The Development of Adaptive Decision Making: Recognition-Based Inference in Children and Adolescents

Sebastian S. Horn
Max Planck Institute for Human Development, Berlin, Germany

Azzurra Ruggeri
Max Planck Institute for Human Development, Berlin, Germany, and University of California, Berkeley

Thorsten Pachur
Max Planck Institute for Human Development, Berlin, Germany

Judgments about objects in the world are often based on probabilistic information (or cues). A frugal judgment strategy that utilizes memory (i.e., the ability to discriminate between known and unknown objects) as a cue for inference is the recognition heuristic (RH). The usefulness of the RH depends on the structure of the environment, particularly the predictive power (validity) of recognition. Little is known about developmental differences in use of the RH. In this study, the authors examined (a) to what extent children and adolescents recruit the RH when making judgments, and (b) around what age adaptive use of the RH emerges. Primary schoolchildren ($M = 9$ years), younger adolescents ($M = 12$ years), and older adolescents ($M = 17$ years) made comparative judgments in task environments with either high or low recognition validity. Reliance on the RH was measured with a hierarchical multinomial model. Results indicated that primary schoolchildren already made systematic use of the RH. However, only older adolescents adaptively adjusted their strategy use between environments and were better able to discriminate between situations in which the RH led to correct versus incorrect inferences. These findings suggest that the use of simple heuristics does not progress unidirectionally across development but strongly depends on the task environment, in line with the perspective of ecological rationality. Moreover, adaptive heuristic inference seems to require experience and a developed base of domain knowledge.

Keywords: cognitive development, adaptive strategy selection, recognition memory, fast and frugal heuristics, Bayesian cognitive modeling

Supplemental materials: http://dx.doi.org/10.1037/dev0000181.supp

From infancy to old age, humans make judgments based on pieces of information (or cues) that are only probabilistically related to some criterion (Brunswik, 1943). For example, infants may infer the edibility of objects based on probabilistic combinations of color, shape, or smell (Wertz & Wynn, 2014); children judge whether it is safe to cross a street from the distance of oncoming vehicles (Hoffrage, Weber, Hertwig, & Chase, 2003); and taxi drivers must quickly evaluate signs of trustworthiness before giving potential customers a ride (Gambetta & Hamill, 2005). How does the ability to make such inferences develop?[1] A growing body of research suggests that the core cognitive abilities and "building blocks" required for probabilistic inference are developed relatively early (Bonawitz, Denison, Gopnik, & Griffiths, 2014; Denison & Xu, 2014; for reviews, see Reyna & Brainerd, 1994; Schlottmann & Wilkening, 2011). However, the *strategic* and *adaptive* use of probabilistic cues may emerge considerably later (Betsch & Lang, 2013; Mata, von Helversen, & Rieskamp, 2011).

The way humans utilize cues for inference has often been described in terms of strategies, that is, goal-directed mental operations used to

[1] We use the term "inference" to refer to probabilistic judgments about magnitudes on a criterion dimension that is not directly accessible, based on available pieces of information (i.e., cues; Brunswik, 1943). The inference task in our investigation thus differs conceptually from the inference problems typically examined in the context of sentence and language comprehension (e.g., Reyna & Kiernan, 1994), which are not the topic of this article.

solve a problem (e.g., Bereby-Meyer, Assor, & Katz, 2004; Betsch & Lang, 2013; Davidson, 1991a; Mata et al., 2011; Mata, Josef, & Lemaire, 2015). The discovery and application of such strategies are key aspects during cognitive development, as children encounter many unfamiliar situations that require efficient use of information (Siegler, 1999). In this view, children acquire a repertoire of strategies and adaptive behavior arises from selecting an appropriate strategy as a function of the situation—an idea that also figures prominently in the adaptive-toolbox perspective on judgment and decision making (Gigerenzer, 2003). Perhaps the most basic strategy in the context of cue-based inference is to rely on recognition memory, that is, to consider whether or not an object has been encountered before. The capacity for recognition appears to be early developed and robust. Already newborns are able to discriminate between familiar and unfamiliar faces or voices, and infants' recognition memories for objects can last several weeks or months (Bjorklund, 2011; Fagan, 1970; Schneider, 2015). Importantly, recognition—or a lack thereof—is an informative cue in many natural environments, where it is systematically rather than randomly distributed (Goldstein & Gigerenzer, 2002). Recognized objects thus often differ from novel ones on relevant dimensions. Life-span research has shown that both younger and older adults strategically rely on recognition when making inferences. Moreover, adults' reliance on recognition seems to be largely adaptive, in the sense that they are more likely to follow recognition in environments where it is a good rather than poor cue (Pachur, Mata, & Schooler, 2009; Pohl, 2006). Notably, it is an open question whether (and to what extent) individuals at the beginning of their life span are already able to strategically rely on recognition for their judgments and whether they adapt their reliance on this cue across contexts, depending on its predictive power.

In this study, we investigate developmental differences in the use of the recognition heuristic (RH), which models people's strategic use of recognition when making inferences about objects in the world (Goldstein & Gigerenzer, 2002). We presented school-age children, younger adolescents, and older adolescents with an inference task in two environments—one in which recognition was a good cue, and one in which it was a poor cue. We then used a hierarchical cognitive-modeling approach to measure participants' reliance on the RH when making these inferences. In contrast to most developmental studies on cue-based inference, which have trained children in the lab to consider a set of predefined cues (e.g., Betsch & Lang, 2013; Davidson, 1991a; Mata et al., 2011), we took a complementary, ecological approach (Brunswik, 1943; see also Bronfenbrenner, 1979) and examined children's and adolescents' use of cue knowledge acquired in the real world, such as whether or not they had heard of certain names of cities before. In the following, we first outline previous research on developmental trends in the use of judgment and decision strategies. We then turn to the development of recognition memory and discuss the RH in more detail. Finally, we present the research questions and hypotheses of the current study.

Developmental Trends in the Use of Judgment and Decision Strategies

Decision performance shows substantial improvement across childhood and adolescence (Jacobs & Klaczynski, 2005). One way to understand these changes is in terms of developments in strategy use, which have been shown to be closely coupled to changes in basic-level processes of cognitive control and attentional resources

(Bjorklund, 2011). Due to constraints in working memory capacity and processing speed (Fry & Hale, 2000), younger children may rely more on simple, information-frugal decision strategies than older children or adolescents do, even in situations in which more complex strategies (i.e., so-called compensatory strategies that weigh and integrate multiple pieces of information) would yield higher performance. Developmental research on probabilistic, cue-based judgments and on multiattribute decisions has indeed shown that younger children sometimes overrely on simple strategies.[2] For instance, Bereby-Meyer et al. (2004) found in multiattribute choice between consumer products (e.g., bicycles, computer games) that 8- to 9-year-olds relied more frequently on noncompensatory strategies than 12- to 13-year-olds did. Noncompensatory strategies consider only one piece of information at a time and ignore any further information that might be available. Several studies with adults have found that decision makers are more likely to select such simple strategies when resources are scarce (e.g., under working memory load, Bröder & Schiffer, 2006; or under time pressure, Hilbig, Erdfelder, & Pohl, 2012; Pachur & Hertwig, 2006). Relatedly, Mata, Schooler, and Rieskamp (2007) observed reduced information search and increased reliance on simple heuristic strategies in older adults and found that it was attributable to age-related reductions in fluid cognitive abilities. From a life span view, these latter findings are also of interest for developmental research on children's inferences, because older adults and younger children can show similar patterns of performance in cognitive tasks (e.g., Dempster, 1992).

On the contrary, there also is evidence suggesting that younger children have greater difficulty than older children in implementing simple strategies. Mata et al. (2011) examined cue-based decisions in 9- to 10-year-olds, 11- to 12-year-olds, and young adults. Their participants were asked to infer, on the basis of six cues (that were probabilistically related to the criterion), which of three cars would win a race. Younger children had more difficulty than older children in using an information-frugal heuristic strategy, and this finding was attributed to their less-developed abilities to attend selectively to relevant information. Davidson (1991b) reported that younger children were less likely than older children to focus on important information, and Betsch, Lang, Lehmann, and Axmann (2014) found that only adults—but neither preschoolers ($M = 6$ years) nor elementary schoolchildren ($M = 10$ years)—were guided in their search by the validity of the cues. Davidson (1991a) found that younger children (aged 7–9 years) searched information less systematically and more exhaustively than older children (aged 10–11 and 13–14 years), often attending to irrelevant information. As suggested by Gregan-Paxton and John (1997), such patterns could be due to younger children being less sensitive to the search cost required by more complex strategies. These authors have shown that 9-year-olds benefit from the

---

[2] In this section, we consider strategy use in both (nonprobabilistic) multiattribute tasks and (probabilistic) multiple-cue tasks. In both task types, individuals are exposed to multiple pieces of information that may inform judgment or choice. Note, however, that adaptive search abilities may develop earlier in nonprobabilistic than in probabilistic tasks (Betsch, Lehmann, Lindow, Lang, & Schoemann, 2016). Relatedly, Schlottmann and Wilkening (2011) suggested that judgment tasks are often less complex than choice tasks and younger children could more likely incorporate probabilistic information in the former than in the latter. For further discussion of task-format effects, see Betsch et al. (2016).

imposition of additional cost that prevents them from engaging in overly extensive search.

In sum, adolescents and older children may be *less* reliant on simple heuristics than younger children, because working memory development allows them to use increasingly complex strategies (e.g., Bereby-Meyer et al., 2004). Conversely, there is also evidence suggesting that increasing age could be associated with a *more* pronounced use of simple strategies, because reliance on heuristic strategies can require selective attention (ignoring less important cues) and inhibition of inappropriate responses—abilities that develop only later in childhood (Mata et al., 2011; see also Betsch et al., 2014).

Notably, this latter possibility is also supported by findings showing that access to semantic relations between items (Bjork-lund, 1987) and reliance on meaning-based memory representations ("gist" memories), both of which have been argued to promote the use of heuristics and rules of thumb, seem to increase from childhood to adulthood (Reyna & Brainerd, 1994; see also Jacobs & Klaczynski, 2005). Further, several developmental studies have reported age-related increases in heuristic reliance on information such as "availability" (Davies & White, 1994) or "representativeness" (Jacobs & Potenza, 1991). In the next section, we introduce a specific model of recognition-based inference—the RH—and consider how developmental trends might impact its use (for a discussion of conceptual differences between heuristic use of recognition, fluency, and availability, see Pachur, Todd, Gigerenzer, Schooler, & Goldstein, 2011).

