, $\eta^2 = .21$. All post hoc analyses revealed no difference in terms of number of cues looked up between young adults and older children, whereas younger children looked up more cues than the other two age groups.

The repeated measures ANOVA also showed that, as expected given that the names carry information, participants looked up more cues in the condition where the objects were presented with generic labels, $F(1, 75) = 33.81$, $p < .001$, $\eta^2 = .32$. We also found an interaction effect between age and labels, $F(2, 75) = 14.84$, $p < .001$, $\eta^2 = .29$, and an interaction between task and labels, $F(1, 75) = 18.02$, $p < .001$, $\eta^2 = .20$.

The numbers in Table 8 suggest that in Experiment 2, unlike in Experiment 1, one-cue strategies were used in less than 50% of the trials, except for the car price trials of the old cues manipulation, in the generic labels condition. However, the incentive system was the same as in Experiment 1; that is, participants still needed to pay 1 token for each cue they asked for. We did not find any consistent effect of age on the percentage of trials in which only one cue was selected.

Accuracy

As in Experiment 1, we first looked at the accuracy reached by the participants n the trials where they did not ask for any cues (Table 9). We found that in the generic labels condition, younger children and adults in only few trials made inferences without acquiring any information, whereas the number of trials in which older children randomly guessed was higher (20–24 trials of 75). In the names condition, as already pointed out, the results might indicate inferences from prior knowledge even though we cannot rule out that they were random guesses. Because we are interested mainly in the process of cue selection, we do not report on these results.

The percentages of correct inferences for those problems where the participants looked up at least one cue are shown in Table 9. A repeated measures ANOVA with the factors label (2 levels: names vs. generic labels), task (2 levels: cities vs. cars), and cue set (2 levels: old vs. new cues) revealed no effect

Table 9
Experiment 2: Accuracy (as percentage of correct inferences) in problems where no cues were selected and in problems where at least one cue was selected for the new cues and old cues manipulations.

Group	City population task		Car price task	
	Names	Generic labels	Names	Generic labels
Problems where no cues were selected				
New cues manipulation				
Younger children	76% (21 of 120, *SD* = 18)	44% (9 of 120, *SD* = 19)	33% (12 of 120, *SD* = 18)	50% (8 of 120, *SD* = 19)
Older children	70% (33 of 75, *SD* = 25)	46% (24 of 75, *SD* = 21)	50% (32 of 75, *SD* = 27)	37% (24 of 75, *SD* = 23)
Adults	79% (73 of 180, *SD* = 23)	60% (10 of 180, *SD* = 19)	71% (49 of 180, *SD* = 15)	29% (7 of 180, *SD* = 17)
Old cues manipulation				
Younger children	62% (29 of 120, *SD* = 20)	33% (9 of 120, *SD* = 22)	37% (16 of 120, *SD* = 21)	33% (6 of 120, *SD* = 25)
Older children	75% (49 of 75, *SD* = 25)	38% (21 of 75, *SD* = 24)	70% (33 of 75, *SD* = 22)	60% (20 of 75, *SD* = 30)
Adults	85% (101 of 180, *SD* = 18)	83% (6 of 180, *SD* = 18)	80% (40 of 180, *SD* = 16)	80% (5 of 180, *SD* = 15)
Problems where at least one cue was selected				
New cues manipulation				
Younger children	66% (99 of 120, *SD* = 23)	56% (111 of 120, *SD* = 26)	83% (108 of 120, *SD* = 21)	77% (112 of 120, *SD* = 25)
Older children	69% (42 of 75, *SD* = 25)	69% (51 of 75, *SD* = 25)	70% (43 of 75, *SD* = 30)	72% (51 of 75, *SD* = 27)
Young adults	78% (107 of 180, *SD* = 18)	78% (170 of 180, *SD* = 19)	90% (131 of 180, *SD* = 17)	88% (173 of 180, *SD* = 15)
Old cues manipulation				
Younger children	80% (91 of 120, *SD* = 19)	68% (111 of 120, *SD* = 24)	73% (104 of 120, *SD* = 22)	79% (114 of 120, *SD* = 20)
Older children	54% (26 of 75, *SD* = 22)	63% (54 of 75, *SD* = 23)	71% (42 of 75, *SD* = 15)	71% (55 of 75, *SD* = 22)
Young adults	66% (79 of 180, *SD* = 15)	83% (174 of 180, *SD* = 13)	84% (140 of 180, *SD* = 19)	86% (175 of 180, *SD* = 16)

