Judgment and Decision Making (2023), **Vol. 18**:e38 1–21
doi:10.1017/jdm.2023.36

EMPIRICAL ARTICLE

Preschoolers' use of cue validities as weights in decision-making: Certainty does not substantially change the world

Stefanie Lindow[1], Anne Lehmann[1], David Buttelmann[2] and Tilmann Betsch[1]

[1]Department of Psychology, University of Erfurt, Erfurt, Germany and [2]Department of Psychology, University of Bern, Bern, Switzerland

Corresponding author: Stefanie Lindow; Email: stefanie.lindow@uni-erfurt.de

Received: 13 January 2023; **Revised:** 26 August 2023; **Accepted:** 13 September 2023

Keywords: child decision-making; decision weights; probabilistic inference choice; information board

Abstract
A child's world is full of cues that may help to learn about decision options by providing valuable predictions. However, not all cues are always equally valid. To enhance decision-making, one should use cue validities as weights in decision-making. Prior research showed children's difficulty in doing so. In 2 conceptual replication studies, we investigated preschoolers' competencies when they encounter a cue whose prediction is always correct. We assessed 5- to 6-year-olds' cue evaluations and decision-making in an information-board-game. Participants faced 3 cues when repeatedly choosing between 2 locations to find treasures: A nonprobabilistic, high-validity cue that always provided correct predictions ($p = 1$) paired with 2 probabilistically correct (Study 1: $p = .34$, $p = .17$) or 2 nonprobabilistic, incorrect cues (Study 2: $p = 0$). Participants considered cue validities—albeit in a rudimentary form. In their cue evaluations, they preferred the high-validity cue, indicating their ability to understand and use cue validity for evaluations. However, in their decision-making, they did not prioritize the high-validity cue. Rather, they frequently searched and followed the predictions of less valid (Study 1) and incorrect cues (Study 2). Our studies strengthen the current state of decision research suggesting that the systematic use of cue validities in decision-making develops throughout childhood. Apparently, having appropriate cue evaluations that reflect cue validities is not sufficient for their use in decision-making. We discuss our findings while considering the importance of learning instances for the development of decision competencies.

1. Introduction

A popular decision-making method among children is the use of counting rhymes such as 'eenie meenie miney mo', which leaves the choice up to chance. Although, in a risky world, the outcomes of our choices are indeed determined by chance, competent decision-makers can utilize the laws of chance to enhance decision-making. Specifically, in a risky world, the predictions of cues do not occur at random but instead are probabilistically correct, that is, cues have a specific validity. Such cues can, for example, be features of decision options (e.g., city-size decision task: Gigerenzer et al., 1991) or advice-givers who make predictions on the decision options such as testers of consumer products (e.g., consumer decision task: Glöckner and Betsch, 2008). Comparable to adults, the world of children is full of such cues. Consider, as an example, a common everyday decision of children. The child has misplaced a needed item such as shoes or a stuffed animal. In order to find the lost item, the child can ask for advice in terms of predictions on where to look. Possible advice-givers (i.e., cues) could be family members. Importantly, these cues may differ in their validity in terms of their accuracy of being correct in the past: While the attentive fathers' advice was correct quite often, the grandmother, who is only in the child's home for visits, did not reliably predict locations in the past. In this case, the child could consider only the prediction of the father as the most valid cue and ignore the grandmothers' prediction. This procedure corresponds to the application of the take-the-best heuristic, which is a lexicographic decision strategy (TTB: Gigerenzer and Gaissmaier, 2011). Decision strategies describe the way decision-makers utilize cues in decision-making. Heuristics are a subset of decision strategies 'that ignore(s) part of the information, with the goal of making decisions more quickly, frugally, and/or accurately than more complex methods' (Gigerenzer and Gaissmaier, 2011, p. 454). For example, when applying TTB, decision-makers choose according to the first cue that discriminates between the options after the cues are ordered by their validity. Considering an additional cue would only be necessary if the most valid cue does not differentiate between the decision options (e.g., Gigerenzer and Gaissmaier, 2011). Thus, if one uses cue validity as a weight to guide information search, decision-makers can minimize effort and reduce the amount of considered information. Besides heuristics, complex, normative models of decision-making also posit the use of cue validities as weights in decision-making (expected utility theory: Von Neumann and Morgenstern, 1944). Here, although a comprehensive information sample is necessary to estimate the value of options, cue validities are used as weights to guide information integration for choice. Returning to the example of the misplaced item, the child may ask for both the father's and grandmother's advice in order to obtain a full picture of the situation. However, when evaluating the different locations, the father's advice outweighs the grandmother's advice. In this way, considering cue validities enhances judgment and decision-making (Bröder et al., 2023; Glöckner and Betsch, 2008). The consideration of weights has been proposed to be a 'law' of choice behavior (Bhatia et al., 2021)—at least among adults. This raises the question of its developmental path. The present research addresses the question of how weighting in decision-making develops.