Strategic Use of Recognition

The RH (Goldstein & Gigerenzer, 2002) is a model of comparative judgment on a criterion dimension that is not directly accessible (i.e., on which exact properties or values of objects are unknown). According to the RH, if one item is recognized (e.g., Chicago), but the other not (e.g., Akron), then the recognized item is judged to have the higher value on the criterion (the population of cities). The RH makes the strong assumption that recognition is used as a noncompensatory cue, implying that any further cue knowledge available about the object cannot override the recognition cue. Although people often give recognition precedence over other knowledge, empirical tests indicate that they do not always neglect further information as strictly as predicted by the RH (e.g., Bröder & Eichler, 2006; Hilbig, Erdfelder, & Pohl, 2010; Newell & Shanks, 2004). Moreover, there are considerable individual differences in the use of the RH, with some people following it more consistently than others (Michalkiewicz & Erdfelder, 2016).

The performance that can be achieved by relying on the RH is domain specific and depends on the interplay with the environment, particularly the correlation between recognition and the criterion dimension (Gigerenzer, 2003; Goldstein & Gigerenzer, 2002); as with any other rule of thumb or strategy, it is thus important to consider in which situations the RH fares well and where it fails. The degree to which recognition predicts a criterion in a given environment can be quantified as *recognition validity*.[3] Ecological analyses suggest that name recognition is a valid cue in many domains (e.g., in identifying larger cities, more successful teams, stocks, or colleges). However, when more frequent object mentions (e.g., by people, books, and other media) do not uniquely

imply higher criterion values, the RH tends to fare poorly, for instance, when judging which of two diseases occurs more frequently (Pachur & Hertwig, 2006).

Adaptive Use of Recognition

Environment Adaptivity

Past and present research emphasizes the significance of adaptations to the environment during cognitive development (e.g., Bronfenbrenner, 1979). In the current context, using the RH adaptively means adjusting one's strategy use depending on whether the recognition validity in a given environment is high or low. We refer to this aspect as *environment adaptivity*. Adults are largely sensitive to the predictive power of recognition and rely considerably less on the RH in environments with low recognition validity than in environments with high recognition validity (Pachur et al., 2009, 2011). Although some developmental research suggests that children as young as 7 years adapt their information search strategies to the statistical structure of the environment to some extent and can identify useful cues (Nelson, Divjak, Gudmundsdottir, Martignon, & Meder, 2014; Ruggeri & Katsikopoulos, 2013; Ruggeri & Lombrozo, 2015), other studies have found that 10-year-olds still have pronounced difficulties in doing so (Betsch et al., 2014; Mata et al., 2011). It is thus an open question to what extent adaptive use of the RH can also be observed in children, and when adaptivity in using the RH emerges.

Discriminability in the Use of the RH

Adaptive selection of the RH may also include evaluation checks to selectively suspend its application when item-specific knowledge indicates that its use would lead to an incorrect decision (e.g., realizing that the town Woodstock is famous for lending its name to a music festival and is thus possibly recognized for a reason that has little to do with its size). There is evidence that such evaluation processes require additional time: RH-inconsistent decisions take longer than RH-consistent decisions and are associated with evaluative frontal brain activation (Volz et al., 2006). To quantify people's ability to discriminate between cases where the RH led to correct versus incorrect predictions on specific trials, the discriminability measure d' of signal-detection theory has been proposed (Pachur & Hertwig, 2006; see Footnote 8 for details). In an environment with low recognition validity, older adults had lower discriminability than younger adults did (Pachur et al., 2009). This age difference in discriminability was attributed to older adults' lower scores in tests of fluid cognitive abilities, suggesting that older adults, although having a well-developed base of knowledge and experience (Baltes, Staudinger, & Lindenberger, 1999), were constrained in flexibly retrieving it. It is

---

[3] The recognition validity α (Goldstein & Gigerenzer, 2002) can be calculated as $\alpha = C_{RU}/(C_{RU} + I_{RU})$, where C_{RU} and I_{RU} are the correct and incorrect inferences, respectively, in cases in which *one* of two objects in comparative judgments is recognized (*RU* cases). If $\alpha = 1$, following the recognition cue will always lead to correct inference; if $\alpha = .50$, recognition is uncorrelated with the criterion and following the recognition cue will lead to chance-level performance. As we shall explain below, we used a model-based measure (parameter a) that is formally equivalent to α to quantify the validity of recognition.

currently unclear how the ability to discriminate between cases where the RH would lead to correct or incorrect decisions develops across childhood, and which cognitive abilities underlie any differences between children and adolescents in this respect.

Research Questions of the Current Study

Despite the extensive literature on adults' use of the RH (e.g., Goldstein & Gigerenzer, 2002; Hilbig et al., 2010; Marewski, Pohl, & Vitouch, 2010; Newell & Shanks, 2004; Pachur et al., 2009, 2011; Pohl, 2006), hardly any research has been conducted about children's and adolescents' use of the RH (Pohl, von Massow, & Beckmann-Schumacher, 2016). In this study, we aim to fill this gap by addressing the following research questions. (a) To what extent does recognition knowledge of children and adolescents enable them to make accurate inferences about the world? (b) Do 9-year-olds already use the RH and, if so, (c) how adaptive are children and adolescents in selecting the RH? Finally, (d) how do children's verbal knowledge and fluid cognitive abilities affect their recognition-based inferences?

The Development of Recognition Validity

Developmental differences in recognition validity could arise because children and adolescents are exposed to different environments (e.g., they encounter different objects with different frequencies) or because younger children's recognition memory reflects the environment less accurately. Because recognition performance peaks in early adulthood (Schneider, 2015), and because recognition validity is a function of the accuracy of recognition memory (Pleskac, 2007), it may increase from childhood to adolescence. As outlined above, however, numerous studies on infants' habituation document their early and robust ability to remember and discriminate old from new stimuli (e.g., Fagan, 1970; Schneider, 2015). This may hold particularly for familiarity-based recognition, which is tapped by the RH (Pachur et al., 2009; for further discussion about recollection- vs. familiarity-based recognition processes, see, e.g., Rosburg, Mecklinger, & Frings, 2011; Schneider, 2015). It is thus conceivable that school-age children's recognition abilities will, to some extent, enable them to make accurate inferences in environments where recognition is a valid cue, but that recognition validity nevertheless increases with age.

Do School-Age Children Use the RH?

Recognition emerges rapidly and with little retrieval effort on the mental stage (e.g., Pachur & Hertwig, 2006; Rosburg et al., 2011). It could therefore be a particularly attractive cue for younger children, who have been shown to have difficulty applying more resource-demanding or complex inference strategies (e.g., Bereby-Meyer et al., 2004). We thus expected to find that school-age children already use the RH systematically. Moreover, based on the evidence on working memory development, it seems possible that children might rely more on the RH than younger and older adolescents do, implying a decreasing use as a function of age. On the contrary, previous research on children's use of heuristics that involve availability (Davies & White, 1994) or representativeness (Jacobs & Potenza, 1991) has shown that, al-

though children begin to apply such strategies from the age of 7 years, the use of these heuristics nevertheless increases with age. Moreover, the requirement to selectively attend to a single cue (and to ignore further information) could be more challenging for children than for adolescents (Mata et al., 2011). Thus, to the extent that children have difficulty systematically implementing a noncompensatory strategy, the use of the RH might increase with age.

Developmental Differences in the Adaptive Use of the RH

Because experience and knowledge about domains (e.g., about the usefulness of cues) is important for the adaptive use of the RH, and assuming that this knowledge increases during childhood (e.g., Bjorklund, 2011), environment adaptivity may be lower in children than in adolescents. Relatedly, because younger children have less developed metacognitive skills and possibly less insight into the task structure than adolescents (see Jacobs & Klaczynski, 2005), we expected that children would show lower ability than adolescents to discriminate between situations in which the RH leads to a correct versus incorrect inference. Note that these considerations imply environment-dependent (and possibly diverging) developmental trends in using the RH. Finally, this line of reasoning could suggest that age differences in discriminability between children and adolescents are explained by different types of cognitive abilities than the differences between younger and older adults. Pachur et al. (2009) reported that decrements in adaptive RH use in older age were mediated by fluid abilities—possibly reflecting deficits in older adults' ability to flexibly retrieve knowledge relevant for evaluating whether the RH arrives at a correct inference. In children, by contrast, such knowledge might be only partially available in the first place. In consequence, crystallized abilities might account more strongly for children's lower discriminability.

To test these research questions and hypotheses, we conducted a study in which children and adolescents completed inference tasks in which they judged cities and diseases in terms of their size and frequency of occurrence, respectively.

Method

Participants

The study involved 117 participants: 38 fourth-grade children (M = 9.29 years of age, SD = 0.65; 14 female) from a primary school, 38 younger adolescents (M = 12.42 years, SD = 1.15; 17 female) from a secondary school, and 41 older adolescents (M = 16.80 years, SD = 0.60; 33 female) from a high school. We chose these age groups based on previous developmental studies that reported significant changes in adaptive strategy selection within this range (Mata et al., 2011) and because, according to pilot testing, 9-year-olds were the youngest able to understand all task instructions. All participants were recruited and tested in schools in Livorno, Italy, and were native Italian speakers; participants had diverse socioeconomic backgrounds.

Materials

To develop the materials used in the inference task, we asked samples of 8- to 9-year-olds ($N = 19$), 11- to 12-year-olds ($N = 23$), and younger adults ($N = 19$) who did not participate in the main study to indicate which of 60 United States cities and 23 notifiable diseases they recognized. We aimed at using task environments compatible with those used in previous studies with younger and older adults (Horn, Pachur, & Mata, 2015; Pachur et al., 2009) in order to permit life span comparisons, but our main goal was to ensure that school-age children would be able to handle the materials selected. Based on the findings of this pilot study, 18 cities and 18 diseases were selected such that the proportion of recognized items was matched between the two environments (in terms of Ms and SDs), thereby making the applicability of the RH comparable across these two conditions. Information on the sizes of the cities was taken from a previous study (Pachur et al., 2009); information on the incidence rates of the diseases was obtained from official statistical databases maintained by the Italian Ministry of Health.[4] Table 1 provides further details about the stimuli.