Note. Number of such problems out of the total number of problems and standard deviations (*SD*) are given in parentheses.

of the cue set manipulation on accuracy. We also found no effects of labels but found a main effect of task on accuracy, $F(1, 75) = 12.52$, $p = .001$, $\eta^2 = .15$; for the most part, participants performed better in the car price task, with the exception being that younger children performed better in the city population task in the names condition with the old cues manipulation.

We also found a main effect of age on accuracy, $F(2, 75) = 26.68$, $p < .001$, $\eta^2 = .43$, and all post hoc analyses confirmed that, overall, adults always performed better than children, and younger children always performed better than older children. Moreover, there was an interaction effect between task and labels, $F(1, 75) = 6.50$, $p = .013$, $\eta^2 = .08$, and an interaction among task, cue set manipulation, and age group, $F(2, 75) = 3.75$, $p = .028$, $\eta^2 = .09$.

The cues selected

Figs. 2 and 3 show the percentages of participants by age group who selected particular cues in the two manipulations for both tasks. In the city population task with the new cues manipulation (Fig. 2A), young adults focused on three main cues, with 44% of them selecting the cue “density of population,” 23% selecting the cue “area” (the least successful cue in the set), and only 29% selecting the cue “number of buildings” (the most successful cue in the set). In contrast, younger and older children showed no preference for any particular cue, being very close to the random distribution of 20% per cue, even though fewer of them selected the cue “altitude.”

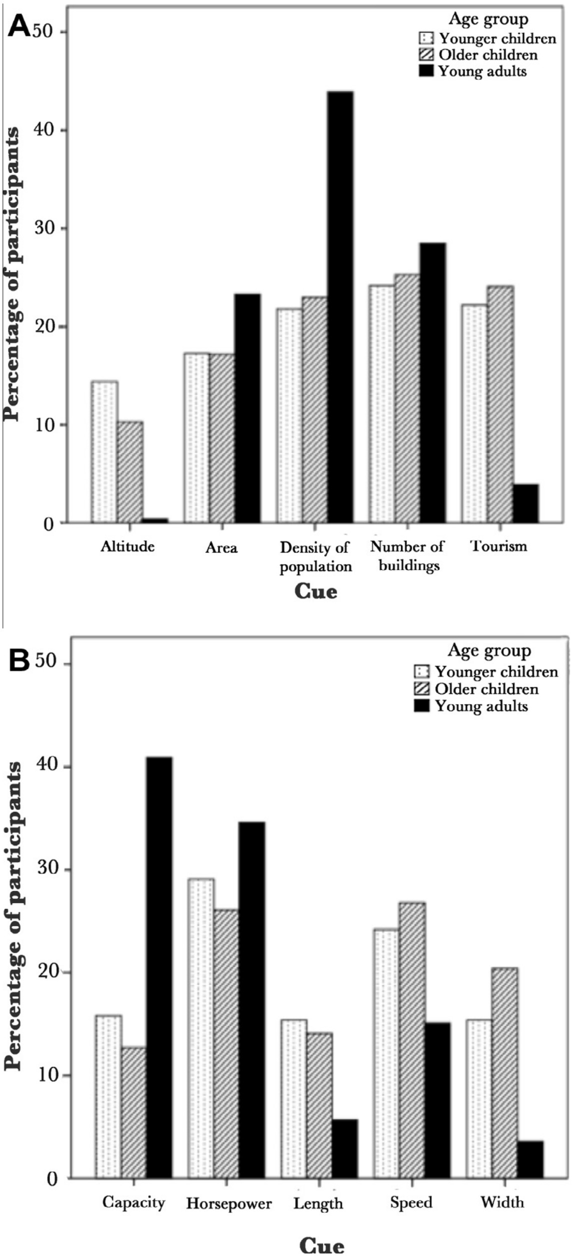

Fig. 2. Experiment 2: Percentages of participants by group who selected particular cues in the city population task (A) and the car price task (B) with the new cues manipulation.