From a theoretical perspective, it is plausible to assume developmental improvement in weighting if weighting underlies cognitive abilities (Löckenhoff, 2018). To differentiate, order, and select cues according to their validity, decision-makers likely require a number of cognitive prerequisites. For example, decision-makers who apply the TTB strategy, must prioritize the high-validity cue and inhibit the less valid cues. This requires executive functions such as selective attention (see attention allocation hypothesis: Mata et al., 2011). Using cue validities for effective information integration in more complex, compensatory strategies has been argued to also draw on cognitive abilities. Decision research with older adults suggests that changes in processing speed and working memory affect successful information integration (Fechner et al., 2019; Löckenhoff, 2018). Also, cognitive resources associated with fluid intelligence are assumed to set an upper limit to the cognitive effort that a decision-maker can invest during integration (Mata et al., 2007). Decades of developmental research generally suggest improvement of basic cognitive abilities such as selective attention and working memory across childhood and even throughout adolescence (overviews: Blumberg et al., 2005; Diamond, 2013).

Along with this notion, findings with standard judgment and decision-making (JDM) paradigms repeatedly showed children's difficulty in systematically weighting information (e.g., IOWA Gambling Task for risk learning: Beitz et al., 2014; gambling tasks for expected value integration: Levin and Hart, 2003; Levin et al., 2007; overview: Schlottmann and Wilkening, 2012). For example, in the treasure hunt game—a child-friendly adaption of the information-board paradigm for probabilistic inference choices—children were asked to find treasures hidden in houses either on a game board (e.g., Betsch and Lang, 2013) or in a matrix presented on the computer (Mousekids, e.g., Betsch et al., 2014, 2016, 2020). Before children decided on a house (i.e., choice options), they could inspect the predictions of different cues regarding the assumed location of the treasure. The cues differed in their validity, which children had previously observed in a series of learning treasure hunts. For example, whereas the high-validity cue had been correct in 5 out of 6 treasure hunts, the lower validity cues had been correct in only 1 or 2 treasure hunts. Typically, the authors used such a noncompensatory environment with a high-validity cue that clearly outweighs the predictions of the other cues. The TTB heuristic to search and choose according to the prediction of the high-validity cue is appropriate in such an environment. Initially, the cue predications are covered behind closed windows on the game board or computer matrix. In this way, information search (in terms of inspecting cue predictions) and subsequent choices (in terms of choosing a house) can be analyzed. Weighting can be assessed by analyzing how systematic the high-validity cue is prioritized. Studies with this paradigm showed that most adults but only some elementary schoolers (8- to 9-year-olds) and no preschoolers (5- to 6-year-olds) systematically prioritized the high-validity cue in their search and choice (e.g., Betsch and Lang, 2013; Betsch et al., 2020; overview: Lindow and Betsch, 2021). In line with these findings, Mata et al. (2011) used a similar information-board decision game to show that, even for 9- to 10-year-olds, applying decision strategies that rely on weighting is not a trivial task. In their study, 9- to 10-year-olds showed a more pronounced reliance on equal weighting compared to the older age groups, indicating that they may neglect cue validities when considering cue predictions. Also, 9- to 10-year-olds showed poorer learning in repeated trials and more application errors accompanied by poorer performance. Following up on this line of research, Betsch et al. (2020) tackled the objection that children's difficulties in using cue validities are merely caused by methodological issues of the paradigms, that is, by faulty procedures of probability induction or the failure to generalize cue validities from the learning to the test phase. Specifically, they ensured probability induction by employing a successful procedure for learning cue validities from trust-in-informants research in their paradigm (e.g., Pasquini et al., 2007). In addition, they expanded the measures and assessed cue evaluations in addition to decision-making. In their cue evaluations, preschoolers prioritized the high-validity cue by frequently preferring it over the other cues. This demonstrates both proper learning and generalizing of cue validities. However, in contrast to adults, preschoolers neglected cue validities in their decision-making.