Design

Participants in each age group were randomly assigned to either the cities or the diseases environment, resulting in a 2×3 (Environment $\times$ Age Group) between-subjects design. They were presented with two computer-based tasks. In the *inference task*, participants were asked to judge which of two cities had a larger population (cities environment) or which of two diseases occurred more frequently per year in Italy (diseases environment). The 18 items were exhaustively paired, resulting in 153 trials. In the *recognition task*, participants indicated which of the items (either 18 cities or 18 infectious diseases) they had heard of before the experiment. The order of the recognition and inference tasks was counterbalanced across participants within each age group.

Procedure

The tasks were administered to groups of six to 10 participants of the same age group. Participants first read an introductory text, and all instructions and explanations were also provided verbally by the experimenter. They then worked on either the inference task or the recognition task, depending on block order. In the inference task, the names of two cities (or diseases) appeared in 24-point sans-serif font, one on the left and one on the right side of the computer screen, surrounded by yellow and green rectangles, respectively. Participants indicated their inference by pressing a corresponding left or right key labeled with a yellow or green sticker, respectively. After a brief practice phase, 153 item pairs were presented in three blocks of 51 trials, with participants being given the opportunity to rest briefly between blocks to avoid fatigue. Each participant saw the item pairs in a different, random sequence. Moreover, one half of the participants (randomly determined) saw a specific item of a name pair on the left side of the screen; the other half saw this item on the right side. Participants were told that they would earn 1 point for each correct judgment and lose 1 point for each incorrect judgment, but feedback was provided only after the experiment. They were also told that the

best two participants (in each age group and condition) would receive book vouchers worth 50 EUR (~56 USD).

In the recognition task, the names of the 18 diseases (or cities) were presented sequentially and in random order. Participants indicated whether they had heard of the disease (or the city) before the experiment by pressing one of two keys (marked with red and blue stickers).

In a final task block, we measured participants' fluid cognitive abilities and verbal comprehension using pencil-and-paper versions of the digit-symbol substitution and vocabulary subtests of the Italian translation and adaptation of the Wechsler Intelligence Scale for Children (Wechsler, 1991). These tests scores served as indicators for participants' speed of information processing and their verbal item knowledge/concept formation, respectively.

The entire session lasted about 30 min. All experimental procedures were approved by the Ethics Committee of the Max Planck Institute for Human Development, Berlin, Germany, and we obtained informed consent from the teachers and parents of all participating children.

Results

An alpha level of .05 was chosen as criterion of significance for the statistical tests. We followed conventional classification in characterizing effect size estimates as "small" ($\eta_p^2 < .06$), "medium" ($\eta_p^2 \geq .06$), or "large" ($\eta_p^2 \geq .14$).

Proportion of Recognized Objects

How many objects did participants recognize? And were there differences between the age groups? To address these questions, we examined the proportion of recognized objects in a 2×3 (Environment $\times$ Age Group) analysis of variance (ANOVA). As shown in Table 2, 9-year-olds recognized around 45% of the cities and diseases, and there was an age-related increase in the proportion of recognized objects, as indicated by a main effect of age group, $F(2, 111) = 10.42$, $p < .01$, $\eta_p^2 = .16$. There was neither an interaction of age group with environment, $F < 1$, nor a significant effect of environment, $F(1, 111) = 2.38$, $p = .13$, $\eta_p^2 = .02$. Next, we examined the proportion of cases in the inference task in which the RH was applicable (i.e., the proportion of *RU* cases, in which one of two objects was recognized). As intended, there were no differences between the environments, $F(1, 111) = 1.05$, $p = .31$, $\eta_p^2 = .01$. There was a marginally significant age-related increase, $F(2, 111) = 3.05$, $p = .05$, $\eta_p^2 = .05$, and this increase did not differ between the two environments ($F < 1$ for the interaction between age group and environment).

Performance in the Inference Task

Table 2 reports the proportion of accurate judgments in the inference task, separately for the three age groups. As indicated by a main effect of age, inferential accuracy increased from children to older adolescents, $F(2, 111) = 15.36$, $p < .01$, $\eta_p^2 = .22$. Further, accuracy was higher in the cities than in the diseases environment, $F(1, 111) = 69.03$, $p < .01$, $\eta_p^2 = .38$. These factors did not interact ($F < 1$). Moreover, an analysis of participants'

---

[4] http://www.salute.gov.it/

Table 1

Stimulus Materials Used In The Inference Task and Recognition Rates as Function of Age Group

| | Cities in the United States | | | | | Infectious diseases | | | | |
| | | Recognized by % of participants | | | | | Recognized by % of participants | | | |
Object	Criterion[a]	Children	Younger adolescents	Older adolescents	Object[b]	Criterion[c]	Children	Younger adolescents	Older adolescents
Chicago	2,934,900	84	78	100	Viral hepatitis	7,102	47	55	90
Philadelphia	1,522,900	68	83	100	Tuberculosis	4,695	37	90	100
Dallas	1,234,200	63	78	75	Brucellosis	1,315	42	40	10
San Antonio	1,188,500	79	83	60	Malaria	843	89	85	100
Detroit	960,300	42	67	75	Meningitis and encephalitis	747	42	95	90
Jacksonville	761,900	26	33	55	Syphilis	470	26	15	67
Boston	597,700	84	78	100	Gonorrhea	336	26	30	33
Washington	577,800	68	72	100	Tetanus	96	53	60	95
Portland	542,300	42	56	65	Leptospirosis	53	37	30	14
New Orleans	488,600	58	72	100	Botulism	33	26	15	38
Norfolk	241,000	26	33	5	Trichinosis	12	21	20	5
Rochester	222,400	21	33	55	Tularemia	8	32	20	10
Akron	218,100	32	28	5	Cholera	2	47	90	95
Madison	211,000	26	50	50	Leprosy	1	74	75	95
Montgomery	203,800	32	39	55	Typhoid	0	37	60	100
Richmond	203,400	32	50	75	Poliomyelitis	0	37	35	29
Shreveport	201,800	16	17	10	Hemorrhagic fever	0	79	65	71
Paradise	201,500	26	61	25	Diphtheria	0	16	10	10

[a] City population.   [b] Target objects translated from Italian (language of presentation during study).   [c] Annual incidence rate.

response times (RTs) with a $2 \times 3 \times 2$ (Environment $\times$ Age Group $\times$ Choice) ANOVA indicated that inferences became faster with increasing age, $F(2, 111) = 3.83$, $p = .02$, $\eta_p^2 = .06$ (see Table 2, which reports RTs separately for choices of recognized and unrecognized objects).[5] RTs did not differ significantly between the two environments, and this held irrespective of age group (indicated by a nonsignificant interaction between environment and age group, $F < 1$). Notably, participants' inferences in *RU* cases were faster when the recognized object was chosen than when the unrecognized object was chosen, $F(1, 111) = 20.24$, $p < .01$, $\eta_p^2 = .15$. This extends previous findings with adults (e.g., Volz et al., 2006) and indicates an analogous choice effect on children's and adolescents' RTs. The effect did not interact with age group (all $Fs < 1.39$, $ps > .25$, for interactions involving age group), but was more pronounced in the cities than in the diseases environment, as indicated by an Environment $\times$ Choice interaction, $F(1, 111) = 6.47$, $p = .01$, $\eta_p^2 = .06$.

Formal Modeling Analysis

We used a cognitive-modeling approach to decompose observed responses in the inference task. Specifically, we adopted a multinomial processing tree (MPT) model to measure use of the RH and the validity of recognition and of further knowledge. In the inference task, the choice of a recognized object may result from reliance on the RH, but also from guessing, reliance on another strategy, or other knowledge/cues correlated with the criterion. Consequently, the proportion of inferences in which people choose the recognized object (the adherence rate) is a confounded measure of RH use (Hilbig et al., 2010). A multinomial analysis allows us to address this issue by disentangling pure reliance on recognition (as assumed by the RH) from the use of further knowledge (or any other strategy). Moreover, MPT modeling provides a well-

developed statistical machinery for model comparison and goodness-of-fit tests (see Batchelder & Riefer, 1999; Erdfelder et al., 2009, for reviews). The *r*-model used in this study provides measures for the probability of using the RH (parameter *r*), for the cue validity of recognition (parameter *a*), and of further knowledge (parameter *b*). Further details are described in Appendix A (see also Hilbig et al., 2010; Horn et al., 2015).

The parameter estimates in MPT models are often based on the pooled data from a group of participants. However, ignoring diversity between individuals may lead to severe problems when drawing conclusions about strategy use (Siegler, 1987) or when populations are likely to be heterogeneous, as in developmental or clinical settings (Arnold, Bayen, & Böhm, 2015; Batchelder & Riefer, 1999). To address these issues, we used a hierarchical model implementation that has been successfully applied in previous research (Horn et al., 2015) and that accounts for differences and similarities across individuals by specifying an overarching group-level distribution. An advantage of this approach is that the group distribution constrains individual-level estimates in a theoretically principled way, which promises to yield more reliable individual parameters (Gelman, Carlin, Stern, & Rubin, 2004). Here, we used Smith and Batchelder's (2010) beta-MPT implementation that assumes that the individual-level estimates for each parameter of an MPT model (e.g., parameters *a*, *b*, *g*, *r*, in the *r*-model) stem from independent group-level beta distributions, where the variance of the distributions captures the diversity

---

[5] The reaction time (RT) analyses were based on the individual medians of untrimmed RTs; various further methods for dealing with extreme RTs would have led to the same conclusions. We focus on the RTs in *RU* cases, in which the RH was applicable (i.e., in which only one object in the pair comparison was recognized).

between individuals (for related model developments see, e.g., Matzke, Dolan, Batchelder, & Wagenmakers, 2015).