In the car price task of this manipulation (Fig. 2B), young adults showed a similar pattern, focusing on mainly two cues: "capacity" (41%) and "horsepower" (35%)—the most successful cues in the set. Children were a bit more selective in this task, focusing on the cues "horsepower" (29% of the younger children and 26% of the older children) and "speed" (24% of the younger children and 27% of the older children).

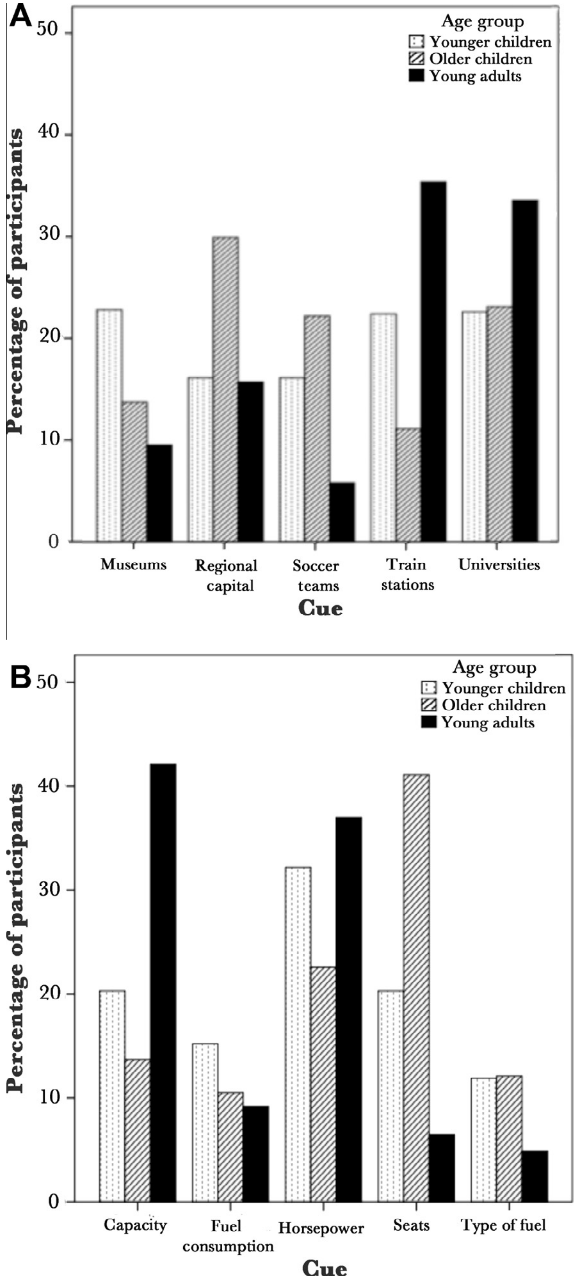

Fig. 3. Experiment 2: Percentages of participants by group who selected particular cues in the city population task (A) and the car price task (B) with the old cues manipulation.

In the city population task of the old cues manipulation (Fig. 3A), young adults focused mainly on the cues "train stations" (35%) and "universities" (34%). Older children selected the cue "regional capital" a bit more often (29%), whereas younger children's selection of cues was close to the random distribution. In the car price task, all of the age groups seemed to be more selective; young adults

Table 10
Experiment 2: Average success of the cues selected by the participants for the new cues and old cues manipulations.