In sum, this line of research suggests that the use of cue validities as weights in decision-making appears to be a challenge for children. Maturation was reported to occur across school age. The direction of this developmental trend could be explained by a link of decision-making and the maturation in general cognitive abilities (e.g., attention allocation hypothesis: Mata et al., 2011). At the same time, however, the late onset of weighting abilities with serious difficulties still occurring at school age is somewhat surprising. School-aged children do have a basic set of cognitive abilities and can, for example, selectively attend to information (e.g., apply the recognition heuristic: Horn et al., 2016) and integrate multiple information at least to some extent (e.g., Ebersbach, 2009; overviews: Blumberg et al., 2005; Diamond, 2013). In addition, research using familiar, experience-based contexts has demonstrated weighting abilities in preschoolers' decision-making (Koenig and Harris, 2005; Lindow and Betsch, 2021; Pasquini et al., 2007). For example, it is a well-replicated finding that, after having observed informants' performances on an object labeling task (i.e., the cue validity), children as young as 4 years of age follow a high-validity informant over a less valid one when subsequently asked to label other objects themselves (e.g., Koenig and Harris, 2005; Pasquini et al., 2007). In contrast, the novel decision situations used in JDM paradigms may trigger exploration approaches in children (see, e.g., Bjorklund, 2018, for the functionality of play for development and maturation) that may hamper cue weighting. As soon as children recognize the fallibility of a cue in a novel environment, they might become cautious. Despite knowing who the most valid cue is, they might still search and follow the predictions of less valid cues for exploration purposes. Conceptually speaking, children might equalize cue validities and use the same, 'unsure' weight for all probabilistic cues involved. In this case, the subjective cue validities would differ widely from the objective, optimal ones (e.g., lens model: Brunswik, 1952) and, consequently, substantially change the decision environment from noncompensatory to compensatory with equal cue weights.

1.1. Research approach and expectations

If children's difficulties in cue weighting are caused by an equalization of probabilistic cue validities, they should be resolved if children encounter a nonprobabilistic cue whose prediction is always correct. To test this assumption, we modified the treasure hunt paradigm of Betsch et al. (2020). Specifically, in

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

our conceptual replication studies, we used a cue that was consistently correct (i.e., a nonprobabilistic, high-validity cue). Following the original procedure, we applied a noncompensatory structure of cue validities. That is, the high-validity cue ($p_1 = 1$) clearly outweighed the other, less valid cues (Study 1: $p_2 = .34$, $p_3 = .17$; Study 2: $p_2 = 0$, $p_3 = 0$). In such an environment, the application of the TTB strategy is highly adaptive: Decision-makers can save effort and reduce the amount of information to consider by prioritizing the high-validity cue and inhibiting the less valid cues. Prioritization and inhibition are an expression of using cue validity as a weight to guide information search and choice.

As in the original study, in our conceptual replications, we assessed cue evaluations and decision-making (in terms of information search and choice). We expected to replicate Betsch et al.'s (2020) findings concerning cue evaluations. That is, we expected preschoolers to prioritize the high-validity cue by preferring it over the other cues in their evaluations. Contrary to the findings of Betsch et al., we hypothesized that, when presented with a nonprobabilistic, high-validity cue in our study, preschoolers also demonstrate cue weighting in decision-making. That is, we expected preschoolers to prioritize the high-validity cue in their information search and systematically follow its predictions for choice.

Our studies were not preregistered. As conceptual replications, whenever possible, we followed Betsch et al. (2020).

2. Study 1

In Study 1, we investigated partially certain decision environments. That is, the 2 cues that we used in addition to the nonprobabilistic, high-validity cue were probabilistically valid ($p_1 = 1$, $p_2 = .34$, $p_3 = .17$).