Parameter estimation relied on a Bayesian approach. We used the Markov chain Monte Carlo (MCMC) methodology for posterior sampling[6] to determine the most probable value ranges of the parameters in the posterior distributions, given the data (see Lee & Wagenmakers, 2013, for an overview). In the following, we focus on group-level differences of the means between age groups and environments.[7]

Validity of recognition and of further knowledge. We first consider the validity of recognition, as measured by the MPT model parameter a. Figure 1a shows the group means and the distribution of the individual estimates of this parameter separately for the three age groups and two environments. As expected per our design, the posterior difference in parameter a, $\Delta\mu_a = \mu_{a,\,\text{cities}} - \mu_{a,\,\text{diseases}}$, indicated substantially higher recognition validity in the cities than diseases environment, and this held for all age groups [that is, all $Ps\,(\Delta\mu_a > 0|D) > .95$]. Importantly, for cities, recognition validity did not differ between any of the age

Table 2

Performance in the Inference Task as a Function of Age Group and Task Environment

	Ms		SDs	
Measure and age group	Cities	Diseases	Cities	Diseases
Proportion of recognized objects (recognition test)				
Children	.46	.43	.21	.18
Younger adolescents	.56	.49	.13	.15
Older adolescents	.62	.58	.13	.09
Proportion of *RU* cases in the inference task				
Children	.43	.45	.12	.10
Younger adolescents	.49	.48	.06	.05
Older adolescents	.47	.50	.08	.05
Proportion of accurate judgments in the inference task				
Children	.60	.52	.08	.05
Younger adolescents	.62	.51	.09	.06
Older adolescents	.68	.58	.05	.05
RT$_{RU}$ (choice of recognized object)				
Children	1,984	2,332	677	1,138
Younger adolescents	1,608	1,812	642	644
Older adolescents	1,258	1,591	181	490
RT$_{RU}$ (choice of unrecognized object)				
Children	2,346	2,417	1,249	1,059
Younger adolescents	2,366	2,030	2,301	750
Older adolescents	2,214	1,862	977	711
Discriminability d'_{RH}				
Children	.14	.04	.55	.37
Younger adolescents	.36	−.01	.53	.38
Older adolescents	.78	.44	.61	.50
Proportion of choices of recognized object (recognition adherence)				
Children	.70	.63	.21	.18
Younger adolescents	.67	.65	.19	.18
Older adolescents	.86	.58	.08	.15

Note. RT = response time in milliseconds; for RTs, the group *M*s and *SD*s are based on individuals' median RT in a condition; *RU* cases = trials in the inference task in which one of the two objects was recognized; RH = recognition heuristic.
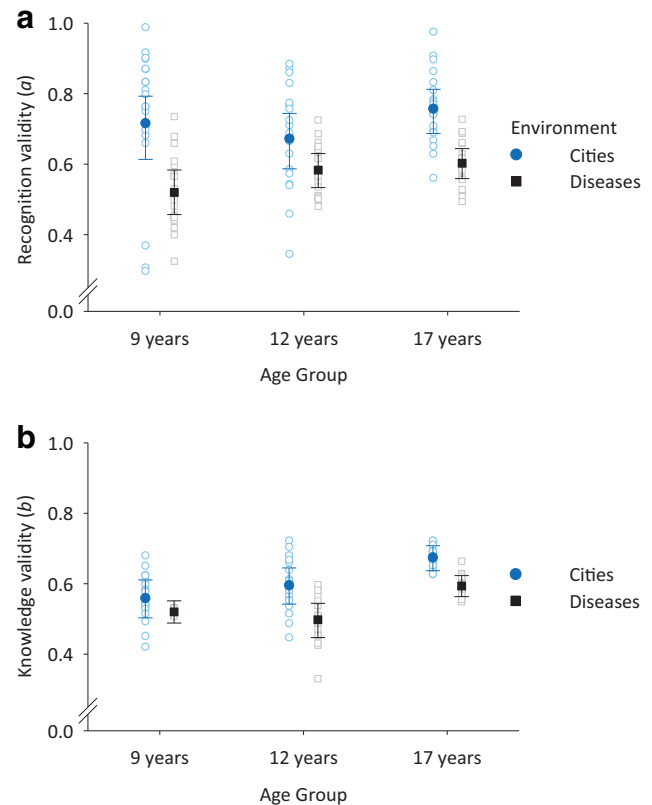
Figure 1. The figures show posterior estimates of the group-level means (filled symbols) and the individual-level parameters (nonfilled symbols). Error bars represent 95% credible intervals for the group means. (a) Predictive power of recognition (recognition cue validity as measured with multinomial processing tree [MPT] model parameter a) as a function of age group and environment. (b) Validity of further knowledge (MPT model parameter b) as a function of age group and environment. See the online article for the color version of this figure.

groups and was similarly high for the 9-year-olds as for the 17-year-olds, $P\,(\Delta\mu_a > 0|D) \approx .22$; 95% CI $[-0.16, 0.06]$, for this difference. In other words, recognition was already a (relatively) useful cue for the youngest age group in this environment, with a validity far above chance level, $P\,(\mu_a > 0.5|D) > .99$. In the diseases environment, in which recognition validity was gen-

---

[6] The method requires specification of prior distributions for the parameters. Following Smith and Batchelder (2010), we used uniform ("vague") priors for the group-level parameters that are easily overruled by data and thus should not have substantial impact on statistical inference. The resulting posterior distributions represent the belief about the parameters after having seen the data. For parameter estimation, we report the medians of the MCMC chains and 95% Bayesian confidence/credible intervals (CIs). A graphical representation of the hierarchical MPT model, further details about convergence of the MCMC sampling, and model fit are provided in the online supplemental materials.

[7] To conduct contrasts between environments and age groups for any of the MPT r-model parameters $\theta = (a, b, g, r)$ we calculated the difference of the group-level means, $\Delta\mu_\theta = \mu_{\theta,1} - \mu_{\theta,2}$, from the posterior distributions. For any difference of interest in a parameter θ, we then calculated the mass of the posterior distribution above zero, given the data D, to quantify the evidence in favor of an effect: $P\,(\Delta\mu_\theta > 0|D)$.

erally low, validity did not differ between 12- and 17-year-olds, 95% CI [−0.08, 0.04], but increased from the youngest to the oldest age group, $P(\Delta\mu_a > 0|D) > .95$; the patterns of recognition knowledge (see Table 1) suggest that this finding might be due to the youngest age group not yet having heard of the more frequent diseases (e.g., hepatitis, tuberculosis).

The model parameter b, expressing the validity of further knowledge (or of any other information beyond recognition), tended to be higher for cities than for diseases (Figure 1b), suggesting that participants had somewhat better knowledge about the cities than about the diseases [for younger and older adolescents, $Ps (\Delta\mu_b > 0|D) > .95$; for children, $P(\Delta\mu_b > 0|D) \approx .89$]. Within both the cities and the diseases environment, the b parameter increased from the youngest to the oldest age group, and the 17-year-olds had more valid knowledge than either the 9- or the 12- year-olds, $Ps (\Delta\mu_b > 0|D) > .95$.

Use of the RH and environment adaptivity. Is there evidence that the school-age children in our study already applied the RH? And to what extent did the three age groups adjust their reliance on the RH between environments with high and low recognition validities? Figure 2 shows the group means and the distributions of the individual estimates of the r parameter. As can be seen, the means clearly exceeded zero for 9- and 12-year-olds [all $Ps (\mu_r > 0|D) > .99$], indicating a systematic use of the RH. This also indicates that even in the diseases environment (where the younger age groups' accuracy was around chance level), their responses did not merely reflect guessing behavior. Rather, responses resulted—at least to some extent—from the systematic use of a judgment strategy (which, however, rested on a cue that had low validity in this environment). The finding that 9-year-olds already relied on the RH is consistent with the idea that the basic cognitive abilities required to apply this simple heuristic (i.e., discrimination between novel and familiar objects) are developed from an early age. Crucially, however, in both younger age groups, there was no difference in reliance on the RH between the high- and low-recognition validity environments: Analyses of the posterior difference in the r parameter between environments, $\Delta\mu_r = \mu_{r, cities} - \mu_{r, diseases}$, indicated no effect of environment on RH use in children, $P(\Delta\mu_r > 0|D) \approx .89$, or in younger adolescents, $P(\Delta\mu_r > 0|D) \approx .49$. In contrast, there was a large

effect of environment in older adolescents, $P(\Delta\mu_r > 0|D) > .99$. That is, older adolescents used the RH substantially more often in the environment with high than low recognition validity, and their RH use even fell below that of the younger age groups when validity was low [$P(\Delta\mu_r > 0|D) < .05$ and $P(\Delta\mu_r > 0|D) \approx .07$, for the comparison with 12- and 9-year-olds, respectively]. These findings suggest that only the older adolescents— but neither children nor younger adolescents—adaptively adjusted their reliance on the RH between environments with different recognition validities. Finally, we note that the younger age groups were highly diverse in their strategy use (underlining the relevance of modeling heterogeneity), whereas older adolescents were more homogeneous in either using or suspending the RH, depending on the task environment.

Discriminability in RH Use

To evaluate participants' ability to discriminate between cases in which recognition leads to a correct versus an incorrect inference, we used the d'_{RH} index mentioned above.[8] Higher d'_{RH} scores imply increased reliance on the RH when it leads to a correct inference and decreased reliance when it leads to an incorrect inference. Overall, discriminability (see Table 2) increased across age groups, $F(2, 111) = 12.60, p < .01, \eta_p^2 = .19$, and was higher in the cities than the diseases environment, $F(1, 111) = 8.48, p < .01, \eta_p^2 = .07$; these two factors did not interact ($F < 1$).

Relation to Measures of Cognitive Speed and Verbal Knowledge

A previous study reported that adult age differences in discriminability in RH use (d'_{RH}) were mediated through measures of cognitive speed (Pachur et al., 2009). Here, we explored the same relations at the beginning of the life span. As argued above, however, findings of lower discriminability in children than in adolescents might instead result from differences in their knowledge base. Overall, test scores for both verbal knowledge and cognitive speed increased with age (Table 3).[9] Does performance in either or both of these measures account for developmental
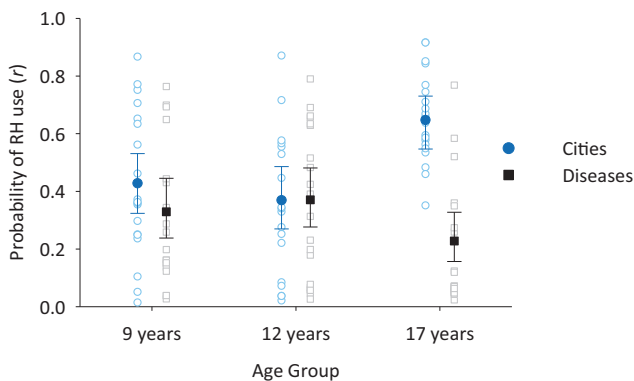
Figure 2. Probability of using the recognition heuristic (model parameter r) as a function of age group and environment. See the online article for the color version of this figure.