Group	City population task		Car price task	
	Names	Generic labels	Names	Generic labels
New cues manipulation				
Younger children	.69 (.08)	.68 (.06)	.83 (.02)	.83 (.02)
Older children	.69 (.12)	.72 (.08)	.82 (.03)	.81 (.03)
Young adults	.69 (.11)	.71 (.10)	.86 (.02)	.85 (.02)
Old cues manipulation				
Younger children	.69 (.03)	.70 (.03)	.76 (.06)	.76 (.07)
Older children	.63 (.08)	.64 (.06)	.67 (.08)	.67 (.06)
Young adults	.71 (.04)	.72 (.04)	.82 (.05)	.82 (.04)

Note. Standard deviations are in parentheses.

chose mainly to look at the cues ‘‘capacity’’ (42%) and ‘‘horsepower’’ (37%), older children to look at the cue ‘‘number of seats’’ (41%), and younger children to look at the cue ‘‘horsepower’’ (32%).

Success of selected cues

In Table 10, we can see the average success of the cues selected by the participants. Overall, a repeated measures ANOVA with the factors label (2 levels: names vs. generic labels), task (2 levels: cities vs. cars), and cue set (2 levels: old vs. new cues) revealed no differences between the two conditions (names vs. generic labels). The quality of the cues selected in the car price task was higher than that selected in the city population task, $F(1, 55) = 130.03$, $p < .001$, $\eta^2 = .71$. This is not surprising given the more successful cues available in this task. The average success of the cues consulted in the new cues manipulation was also higher, $F(1, 55) = 32.12$, $p < .001$, $\eta^2 = .38$; again, this is due to the fact that the cues available in this manipulation were, on average, more successful.

The analysis revealed a main effect of age on quality, $F(2, 55) = 17.53$, $p < .001$, $\eta^2 = .40$; all post hoc analyses showed that young adults, overall, selected more successful cues than younger ad older children, and younger children selected more successful cues than older children.

We also found an interaction effect between task and the cue set manipulation, $F(1, 55) = 17.74$, $p < .001$, $\eta^2 = .25$, and an interaction between age group and cue set manipulation, $F(2, 55) = 8.84$, $p < .001$, $\eta^2 = .25$.

Summary of results of Experiment 2

In Experiment 2, we tested children and young adults on the same inference problems used in Experiment 1. In this experiment, however, we adopted an experimental design that has been predominantly used when testing cue-based inferences; participants needed to choose which cues to consult from a preliminarily defined fixed set. These five cues in the new cues manipulation were the top five cues generated by the participants in Experiment 1; in the old cues manipulation, they were either five of the cues used in the original city population task (Gigerenzer et al., 1999) or the first five cues shown in car descriptions on a popular Italian website advertising and selling used cars. As in the first experiment, we focused on the frugality and accuracy of the inferences made as well as on the specific cues generated and on their success.

We found that (a) participants used one-cue strategies in less than 50% of the trials in nearly all of the conditions; (b) young adults always outperformed the other two age groups in making cue-based inferences; (c) young adults always focused on some specific cues even though they were not always the most successful ones (as in the city population task), whereas children did not focus on any cues in particular; and (d) young adults selected more successful cues than children. Note that these results are different from those of Experiment 1, where participants were free to generate cues.

General discussion

In this study, we tested children and young adults on inference problems by using both a new experimental procedure and an established one; in Experiment 1 participants were prompted to generate their own cues, whereas in Experiment 2 they needed to select from a list of available cues, either old cues that have been used in the previous literature (Gigerenzer et al., 1999) or new cues that were generated by other participants. This comparison allowed us to investigate whether and under what conditions children are capable of focusing on the most informative cues and, in general, to explore the generation of cues that people find informative, a so far little-studied issue in cue-based inference.

First, we found that in Experiment 1 all participants, and especially younger children, mostly generated one cue, whereas in Experiment 2 younger children looked up more cues even though in both experiments participants needed to pay 1 token for each cue consulted. This finding is inconsistent with previous research showing that younger children are less efficient in their search behavior than older children because they gather more information (Davidson, 1991b; Howse et al., 2003). There are two possible explanations for this result that are not mutually exclusive but compatible. On the one hand, it may have been that selecting cues from a given set was much less effortful than generating them from scratch, imposing a cognitive cost that led children to search more frugally (Katz et al., 2010). On the other hand, in Experiment 1, participants could only compare cues against others in their own *cuebox*, which we define as the set of cues available and usable for making an inference. Assuming that they had generated the most successful cue in their cuebox first, participants might have been so confident in its predictive power that they thought the inference could be based on only that cue. In Experiment 2, in contrast, participants, presented with cues that they would not have considered otherwise, might have doubted the ranking of the most successful cues in their cuebox. This, together with the fact that the cues given to them were all informative cues, may have prompted participants to look up more cues to gain more confidence in which inference to make.