2.1. Method

2.1.1. Participants

Thirty-four 5- to 6-year-olds participated (41.2% female; age: Mdn = 77.5 months, M = 76.4, SD = 5.1). The selection of the age range and sample size (aimed at $N > 30$) followed prior research (e.g., Betsch et al., 2020; Betsch and Lang, 2013) and allows for detecting medium effects with $\alpha = .05$ and at least $1 - \beta \geq .85$ in the conducted tests (i.e., repeated-measures analysis of variance, one-sample, one-tailed *t*-test, one-sample, two-tailed binomial test; a priori power analysis conducted with G*Power, Faul et al., 2007). Children (all native German speakers) were recruited in 4 daycare centers located in middle-class areas of 2 moderately large cities in Germany. Parents had previously provided written consent. Participants received a personalized certificate of participation. This research was approved by the research ethics board of the University of Erfurt; project title 'Probabilistic decisions in children'.

2.1.2. Material

The material is publicly available at the OSF project page: https://doi.org/10.17605/OSF.IO/HE7ZR; https://osf.io/he7zr/?view_only=184f9b892ea74db89b9b5d914fde9dd1.

We used the Mousekids decision game from Betsch et al. (2020). All modifications concern changing the high-validity cue with an initial accuracy of 83% (i.e., providing correct predictions 5 out of 6 times) into a nonprobabilistic, high-validity cue (i.e., always provides correct predictions).

In the learning session of the decision game, participants were introduced to 3 cues. They observed the cues' performance when labeling 6 well-known objects (i.e., toys: ball, puzzle, dice, building blocks, cuddly toy, and toys). The high-validity cue labeled all 6 objects correctly (100%); the medium validity cue labeled 2 out of 6 objects (34%) correctly; and the low-validity cue labeled 1 out of 6 objects (17%) correctly. To learn these validities, participants watched video clips in which 3 female actors sat next to each other behind a table (Figure 1a). Each clip began with showing the object positioned on the middle of the table (e.g., a ball). Then, each cue, one after the other, and starting from left to right, labeled the object (e.g., correct label: ball, incorrect label: doll). To facilitate discrimination, the cues wore shirts and caps of different colors (i.e., yellow, green, and blue). We used a pool of 36 possible clip

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

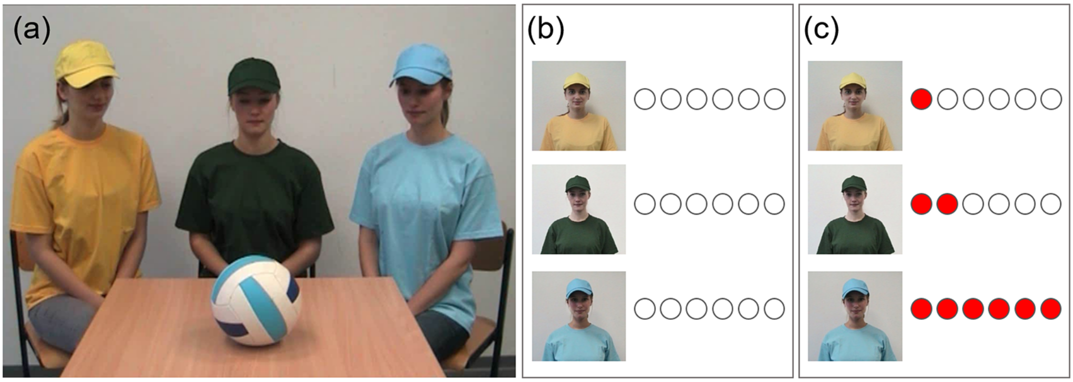

Figure 1. *Material used during the learning session: (a) shows an example frame from the video clips, (b) shows the blank documentation sheet, and (c) shows the colored documentation sheet at the end of the learning session.*

orders to counterbalance the order of objects and the position of the cues at the table (6 variants each). We used the same video material as Betsch et al. (2020) but partly rearranged the material to increase the validity of the high-validity cue to certainty. This learning procedure is a standard procedure in trust-in-informants research (overview: Koenig et al., 2019) and has been shown to successfully ensure probability induction in the treasure hunt game (Betsch et al., 2020).