---

[8] The discriminability measure d'_{RH} is derived from an equal-variance signal detection model and uses each participant's hit rate (HR) and false-alarm rate (FA) on one-recognized (RU) trials (Pachur et al., 2009). In this context, HR is defined as the proportion of RU cases in which the decision maker chooses the recognized object when it scores higher on the criterion, thus leading to a correct inference; FA is defined as the proportion of RU cases in which the decision maker chooses the recognized object when it scores lower on the criterion, thus leading to an incorrect inference; d'_{RH} is then calculated for each participant as $d'_{RH} = z(HR) - z(FA)$, where $z(HR)$ and $z(FA)$ are the corresponding z values that cut off areas of HR and FA, respectively, from the standard normal distribution. Note that significant correlations between d'_{RH} and model parameter b (validity of further knowledge) can be expected, as both measures capture processes beyond exclusive use of the RH (e.g., Horn et al., 2015). The online supplemental materials include tables of intercorrelations among the measures used in this study.

[9] Main effects of age emerged for all cognitive measures; we did not plan to test specific contrasts between age groups. We notice, however, that the verbal test scores between 9- and 12-year-olds did not differ significantly, which was unexpected. We are therefore cautious in basing conclusions on specific comparisons between these two age groups with that measure.

Table 3

Participants' Characteristics and Individual Difference Measures

Variable	Children		Younger adolescents		Older adolescents		Statistical test[a]	
	M	SD	M	SD	M	SD	F	R^2
Age (years)	9.29	.65	12.42	1.15	16.80	.60		
Verbal knowledge[b]	28.87	9.06	26.74	9.30	40.98	8.80	28.74	.34*
Cognitive speed[b]	37.55	9.48	54.41	15.26	62.11	11.27	40.15	.42*

[a] Main effect of age group.   [b] Vocabulary and digit-symbol substitution subtests from the Italian adaptation of the Wechsler Intelligence Scale for Children (Wechsler, 1991).
* $p < .001$.

differences in d'_{RH} in the present study? To test this possibility, we conducted a series of regression and mediation analyses (e.g., Hayes, 2013) predicting participants' discriminability. As shown in Table 4, an indirect effect of age through verbal knowledge on d'_{RH} emerged ($z = 2.44$, $p < .05$). Moreover, verbal knowledge scores accounted for age differences in inference accuracy ($z = 2.96$, $p < .01$) and in the validity of further information (model parameter *b*), $z = 2.90$, $p < .01$. These results are in line with the notion that item-specific knowledge helps adolescents to adaptively suspend the RH on particular trials. In contrast, cognitive speed scores did not mediate effects of age on any of these variables (all zs < 0.70; ps $> .48$). Further details of these path analyses are in given Appendix B.

Notably, we also observed an association between verbal knowledge scores and use of the RH (parameter *r*), but the mediating role of knowledge differed significantly between task environments. That is, we found conditional indirect effects of age through verbal knowledge—but not through speed—on use of the RH when task environment was included as moderator in the mediation analyses. The effect was positive for cities (more knowledge implying increased RH use) but negative for diseases (more knowledge implying reduced RH use), in line with the assumption that knowledge is associated with adaptive (task-dependent) selection of the RH (Table B2). We note, however, that our analyses with concurrent data may differ from longitudinal mediations (e.g., Lindenberger, von Oertzen, Ghisletta, & Hertzog, 2011) and serve to test necessary but not sufficient conditions for mediation mechanisms. We thus refrain from any causal interpretation of processes developing over time.

Discussion

It has previously been shown that recognition information—that is, knowledge about whether or not an object has been encountered before—can be an informative cue for making inferences about properties of that object. In this study, we investigated to what extent school-age children and adolescents can and do make use of the RH, which relies on real-world recognition knowledge. Further, as the success of a heuristic strongly hinges on its match with the environment, we investigated developmental trends in the adaptive selection of the RH when cue validities differed between domains. The main findings can be summarized as follows: First, primary schoolchildren and younger adolescents already possessed valid recognition knowledge and relied on the RH to some extent, as indicated by a formal measurement model. Second, however, environment adaptivity of RH use was observed only in the older

adolescents, who *increased* their reliance on the RH substantially in an environment with high recognition validity, but adaptively *reduced* their reliance when recognition validity was low. Third, these developmental differences in environment adaptivity were accompanied by differences in having valid further knowledge and in discriminability (d'_{RH}) between correct and incorrect inferences (including the ability to suspend the use of the RH for a specific judgment). Specifically, older adolescents were better able to discriminate between trials in which the RH led to a correct or an incorrect inference than were younger adolescents or children. As a corollary of these differences in adaptive RH use, older adolescents made more accurate inferences in both environments than the younger age groups did. Finally, we obtained some evidence that developmental differences in discriminability (d'_{RH}) were mediated through participants' verbal knowledge. In consequence, younger children's lower adaptivity could be partially attributed to a lack of item-specific knowledge to discriminate between situations in which the RH arrives at a correct versus an incorrect inference. In the following, we discuss further implications of these results.

Developmental Differences in Adaptive Strategy Selection

Given that a central topic of our investigation was the development of adaptive RH use—which presupposes information about the cue validity of recognition—the question arises how such information might be acquired. One possibility is that both the size and direction of the correlation between recognition and a criterion are learned from experience (e.g., Gigerenzer, 2003).[10] A stream of research indicates that individuals can encode event frequencies and co-occurrences rather accurately and with little effort, implying that some ingredients of cue learning may have a robust cognitive foundation (Sedlmeier & Betsch, 2002, provide an overview). Moreover, in laboratory studies, the learning of cue validities—including recognition—is often successfully implemented as experiential learning, given sufficient feedback and possibilities for active exploration (Newell, Lagnado, & Shanks, 2007; Newell & Shanks, 2004; see also Klayman, 1985). Notably, however, it has been emphasized that in realistic situations outside the laboratory, learning of cue validities could be very taxing, as individuals might need to monitor a vast number of associations and

---

[10] Apart from individual learning, adaptive use of recognition could also result from social and evolutionary learning and may rely more on approximate discrimination between contexts rather than exact computation of cue values.

Table 4

Hierarchical Linear Regressions With Discriminability in RH Use and Accuracy of Inference as Dependent Variables

Predictor variables	Discriminability d'_{RH}				Accuracy of inference			
	R^2	ΔR^2	F^a	Indirect effect $a \times b$	R^2	ΔR^2	F^a	Indirect effect $a \times b$
Step 1								
Age	.15		19.39***		.11		13.76***	
Verbal knowledge	.16		21.47***		.15		20.23***	
Speed	.07		8.71**		.07		8.54**	
Step 2								
Verbal Knowledge + Age	.22	.05	7.34**	**.13 [.03, .25]**	.18	.03	4.09*	**.14 [.05, .24]**
Speed + Age	.15	.08	9.86**	.01 [−.14, .15]	.11	.04	5.32*	.06 [−.11, .22]

Note. In Step 2, ΔR^2 and corresponding F values refer to the difference in explained variance between the model with two predictor variables and the model with only the cognitive test measure as predictor in Step 1. These comparisons indicate whether Age adds to the explained variance when each cognitive measure has been considered. In addition, bootstrap mediation analyses (e.g., Hayes, 2013; with $N = 50,000$ resamples) were conducted for the models in Step 2, where $a \times b$ represents the standardized mean path weight of the indirect effect of age (through knowledge or speed, respectively, as mediators) on the dependent variable; 95% confidence intervals are in brackets and indicate significant mediation if the interval does not include the value zero (shown in boldface). See Appendix B for further details. RH = recognition heuristic; d'_{RH} = discriminability between correct and incorrect predictions of the RH.
[a] dfs are $F(1, 111)$ in Step 1 and $F(1, 110)$ in Step 2; four participants were excluded from all regression analyses because their scores on the digit-symbol substitution test were unavailable.
* $p < .05$. ** $p < .01$. *** $p < .001$.

feedback could be lacking or unreliable (Dougherty, Franco-Watkins, & Thomas, 2008). Individuals may thus have only fuzzier intuitions about the validity of recognition in different environments and adjust their strategy use accordingly (Pachur, 2011); even though these intuitions may not be perfect, they could still capture relative differences in recognition validities between environments in a surprisingly robust fashion (Wright & Murphy, 1984). Importantly, adaptivity to different domains could also arise through insight into the structure of a task (Newell et al., 2007) or beliefs and subjective theories about these domains. According to Klaczynski (2005), the evaluation (and justification) of such beliefs requires metacognitive abilities of monitoring and controlling information processing that change across development.

One interesting implication of our findings is that the use of simple heuristics does not progress unidirectionally across childhood (Jacobs & Klaczynski, 2005), but may instead depend on features of the task environment and on the decision maker's sensitivity to those features. Specifically, the results indicate that younger children do not generally resort to the simple RH due to limitations in their information processing skills and cognitive resources (Bereby-Meyer et al., 2004; see also Payne, Bettman, & Johnson, 1993). In fact, when recognition validity was high, the 9- and 12- year-olds applied the RH less frequently than older adolescents did. Together with the observation that their RH use did not differ between environments, this finding indicates that younger children may have difficulties in focusing selectively on a single relevant cue (i.e., recognition). This is in line with developmental research on the use of other cue-based strategies (Mata et al., 2011; see also Betsch et al., 2014, Betsch, Lehmann, Lindow, Lang, & Schoemann, 2016), which indicates that the adaptive selection of simple heuristic strategies (such as the RH) appears to be harder for children than initially assumed—even if the required cognitive ingredients or building blocks needed to carry out a strategy's subcomponents are developed relatively early.

The Role of Recognition Knowledge

Application of the RH requires some level of ignorance (i.e., individuals who recognize all or none of the objects cannot apply it). In this study, developmental differences in the amount of recognition knowledge had little impact on the applicability of the RH because the proportion of recognized cities and diseases varied only moderately between age groups (from 43% to 62%). Consequently, opportunities (proportion of cases) where the RH could be applied were relatively similar across age groups.

Could age differences in recognition validity account for the differential use of the RH observed between the groups? In principle, developmental differences in the predictive power of recognition knowledge could emerge because children and adolescents dwell in environments with different statistical structures, leading to different patterns of recognized and unrecognized objects. Second, validity differences could emerge if younger children's recognition memory reflects the environment less accurately than that of adolescents (Pleskac, 2007). Finally, because response biases appear to decline from early childhood to young adulthood (Reyna & Kiernan, 1994), younger children could have a stronger tendency to claim to recognize objects that they have never encountered before ("false alarms"), thereby inflating estimates in the recognition test (and thus decreasing the recognition validity). Notably, our results suggest that these possibilities had little impact in the present study because the recognition validities were rather similar across age groups in the cities environment. Therefore, the younger age groups' infrequent use of the RH in this environment does not reflect an adaptive adjustment to potentially less-accurate memories.