Second, in Experiment 1, for problems in which objects were presented with generic labels instead of names, younger children outperformed the other two age groups in making cue-based inferences in both tasks. On the other hand, in Experiment 2, young adults always outperformed children. We think that the accuracy in cue-based inferences depends on the success of the cues generated/selected (see below) and on the ability of the participants to interpret the cues correctly, that is, to understand the direction in which the cues point.

Third, we found that in Experiment 1 children focused on one or two specific cues, generating them more frequently than other cues, whereas in Experiment 2 overall they did not focus on any of the available cues in particular. Moreover, we found that in both experiments young adults asked for task-specific cues that were somewhat "technical" and referred to hidden features of the objects, for example, horsepower or capacity in the car price task and density of population or area in the city population task. In contrast, the cues that younger children focused on in Experiment 1 were more perceptual, observable cues, for example, number of buildings in the city population task and length or height in the car price task.

Indeed, children, being novices in most domains of inference, may first learn mostly perceptual cues because young children are often "concrete" or "perceptually bound" (Flavell, 1985; Wartella, Daniel, Scott, Jacob, & Allison, 1979), even though this is less of a limitation than originally believed (Keil, 1989; Wellman & Gelman, 1988). As they grow up, children face new inference domains and deepen their knowledge of old domains, continually acquiring and reorganizing domain-specific knowledge (Carey, 1985). As children understand and grasp more concepts and terms, they also learn cues that go beyond appearances, cues related to hidden features and structures (e.g., horsepower, density of population) that are possibly more "technical." This learning process leads to an update of cues in the cuebox that we argue is biased; the more technical cues are intuitively assumed to be more informative than the general perceptual cues that were there before because they are more elaborate and appropriate—even when we do not know exactly what they mean or how to interpret them. By *bias,* we mean an asymmetrical tendency that is not good or bad per se; a bias that favors technical cues often leads to more accurate inferences, but these technical cues are not always the most informative.

Indeed, our fourth main result shows that in Experiment 1 the children, somehow naturally biased toward perceptual cues, generated cues that were overall as successful as those generated by the young adults and even better in the city population task, where perceptual cues *happened* to be more successful.

In conclusion, we have shown that "making your own kind of cues," paraphrasing Mama Cass Elliot, matters. Compared with the setup of Experiment 1, the set of cues presented in Experiment 2 indeed had a very different, if not opposite, effect on our participants. Needing to compare cues they would not have considered otherwise might have been confusing for all of the participants, who felt tempted to look up more cues to increase their confidence about their inference. However, young adults recognized that some of the cues offered were even more successful than any they would have thought of and selected them. In contrast, children, probably unfamiliar with some of the cues presented, could not recognize which cues were the most successful. Thus, the procedure used in Experiment 2, by providing all of the participants with a same cuebox *unbiased* them; children were no longer biased toward perceptual cues, and young adults were no longer biased toward technical cues. Although this may result in an improvement for adults, who are able to identify in a set of given cues those that are the most successful, it might be disadvantageous for children, who lose the advantage that the "perceptual bias" gives them in some tasks without being able to compensate for the loss. However, in this first exploration of cue generation, a question our two experiments cannot answer is whether the children and young adults in Experiment 1 generated the best cues they could or the only ones they had available in their cuebox. Moreover, if the children's perceptual bias has been already thoroughly investigated, both as limit and advantage (John & Sujan, 1990; Springer, 2001), future research should further investigate and test our conjecture about adults' bias toward technical cues.

Finally, we find the generation paradigm to be a stimulating and promising methodology. From a perspective of ecological rationality (Todd, Gigerenzer, & ABC Research Group, 2012), we believe that generating cues from scratch, instead of choosing them from a catalogue, can be a crucial paradigm breakthrough in studying cue-based decision making and, in general, the process of searching for information that actually takes place when people make choices and decisions, assess preferences, generate hypotheses, and trace causes.