After each clip, participants documented the cue validity on a documentation sheet. The sheet showed pictures of the cues next to a string of 6 blank circles each (Figure 1b, 6 versions of different picture orderings were used that matched the video clips). For each correct label, participants colored one circle of the corresponding cue. Thus, at the end of the learning session, the distribution of colored and blank circles visualized the cue validities (Figure 1c).

For the test session of the decision game, the computerized Mousekids were used (Betsch et al., 2020). Participants repeatedly chose between 2 houses that contained either a treasure or a spider. For each treasure found, participants received a treasure point. The goal of the game was to collect as many treasure points as possible. Participants could inspect the cues' predictions—that is, what cues reported were hidden in the house (binary: treasure vs. spider). This information was presented to participants in an information-board matrix. As shown in Figure 2, the choice options (i.e., the houses) appeared in rows, whereas the cues appeared in columns. To visualize and transfer the learned cue validities, the pictorial representation of colored/blank circles that was introduced in the learning session appeared to the left of each cue. To the right, in the matrix cells, the prediction appeared in the form of spider and treasure pictures. The predictions were initially covered behind closed windows that could be opened. The order of cues was always as shown in Figure 2, with the high-validity cue positioned in the bottom row (to avoid confounds of reading direction with validity).

Following Betsch et al. (2020), we used 3 different prediction patterns across the decision tasks. In decisions of pattern Type 3, the high-validity cue was consistently contradicted by the 2 other cues (Figure 3). In decisions of pattern Type 1 and 2, the medium validity cue made the same prediction for both options (i.e., 2 spiders in Type 1 and 2 treasures in Type 2), whereas the high-validity cue and the low-validity cue contradicted each other. If one considers cue validity, the pattern type should have no effect on decision-making. The prediction of the high-validity cue cannot be outweighed by the lower validity cues—regardless of whether they consistently contradict the high-validity cue (Betsch et al., 2020). Thus, we compared decision behavior between the pattern types to further understand the use of cue validities.

The test session consisted of 26 decisions. Specifically, 2 practice decisions were followed by 24 test decisions. In the practice decisions, cues made the same prediction on a single option. In the test