Outlook and Limitations

It is possible that the task domains we selected (to permit life span comparisons with data from previous studies with older and younger adults; Horn et al., 2015; Pachur et al., 2009) were relatively unfamiliar to the youngest children (whose inference

accuracy in one of the conditions, the diseases environment, was around chance). It has been emphasized, however, that the RH exploits patterns of ignorance (e.g., Goldstein & Gigerenzer, 2002) and may play out its strength in uncertain environments, where knowledge is scarce. It is thus an interesting question whether children would make more accurate inferences in domains in which they have more expertise (e.g., animals, comic characters, etc.) or whether this knowledge would ironically prevent them from using simpler strategies that can fare surprisingly well (Goldstein & Gigerenzer, 2002, p. 79, and Pachur, 2010, provide further discussion of such "less-can-be-more" effects).

In this study, we focused on children's strategic use of recognition, with a particular interest in adaptivity of RH use. The results suggest that school-age children's RH use, though already present, was not very pronounced. Hence, an interesting question is what other strategies children might have employed whenever they did *not* apply the RH. The model estimates suggest that the validity of further information available on those occasions was relatively low. One possibility, therefore, is that children resorted to guessing because further knowledge was scarce (particularly in the diseases environment). Another possibility is that children were influenced by further information associated with the objects that was not very helpful (or even irrelevant) and that might have been automatically activated (Khader, Pachur, & Jost, 2013). Finally, children may have used various idiosyncratic strategies, such as reliance on specific visual, phonemic, or orthographic features of the presented items. The employed MPT-modeling approach does not allow us to disentangle guessing, errors in strategy execution, and the use of alternative knowledge-based strategies. Further research could address and dissociate these possibilities through experimental extensions of the comparative-judgment paradigm and further modeling developments.

Our experimental design followed an ecological approach to investigate how children's and adolescents' natural knowledge acquisition (in the environments to which they are exposed) equip them with the ability to make recognition-based inferences adaptively. We thus selected names of real-world objects as stimuli and relied on participants' recognition knowledge acquired outside the lab. A raison d'être for this approach is that the RH is assumed to exploit regularities in exposure frequencies to objects in natural environments (Goldstein & Gigerenzer, 2002) and there is some evidence that individuals might treat recognition induced in the lab differently from naturally acquired recognition when making inferences (Pachur et al., 2011). Nonetheless, an important limitation of this approach is the lack of experimental control over stimuli (e.g., further knowledge about items) and cue learning. Therefore, testing to what extent our conclusions generalize to situations in which recognition and further knowledge are induced experimentally in the lab (Bröder & Eichler, 2006; Newell & Shanks, 2004) would provide an important complementary contribution.

Finally, given our cross-sectional design, it remains an open question whether the adaptive use of the RH follows a discontinuous staircase pattern across development (as implied by traditional models; e.g., Piaget & Inhelder, 1951), or a gradually changing pattern, possibly resulting from a repertoire of different rules or algorithms that are at a child's disposal at any given time (Siegler, 1999). In light of the high diversity in strategy use that we

observed in school-age children, longitudinal analyses with frequent measurement time points could provide an exciting avenue for future research.

Conclusion

Making good decisions in an uncertain world requires the adaptive selection of strategies that are attuned to the structure of the environment. We examined developmental differences in adaptive strategy selection, focusing on a prime example of an ecologically rational inference strategy: the RH. Our results are in line with the assumption that core cognitive abilities required to implement this strategy are developed relatively early; we observed that 9- and 12-year-olds make systematic use of the RH. Crucially, however, only older adolescents could use the RH adaptively across different task environments. Our findings thus highlight the importance of examining the interplay between cognition and environment to gain insight into the development of decision making.

References

Arnold, N. R., Bayen, U. J., & Böhm, M. F. (2015). Is prospective memory related to depression and anxiety? A hierarchical MPT modelling approach. *Memory, 23,* 1215–1228. http://dx.doi.org/10.1080/09658211.2014.969276

Baltes, P. B., Staudinger, U. M., & Lindenberger, U. (1999). Lifespan psychology: Theory and application to intellectual functioning. *Annual Review of Psychology, 50,* 471–507. http://dx.doi.org/10.1146/annurev.psych.50.1.471

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6,* 57–86. http://dx.doi.org/10.3758/BF03210812

Bereby-Meyer, Y., Assor, A., & Katz, I. (2004). Children's choice of strategies: The effects of age and task demands. *Cognitive Development, 19,* 127–146. http://dx.doi.org/10.1016/j.cogdev.2003.11.003

Betsch, T., & Lang, A. (2013). Utilization of probabilistic cues in the presence of irrelevant information: A comparison of risky choice in children and adults. *Journal of Experimental Child Psychology, 115,* 108–125. http://dx.doi.org/10.1016/j.jecp.2012.11.003

Betsch, T., Lang, A., Lehmann, A., & Axmann, J. M. (2014). Utilizing probabilities as decision weights in closed and open information boards: A comparison of children and adults. *Acta Psychologica, 153,* 74–86. http://dx.doi.org/10.1016/j.actpsy.2014.09.008

Betsch, T., Lehmann, A., Lindow, S., Lang, A., & Schoemann, M. (2016). Lost in search: (Mal-)adaptation to probabilistic decision environments in children and adults. *Developmental Psychology, 52,* 311–325. http://dx.doi.org/10.1037/dev0000077

Bjorklund, D. F. (1987). How age changes in knowledge base contribute to the development of children's memory: An interpretive review. *Developmental Review, 7,* 93–130. http://dx.doi.org/10.1016/0273-2297(87)90007-4

Bjorklund, D. F. (2011). *Children's thinking: Cognitive development and individual differences.* Belmont, CA: Wadsworth.

Bonawitz, E., Denison, S., Gopnik, A., & Griffiths, T. L. (2014). Win-stay, lose-sample: A simple sequential algorithm for approximating Bayesian inference. *Cognitive Psychology, 74,* 35–65. http://dx.doi.org/10.1016/j.cogpsych.2014.06.003

Bröder, A., & Eichler, A. (2006). The use of recognition information and additional cues in inferences from memory. *Acta Psychologica, 121,* 275–284. http://dx.doi.org/10.1016/j.actpsy.2005.07.001

Bröder, A., & Schiffer, S. (2006). Stimulus format and working memory in fast and frugal strategy selection. *Journal of Behavioral Decision Making, 19,* 361–380. http://dx.doi.org/10.1002/bdm.533

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review, 50,* 255–272. http://dx.doi.org/10.1037/h0060889

Davidson, D. (1991a). Children's decision-making examined with an information-board procedure. *Cognitive Development, 6,* 77–90. http://dx.doi.org/10.1016/0885-2014(91)90007-Z

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

Davies, M., & White, P. A. (1994). Use of the availability heuristic by children. *British Journal of Developmental Psychology, 12,* 503–505. http://dx.doi.org/10.1111/j.2044-835X.1994.tb00651.x

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review, 12,* 45–75. http://dx.doi.org/10.1016/0273-2297(92)90003-K

Denison, S., & Xu, F. (2014). The origins of probabilistic inference in human infants. *Cognition, 130,* 335–347. http://dx.doi.org/10.1016/j.cognition.2013.12.001

Dougherty, M. R., Franco-Watkins, A. M., & Thomas, R. (2008). Psychological plausibility of the theory of probabilistic mental models and the fast and frugal heuristics. *Psychological Review, 115,* 199–211. http://dx.doi.org/10.1037/0033-295X.115.1.199

Erdfelder, E., Auer, T. S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models: A review of the literature. *Zeitschrift für Psychologie/Journal of Psychology, 217,* 108–124. http://dx.doi.org/10.1027/0044-3409.217.3.108

Fagan, J. F., III. (1970). Memory in the infant. *Journal of Experimental Child Psychology, 9,* 217–226. http://dx.doi.org/10.1016/0022-0965(70)90087-1

Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology, 54,* 1–34. http://dx.doi.org/10.1016/S0301-0511(00)00051-X

Gambetta, D., & Hamill, H. (2005). *Streetwise: How taxi drivers establish customers' trustworthiness*. New York, NY: Russell Sage Foundation.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis*. Boca Raton, FL: Chapman & Hall.

Gigerenzer, G. (2003). The adaptive toolbox and life span development: Common questions? In U. M. Staudinger & U. Lindenberger (Eds.), *Understanding human development: Dialogues with lifespan psychology* (pp. 423–435). Boston, MA: Kluwer. http://dx.doi.org/10.1007/978-1-4615-0357-6_20

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109,* 75–90. http://dx.doi.org/10.1037/0033-295X.109.1.75

Gregan-Paxton, J., & John, D. R. (1997). The emergence of adaptive decision making in children. *Journal of Consumer Research, 24,* 43–56. http://dx.doi.org/10.1086/209492

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press.

Hilbig, B. E., Erdfelder, E., & Pohl, R. F. (2010). One-reason decision making unveiled: A measurement model of the recognition heuristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 123–134. http://dx.doi.org/10.1037/a0017518

Hilbig, B. E., Erdfelder, E., & Pohl, R. F. (2012). A matter of time: Antecedents of one-reason decision making based on recognition. *Acta Psychologica, 141,* 9–16. http://dx.doi.org/10.1016/j.actpsy.2012.05.006

Hoffrage, U., Weber, A., Hertwig, R., & Chase, V. M. (2003). How to keep children safe in traffic: Find the daredevils early. *Journal of Experimental Psychology: Applied, 9,* 249–260. http://dx.doi.org/10.1037/1076-898X.9.4.249

Horn, S. S., Pachur, T., & Mata, R. (2015). How does aging affect recognition-based inference? A hierarchical Bayesian modeling approach. *Acta Psychologica, 154,* 77–85. http://dx.doi.org/10.1016/j.actpsy.2014.11.001

Jacobs, J. E., & Klaczynski, P. A. (Eds.). (2005). *The development of judgment and decision making in children and adolescents*. Mahwah, NJ: Erlbaum.