Acknowledgments

Thanks go to Uwe Czienskowski for programming the experiments, to Francesca Ricci, Eleonora Gilli, and Claudia Mazzeranghi for collecting the data, and to the teachers of the schools used for recruitment for their support, especially Maria Rosaria Angioni and Simona Preti. Thanks also go to Henrik Olsson, Nicolai Bodemer, Michaela Gummerum, Katherine McMahon, and Anita Todd for their comments on the many versions of this article.

References

Bereby-Meyer, Y., Assor, A., & Katz, I. (2004). Children's choice strategies: The effects of age and task demands. *Cognitive Development, 19*, 127–146.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Chouinard, M. M. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development*, *72*(1, Serial No. 286).

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.

Davidson, D. (1996). The effects of decision characteristics on children's selective search of predecisional information. *Acta Psychologica, 92*, 263–281.

Flavell, J. (1985). *Cognitive development*. Englewood Cliffs, NJ: Prentice Hall.

Gigerenzer, G., & Todd, P. M.ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford 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*, 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*, 524–536.

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.

Gregan-Paxton, J., & Roedder John, D. (1997). The emergence of adaptive decision making in children. *Journal of Consumer Research, 24*, 43–56.

Hagen, J. W., & Hale, G. H. (1973). The development of attention in children. In A. D. Pick (Ed.). *Minnesota Symposium on child psychology* (Vol. 7, pp. 117–140). Minneapolis: University of Minnesota Press.

Howse, R. B., Best, D. L., & Stone, E. R. (2003). Children's decision making: The effects of training, reinforcement, and memory aids. *Cognitive Development, 18*, 247–268.
John, D. R., & Sujan, M. (1990). Children's use of perceptual cues in product categorization. *Psychology and Marketing, 7*, 277–294.
Katz, I., Bereby-Meyer, Y., Assor, A., & Danziger, S. (2010). Children's adaptive pre-decisional search behavior: Effects of memory and number of alternatives. *Journal of Economic Psychology, 31*, 17–24.
Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: Bradford Books.
Mata, R., von Helversen, B., & Rieskamp, J. (2011). When easy comes hard: The development of adaptive strategy selection. *Child Development, 82*, 687–700.
Mosher, F. A., & Hornsby, J. R. (1966). On asking questions. In J. S. Bruner, R. R. Olver, & P. M. Greenfield (Eds.), *Studies in cognitive growth* (pp. 86–102). New York: John Wiley.
Newell, B. R., Rakow, T., Weston, N. J., & Shanks, D. R. (2004). Search strategies in decision making: The success of "success". *Journal of Behavioral Decision Making, 17*, 117–137.
Pick, A. D., & Frankel, G. W. (1973). A study of strategies of visual attention in children. *Developmental Psychology, 9*, 348–357.
Ruggeri, A., & Feufel, M. A. (2013). *How object representation and domain affect categorization* Unpublished manuscript, Max Planck Institute for Human Development.
Springer, K. (2001). Perceptual boundedness and perceptual support in conceptual development. *Psychological Review, 108*, 691–708.
Todd, P. M., & Gigerenzer, G.ABC Research Group. (2012). *Ecological rationality: Intelligence in the world*. New York: Oxford University Press.
Vosniadou, S. (1994). Universal and culture-specific properties of children's mental models of the earth. In L. Hirshfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 412–430). Cambridge, UK: Cambridge University Press.
Wartella, E., Daniel, B. W., Scott, W., Jacob, S., & Allison, A. (1979). The young child as consumer. In E. Wartella (Ed.), *Children communicating: Media and development of thought, speech, understanding* (pp. 251–279). Beverly Hills, CA: Sage.
Wellman, H. M., & Gelman, S. A. (1988). Children's understanding of the nonobvious. In R. J. Sternberg (Ed.). *Advances in the psychology of human intelligence* (Vol. 4, pp. 99–135). Hillsdale, NJ: Lawrence Erlbaum.