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

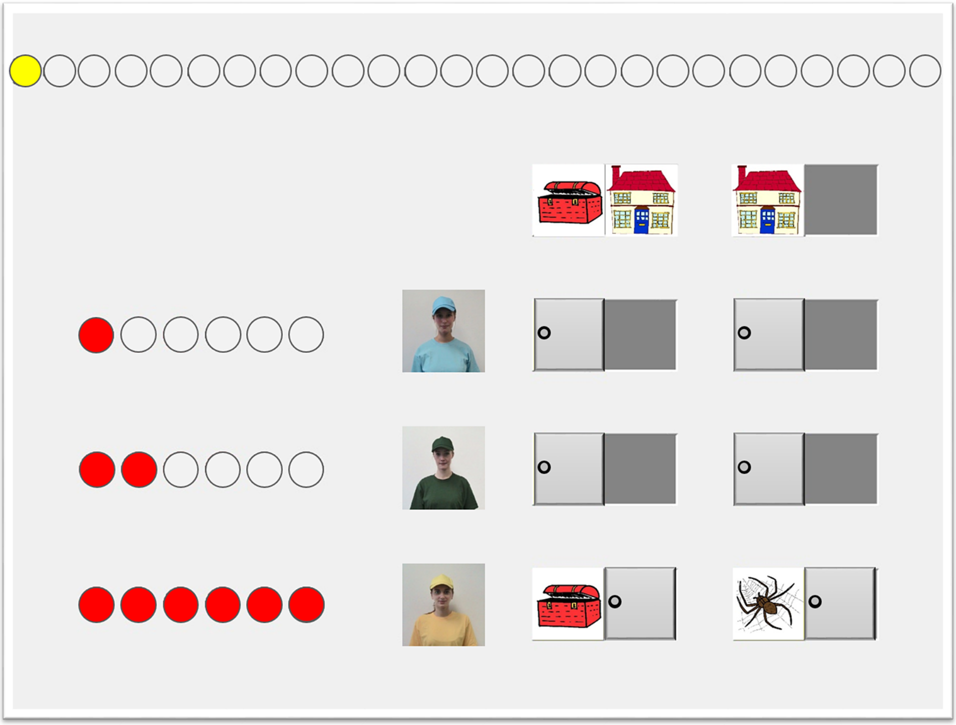

Figure 2. *The treasure hunt game (Mousekids, Betsch et al., 2020). Screenshot shows 1 example decision of the test session. In this example, the participant has searched the prediction of the high-validity cue (i.e., cue with all red circles at the bottom) for both houses (i.e., windows of the bottom row are both opened) and subsequently followed her prediction (i.e., house with the treasure prediction is opened) without searching the prediction of the other cues (i.e., windows in the first and second rows are closed). The participant found a treasure in the house (i.e., treasure icon next to the left house) and received a treasure point (yellow circle in the top line of 26 circles in total).*

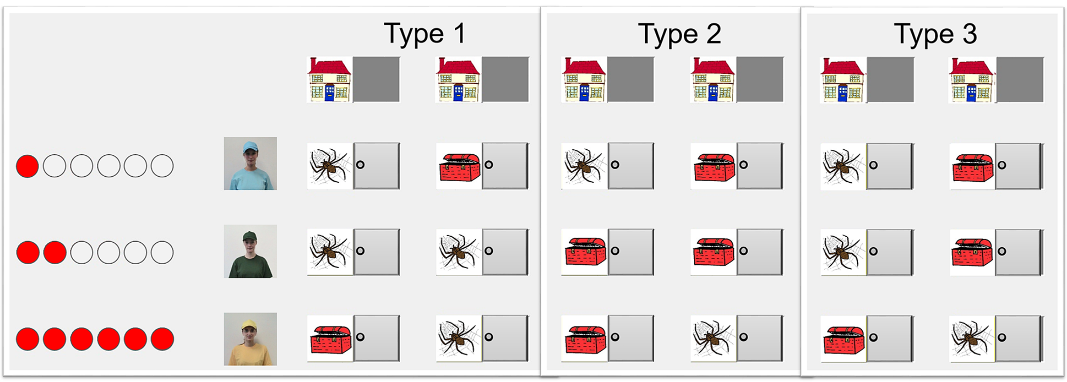

Figure 3. *Pattern types used in Study 1 (replicating Betsch et al., 2020). In pattern Type 3, the high-validity cue is consistently contradicted by both other cues. In pattern Type 1 and 2, the medium validity cue provides the same prediction for both houses, whereas the low-validity cue contradicts the high-validity cue.*

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

decisions, the 3 pattern types (Figure 3) appeared in random order with eight decisions each (i.e., within-subjects manipulation of pattern type). Note that we counterbalanced the order of the house preferred by the high-validity cue. That is, in half of the decisions per pattern type, she predicted the left and right house, respectively. Since her prediction was always correct, each house contained the treasure in half of the decisions. Due to logical constraints arising from the prediction patterns (Figure 3), the cue validities of the 2 lower validity cues in the test session deviated slightly from their cue validities in the learning session (i.e., test session: 31%, 23%, learning session: 34%, 17%).

2.1.3. Procedure

Participants were tested individually in a quiet room in the child's daycare center by 1 of 4 female experimenters. The overall procedure was as follows: First, the game was introduced as a treasure hunt and subsequently explained. Then, the learning session, manipulation check, test session, and cue evaluation followed (replication of Betsch et al., 2020). At the end of the study, the experimenter commended participants for their earned treasure points (regardless of the actual number) and created a personalized certificate with a picture of the participant. After completion, participants participated in a second study with the aim to assess the comprehension of the treasure hunt game. Finally, they were debriefed and received their personalized certificate of participation.

2.1.3.1. Learning session

To learn the cue validities, children observed the cues' performance in a labeling task. The experimenter stated: 'Let's have a look how familiar the women are with toys, so we know how well they can help you with the treasure hunt game'. This was used to connect the learning and test sessions (Betsch et al., 2020). First, the experimenter showed pictures of the objects used in the study and asked participants to name them. This ensured that children knew the labels for the subsequently used objects. The experimenter provided the correct label in case the participant's label deviated. Then, the experimenter began a sequence of 6 video clips that each showed 1 labeling trial. She stated: 'Let's see how the three name different toys. We will count how often they name the toys right'. After each labeling trial, participants recorded the performance of each cue on the documentation sheet along with a verbalization of the cue's performance by the experimenter (Figure 1).