Jacobs, J. E., & Potenza, M. (1991). The use of judgement heuristics to make social and object decisions: A developmental perspective. *Child Development, 62,* 166–178.

Khader, P. H., Pachur, T., & Jost, K. (2013). Automatic activation of attribute knowledge in heuristic inference from memory. *Psychonomic Bulletin & Review, 20,* 372–377. http://dx.doi.org/10.3758/s13423-012-0334-7

Klaczynski, P. A. (2005). Metacognition and cognitive variability: A two-process model of decision making and its development. In J. E. Jacobs & P. A. Klaczynski (Eds.), *The development of decision making in children and adolescents* (pp. 39–76). Mahwah, NJ: Erlbaum.

Klayman, J. (1985). Children's decision strategies and their adaptation to task characteristics. *Organizational Behavior and Human Decision Processes, 35,* 179–201. http://dx.doi.org/10.1016/0749-5978(85)90034-2

Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian modeling for cognitive science: A practical course*. Cambridge, United Kingdom/New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139087759

Lindenberger, U., von Oertzen, T., Ghisletta, P., & Hertzog, C. (2011). Cross-sectional age variance extraction: What's change got to do with it? *Psychology and Aging, 26,* 34–47. http://dx.doi.org/10.1037/a0020525

Marewski, J. N., Pohl, R. F., & Vitouch, O. (Eds.). (2010). Recognition-based judgments and decisions: Introduction to the special issue. *Judgment and Decision Making* (Vol. 1, pp. 207–215). Retrieved from http://journal.sjdm.org/vol5.4.html

Mata, R., Josef, A. K., & Lemaire, P. (2015). Adaptive decision-making and aging. In T. M. Hess, J. Strough, & C. E. Loeckenhoff (Eds.), *Aging and decision making: Empirical and applied perspectives* (pp. 105–122). New York, NY: Elsevier.

Mata, R., Schooler, L. J., & Rieskamp, J. (2007). The aging decision maker: Cognitive aging and the adaptive selection of decision strategies. *Psychology and Aging, 22,* 796–810. http://dx.doi.org/10.1037/0882-7974.22.4.796

Mata, R., von Helversen, B., & Rieskamp, J. (2011). When easy comes hard: The development of adaptive strategy selection. *Child Development, 82,* 687–700. http://dx.doi.org/10.1111/j.1467-8624.2010.01535.x

Matzke, D., Dolan, C. V., Batchelder, W. H., & Wagenmakers, E.-J. (2015). Bayesian estimation of multinomial processing tree models with heterogeneity in participants and items. *Psychometrika, 80,* 205–235. http://dx.doi.org/10.1007/s11336-013-9374-9

Michalkiewicz, M., & Erdfelder, E. (2016). Individual differences in use of the recognition heuristic are stable across time, choice objects, domains, and presentation formats. *Memory & Cognition, 44,* 454–468.

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,* 74–80. http://dx.doi.org/10.1016/j.cognition.2013.09.007

Newell, B. R., Lagnado, D. A., & Shanks, D. R. (2007). *Straight choices: The psychology of decision making*. New York, NY: Psychology Press.

Newell, B. R., & Shanks, D. R. (2004). On the role of recognition in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 923–935. http://dx.doi.org/10.1037/0278-7393.30.4.923

Pachur, T. (2010). Recognition-based inference: When is less more in the real world? *Psychonomic Bulletin and Review, 17,* 589–598.

Pachur, T. (2011). The limited value of precise tests of the recognition heuristic. *Judgment and Decision Making, 6,* 413–422.

Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: Retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 983–1002. http://dx.doi.org/10.1037/0278-7393.32.5.983

Pachur, T., Mata, R., & Schooler, L. J. (2009). Cognitive aging and the adaptive use of recognition in decision making. *Psychology and Aging, 24,* 901–915. http://dx.doi.org/10.1037/a0017211

Pachur, T., Todd, P. M., Gigerenzer, G., Schooler, L. J., & Goldstein, D. G. (2011). The recognition heuristic: A review of theory and tests. *Frontiers in Psychology, 2,* 147. http://dx.doi.org/10.3389/fpsyg.2011.00147

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker.* Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139173933

Piaget, J., & Inhelder, B. (1951). *The origin of the idea of chance in children.* London, UK: Routledge & Kegan Paul.

Pleskac, T. J. (2007). A signal detection analysis of the recognition heuristic. *Psychonomic Bulletin & Review, 14,* 379–391. http://dx.doi.org/10.3758/BF03194081

Pohl, R. F. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making, 19,* 251–271. http://dx.doi.org/10.1002/bdm.522

Pohl, R. F., von Massow, F., & Beckmann-Schumacher, B. (2016). *Developmental differences in using a simple decision strategy: The case of the recognition heuristic.* Unpublished manuscript.

Reyna, V. F., & Brainerd, C. J. (1994). The origins of probability judgment: A review of data and theories. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 239–272). Chichester, UK: Wiley.

Reyna, V. F., & Kiernan, B. (1994). The development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology, 30,* 178–191. http://dx.doi.org/10.1037/0012-1649.30.2.178

Rosburg, T., Mecklinger, A., & Frings, C. (2011). When the brain decides: A familiarity-based approach to the recognition heuristic as evidenced by event-related brain potentials. *Psychological Science, 22,* 1527–1534. http://dx.doi.org/10.1177/0956797611417454

Ruggeri, A., & Katsikopoulos, K. V. (2013). Make your own kinds of cues: When children make more accurate inferences than adults. *Journal of Experimental Child Psychology, 115,* 517–535. http://dx.doi.org/10.1016/j.jecp.2012.11.007

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

Schlottmann, A., & Wilkening, F. (2011). Judgment and decision making in young children. In M. K. Dhami, A. Schlottmann, & M. R. Waldmann (Eds.), *Judgment and decision making as a skill: Learning, development, and evolution* (pp. 55–84). Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139015684.007

Schneider, W. (2015). *Memory development from early childhood through emerging adulthood.* Heidelberg, Germany: Springer. http://dx.doi.org/10.1007/978-3-319-09611-7

Sedlmeier, P. E., & Betsch, T. E. (Eds.). (2002). *Etc.: Frequency processing and cognition.* Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198508632.001.0001

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General, 116,* 250–264. http://dx.doi.org/10.1037/0096-3445.116.3.250

Siegler, R. S. (1999). Strategic development. *Trends in Cognitive Sciences, 3,* 430–435. http://dx.doi.org/10.1016/S1364-6613(99)01372-8

Smith, J. B., & Batchelder, W. H. (2010). Beta-MPT: Multinomial processing tree models for addressing individual differences. *Journal of Mathematical Psychology, 54,* 167–183. http://dx.doi.org/10.1016/j.jmp.2009.06.007

Volz, K. G., Schooler, L. J., Schubotz, R. I., Raab, M., Gigerenzer, G., & von Cramon, D. Y. (2006). Why you think milan is larger than modena: Neural correlates of the recognition heuristic. *Journal of Cognitive Neuroscience, 18,* 1924–1936. http://dx.doi.org/10.1162/jocn.2006.18.11.1924

Wechsler, D. (1991). *WISC-III: Wechsler Intelligence Scale for Children.* New York, NY: Psychological Corporation.

Wertz, A. E., & Wynn, K. (2014). Thyme to touch: Infants possess strategies that protect them from dangers posed by plants. *Cognition, 130,* 44–49. http://dx.doi.org/10.1016/j.cognition.2013.09.002

Wright, J. C., & Murphy, G. L. (1984). The utility of theories in intuitive statistics: The robustness of theory-based judgments. *Journal of Experimental Psychology: General, 113,* 301–322. http://dx.doi.org/10.1037/0096-3445.113.2.301

Appendix A

Description of the Multinomial Processing Tree Model

The class of multinomial processing tree (MPT) models (see Batchelder & Riefer, 1999; Erdfelder et al., 2009, for overviews) treats categorical response frequencies as probabilistic realizations of a set of underlying cognitive states (represented by the model parameters). The r-model (Hilbig, Erdfelder, & Pohl, 2010) used in the present analyses belongs to this class and can be illustrated in form of a tree diagram (Figure A1).

The model considers three possible cases in a comparative judgment task (i.e., RR, RU, and UU trial types), represented by $J = 3$ separate trees. In each of the model trees, possible responses are assigned to one of the K mutually exclusive outcome categories C_{jk}, distinguishing between inference accuracy (correct vs. false) and choice of recognized $(+)$ versus unrecognized $(-)$ objects. In the upper tree, a decision maker recognizes both objects (RR case) and therefore recruits further information beyond recognition, leading to a correct inference with probability b and to an incorrect inference with complementary probability $1 - b$. Parameter b thus indexes the validity of the decision maker's further knowledge (or any other information or strategy beyond recognition) in conceptual equivalence to knowledge validity, β. The second tree represents the situation in which one of the two objects is recognized (RU case) and the RH can thus be applied. With probability r, the decision maker uses the RH and chooses the recognized item. This leads to a correct inference with probability a and to an incorrect inference with probability $1 - a$. Parameter a reflects the strength of association between recognition and the criterion variable and is equivalent to the validity of recognition α (Goldstein & Gigerenzer, 2002). Importantly, with complementary probability $1 - r$, the RH is not applied and the inference is based on further information beyond recognition (or any other strategy). This leads to a correct inference with probability b. In this case, the recognized object is chosen with probability a and the unrecognized object is chosen with probability $1 - a$. With probability $1 - b$, the inference is incorrect. In this case, the unrecognized item is chosen with probability a and the recognized item is chosen with probability $1 - a$. In the bottom tree, neither of the objects is recognized (UU case) and the decision maker has to guess, leading to a correct inference with probability g.

We employed a hierarchical Bayesian implementation of the r-model to account for individual differences in RH use. Further details are in the online supplementary materials.
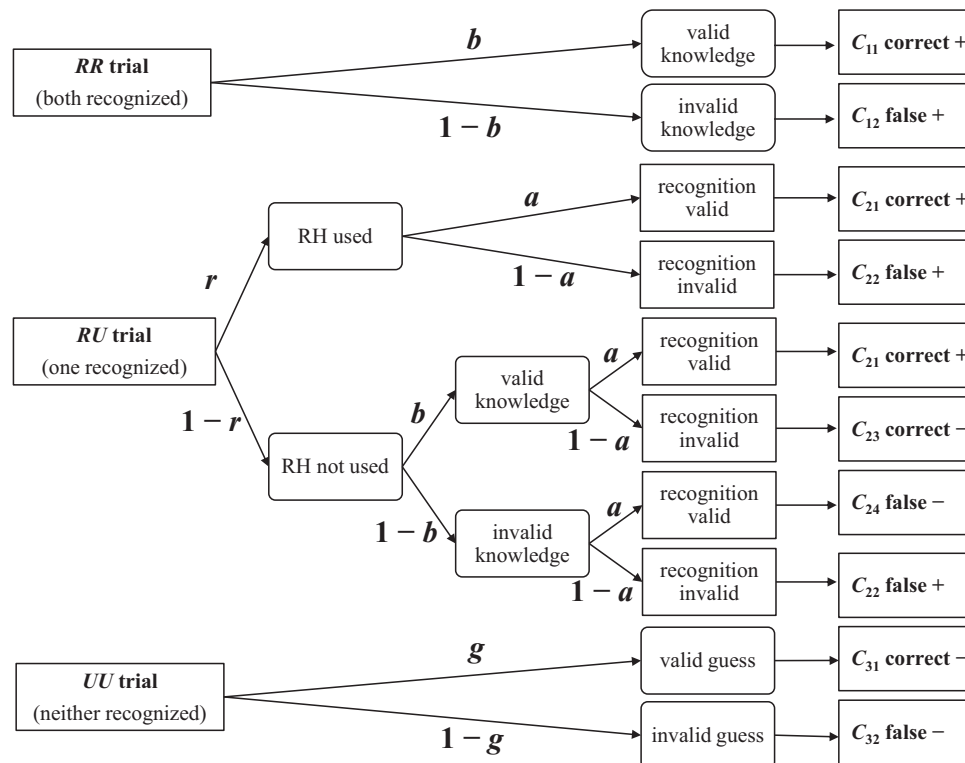

Figure A1. The r-model with parameters b (validity of further knowledge), g (probability of a correct guess), r (probability of applying the recognition heuristic), and a (recognition validity).

(Appendices continue)

Appendix B

Regression Analyses

This section provides further details about the mediation analyses mentioned in the Results section and includes tables of regression coefficients (Table B1) with corresponding model diagrams (see Model 1 in Figures B1 and B2) for ease of reference. The findings can be summarized as follows: significant indirect effects of age through a measure of verbal knowledge emerged on inference accuracy, discrimination ability (d'_{RH}), and validity of further information (parameter b). Analogous indirect effects of age did not emerge through a measure of cognitive speed.

Use of the recognition heuristic was not predicted by age, but strongly hinged on the interaction with environment. Consequently, when task environment was included as a moderator in the mediation model (Model 2 in Figures B1/B2), we also observed an indirect effect of age through knowledge (but not through speed) on adaptive use of the recognition heuristic. Notably, a positive indirect effect (mediation through knowledge) on use of the recognition heuristic emerged in the cities environment, whereas this effect was negative in the diseases environment (see Table B2).

Table B1
Mediation Analyses (Model 1)

Entered variables			Model coefficients			
X	M	Y	a	b	c'	$a \times b$
Age	Verbal	Accuracy	**.45 [.30, .60]**	**.31 [.13, .48]**	**.20 [.01, .38]**	**.14 [.06, .24]**
Age	Verbal	d'_{RH}	**.45 [.30, .60]**	**.29 [.08, .50]**	**.26 [.06, .45]**	**.13 [.03, .26]**
Age	Verbal	Parameter b	**.45 [.30, .60]**	**.27 [.11, .42]**	**.37 [.21, .52]**	**.12 [.05, .22]**
Age	Verbal	Parameter r	**.45 [.30, .60]**	.10 [−.12, .31]	.07 [−.15, 29]	.04 [−.04, .15]
Age	Speed	Accuracy	**.66 [.53, .80]**	.08 [−.16, .33]	**.28 [.03, .52]**	.06 [−.11, .22]
Age	Speed	d'_{RH}	**.66 [.53, .80]**	.03 [−.20, .25]	**.37 [.12, .62]**	.02 [−.15, .15]
Age	Speed	Parameter b	**.66 [.53, .80]**	.07 [−.15, 29]	**.44 [.24, .64]**	.05 [−.10, .20]
Age	Speed	Parameter r	**.66 [.53, .80]**	.08 [−.16, .31]	.06 [−.20, .33]	.05 [−.10, .21]

Note. Model coefficients a, b, c', refer to standardized mean regression weights in Mediation Model 1 (see Figure B2); $a \times b$ is the indirect effect of variable X on Y through M; 95% confidence intervals are in brackets and indicate significant effects if the interval does not include the value zero (significant values $p < .05$ are given in boldface). Indirect effects are based on bootstrap estimates (with $N = 50,000$ resamples and bias-corrected confidence intervals). Parameter r = probability of applying the recognition heuristic; parameter b = validity of further information; d'_{RH} = discrimination ability between correct and incorrect predictions of the recognition heuristic.
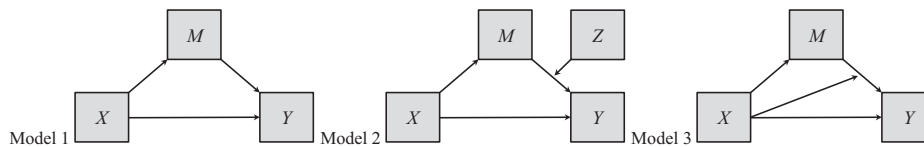

Figure B1. Conceptual diagrams.
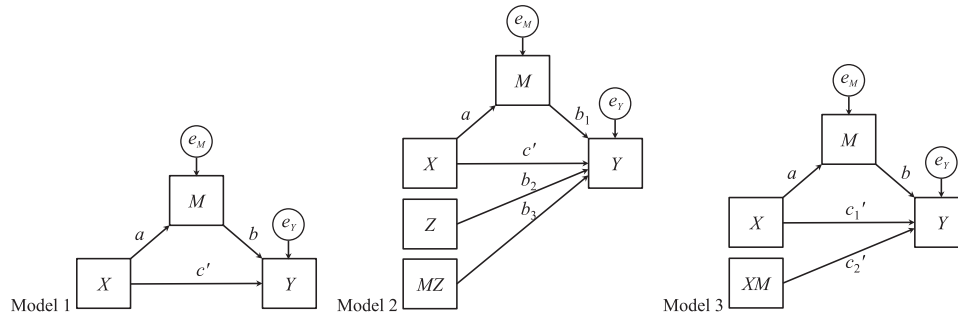
(*Appendices continue*)

Figure B2. Statistical diagrams.

Table B2

Mediation Analyses With Task Environment as Moderator (Model 2)

	Entered variables			Model coefficients						
X	M	Y	Z	b_1	b_2	b_3 (Interaction)[a]	c'	Conditional indirect effect[b] cities	Conditional indirect effect[b] diseases	Index of moderated mediation[c]
Age	Verbal	Parameter r	Environment	.25 [−.03, .53]	**−.71 [−1.05, −.38]**	**−.49 [−.82, −.16]**	.12 [−.08, .32]	**.11 [.01, .26]**	**−.11 [−.25, −.01]**	**−.22 [−.41, −.07]**
Age	Speed	Parameter r	Environment	.17 [−.11, .45]	**−.71 [−1.07, −.36]**	−.33 [−.69, .04]	.11 [−.14, .35]	.11 [−.04, .30]	−.10 [−.34, 13]	−.22 [−.45, .02]

Note. Model coefficients are standardized mean regression weights in Model 2 (see Figure B2). Estimates for model coefficient *a* can be obtained from Table B1; 95% confidence intervals are in brackets and indicate significant effects if the interval does not include the value zero (marked in boldface). Indirect effects are bootstrapped (with $N = 50,000$ resamples and bias-corrected confidence intervals). Parameter r = probability of applying the recognition heuristic.
[a] Model coefficient b_3 represents the interaction between variables M and Z. [b] The conditional indirect effect of X on Y through M at different values of Moderator Z is obtained as $a \times (b_1 + b_3) \times Z$. [c] The product $a \times b_3 \times (z_1 - z_2)$ can be interpreted as an index of moderated mediation (where z_i are the levels of the moderator; see Hayes, 2013). That is, an inference as to whether $a \times b_3 \times (z_1 - z_2)$ is statistically different from zero serves as a formal test of equality of the indirect effects between the task environments.

Table B3

Mediation Analyses with Age as Moderator (Model 3)

	Entered variables		Model coefficients	
X	M	Y	c'_2 interaction[a]	$a \times c'_2$ index of moderated mediation[b]
Age	Verbal	Accuracy	−.03 [−.24, .19]	−.01 [−.10,.08]
Age	Verbal	d'_{RH}	−.02 [−.27, .22]	−.01 [−.11, .11]
Age	Verbal	Parameter b	.07 [−.11, .24]	.03 [−.05, .12]
Age	Speed	Accuracy	.15 [−.03, 34]	.10 [−.02, .23]
Age	Speed	d'_{RH}	.06 [−.16, 27]	.04 [−.09, .18]
Age	Speed	Parameter b	**.20 [.03, .37]**	**.14 [.02, .27]**

Note. Model coefficients are standardized mean regression weights in Model 3 (see Figure B2). Estimates for further model coefficients a, b, c'_1, $a \times b$ can be obtained from Table B1; 95% confidence intervals are in brackets and indicate significant effects if the interval does not include the value zero (marked in boldface). Indirect effects are bootstrapped (with $N = 50,000$ resamples and bias-corrected confidence intervals). Parameter b = validity of further information; d'_{RH} = discrimination ability between correct and incorrect predictions of the recognition heuristic.
[a] Model coefficient c'_2 measures the interaction between variables X and M. [b] The product of model coefficients a and c'_2 can be interpreted as an index of moderated mediation. That is, an inference as to whether $a \times c'_2$ is statistically different from zero serves as formal test of moderation of the indirect effect by the moderator in the model (Hayes, 2013).

Finally, to explore whether mediation differed between age groups, we also considered moderated mediation models with age as both predictor and moderator (Model 3/Table B3). These explorations provided little evidence for systematic moderating effects of age. Finally, please note that all path analyses serve to describe interrelations of variables of interest. With concurrent data, they should be interpreted with caution, cannot provide evidence for any causal relationships, and serve to test necessary (but not sufficient) preconditions for mediation mechanisms (e.g., Lindenberger, von Oertzen, Ghisletta, & Hertzog, 2011).