2.1.3.2. Manipulation check

Correct encoding of the cue validities and the cues' smartness was assessed after the learning and test sessions. As a manipulation check, the experimenter asked participants to indicate which cue had been correct most often in the learning session.

2.1.3.3. Test session

The experimenter started Mousekids on the computer. She then explained the treasure hunt game and the handling of the computer program, which she operated. She stated: 'So, we have 2 houses, you see them up here. The treasure is either in this house or this house. Or maybe there is a treasure in both houses. Or in neither. We don't know. But the women will tell you what house they think the treasures and spiders are hidden in'. Before starting the test decisions, the experimenter repeatedly summarized the meaning and result of the learning session. She repeated that now that they know how often the cues were correct when labeling toys, they also know how well the cues can help in the treasure hunt game (see procedure to ensure the generalization from learning to test session in Betsch et al., 2020). Then, the experimenter began the first decision by stating: 'You are now allowed to open windows to see what the people say (. . .). You may open as many windows as you want'. During the first and second decisions (i.e., the practice decisions), participants were asked to verbalize the cue predictions and were corrected, if necessary. In each of the 26 decisions (2 practice decisions, 24 test decisions), participants opened as many windows as they wished before they made a choice by opening one of the houses (see Figure 2 for an example). If a treasure was found in the house, the experimenter said: 'Oh a treasure, now you get a treasure point' (which was subsequently colored in on the computer, Figure 2). If a

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

Table 1. *Wording of questions for cue evaluation.*

Domain	Question
Labeling	Who knows the names of different animals?
Knowledge	Who knows that deciduous trees lose their leaves in autumn?
Creativity	Who can paint pretty pictures?
Prosocial behavior	Who shares their toys with others?
Friendship	Who would you like to become friends with?
Treasure	Who knows which house has a treasure in it?

spider appeared, she said: 'Oh, unfortunately no treasure. Well, let's try again'. Immediately afterwards the next decision followed. The house that was not chosen and previously unopened windows (i.e., not searched cue predictions) were not uncovered for inspection.

2.1.3.4. Cue evaluation

The experimenter asked participants to indicate the best cue(s) in 6 different domains. For example, she asked 'Who knows the names of different animals?' (Table 1). The experimenter used 1 of 7 categories to protocol responses: high-validity cue, medium validity cue, low-validity cue, 2 cues, all 3 cues, none of the cues, other (e.g., do not know, no conclusion possible).

2.2. Results and discussion

Anonymized data are publicly available at the OSF project page: https://doi.org/10.17605/OSF.IO/HE7ZR; https://osf.io/he7zr/?view_only=184f9b892ea74db89b9b5d914fde9dd1.

2.2.1. Manipulation check of cue validity learning

Following the procedure of Betsch et al. (2020), we asked participants to recall the learning session in order to assess the correct learning of cue validities. All participants passed this manipulation check by responding that the high-validity cue was correct most often. One participant was excluded because she did not finish the study. The final data set we used for analysis included 33 participants.

2.2.2. Cue evaluations

As evident from the percentages of participants reported in Table 2, the majority of preschoolers preferred the high-validity cue when asked for an evaluation of the cue in different domains. For example, 61% of the preschoolers indicated that the high-validity cue knows that deciduous trees lose their leaves in autumn. Only 24% indicated that the medium validity cue would know (all other responses occurred rarely, Table 2).

For inferential analysis, we collapsed evaluations across domains. Table 3 shows how often participants checked a particular response category on average across the domains. Mean values could range from 0 to 5 given the 5 domains. Note that due to infrequent use, we collapsed 4 response categories into an 'other' category (i.e., '2 cues', 'all cues', 'none', and 'other'). The high-validity cue was clearly the most frequent response, substantiated by a large main effect of response category in a repeated-measures analysis of variance ($F(3, 96) = 25.89$, $p < .001$, $\eta^2 = .45$) and the nonoverlapping confidence intervals (Table 3). This indicates that when preschoolers evaluate cues, they consider cue validity.

2.2.3. Information search

We assessed the depth of information search by analyzing the amount of possible cue predictions preschoolers searched. If one uses cue validity as a weight to guide information search in our

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

Table 2. *Cue evaluations in Study 1.*

Domain	High-validity cue	Medium validity cue	Low-validity cue	All cues	None	Other
Labeling	60.6 (20)	9.1 (3)	18.2 (6)	0	0	12.1 (4)
Knowledge	60.6 (20)	24.2 (8)	6.1 (2)	3.0 (1)	0	6.1 (2)
Creativity	54.5 (18)	15.2 (5)	21.2 (7)	6.1 (2)	0	3.0 (1)
Prosocial behavior	54.5 (18)	15.2 (5)	21.2 (7)	3.0 (1)	0	6.1 (1)
Friendship	54.5 (18)	12.1 (4)	15.2 (5)	12.1 (4)	3.0 (1)	3.0 (1)
Treasure	66.7 (22)	9.1 (3)	15.2 (5)	3.0 (1)	3.0 (1)	3.0 (1)
Averaged across domains	58.6	14.2	16.2	4.5	1.0	5.6

Note: The table shows the percentage (number) of participants that responded with the respective category. The response category '2 cues' did not occur and is thus not included in the table.

Table 3. *Frequencies of cue evaluations averaged across domains for Studies 1 and 2.*

	High-validity cue	Medium validity/ incorrect cue*	Low-validity/ incorrect cue**	Other response***
Study 1	3.5 (1.9) [2.9, 4.2]	0.8 (1.0) [0.5, 1.2]	1.0 (1.1) [0.6, 1.4]	0.7 (1.1) [0.3, 1.0]
Study 2	4.5 (1.7) [3.9, 5.0]	0.6 (1.0) [0.3, 1.0]	0.6 (0.8) [0.3, 0.8]	0.3 (0.7) [0.1, 0.6]

Note: Frequencies (M, (SD), [95% CI]) indicate how often participants checked the response categories on average across domains (range 0–5).

*Medium validity cue in Study 1, incorrect cue, middle row in Study 2;

**Low-validity cue in Study 1, incorrect cue, top row in Study 2;

***'Other response' includes '2 cues', 'all cues', 'none', and 'other'.

noncompensatory decision environment, decision-makers can save effort and reduce the amount of information considered by prioritizing the high-validity cue (e.g., when applying TTB). In contrast to such a heuristic approach, overall, preschoolers search was quite exhaustive. They tended to search almost all predictions. That is, on average, they searched 5.2 predictions (Bootstrap 95% CI [4.8, 5.6], SD = 1.2). However, notably, they searched in almost all of their decisions at least one of the predictions of the high-validity cue (M = 95.0%, SD = 9.7, CI [91.1, 97.9]). With this, they had a good information sample available for choice although they did not show a tendency to reduce the amount of information by focusing on the high-validity cue.

Following the analysis procedure of Betsch et al. (2020), we additionally analyzed which cue participants searched first. The start of search is considered a reliable measure for prioritizing the predictions of the high-validity cue (Betsch et al., 2016, 2020). For example, TTB begins with the inspection of the high validities cue prediction. In our study, preschoolers started their search with the high-validity cue in 46.0% of the decisions (CI [36.0, 58.0], SD = 33.6). This is significantly above a chance level of 25%[1] (one-sample *t*-test, one-tailed, test-value 25; $t(32) = 3.59$; $p < .001$, $d = 0.624$). However, the low-validity cue positioned at the top of the information board was also frequently searched as the first cue (44.6%, CI [33.0, 55.4], SD = 33.7). This is noteworthy because her prediction is normatively irrelevant (see Betsch et al., 2020 for a discussion). The observation that preschoolers still frequently searched this cue prediction first indicates their difficulty in weighting cues by validity in information search. Rarely, children started their search with the medium validity cue positioned at the middle (7.2%, CI [3.4, 11.7], SD = 12.5) or searched no cue (2.3%, CI [0.6, 4.4], SD = 5.6).

[1]The variable examined is which cue participants searched first. Chance level is set to 25% because of four possible responses on this variable—that is, a search start with the high-validity cue (correct response), the medium validity cue (incorrect response), the low-validity cue (incorrect response), or no cues searched (incorrect response).

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

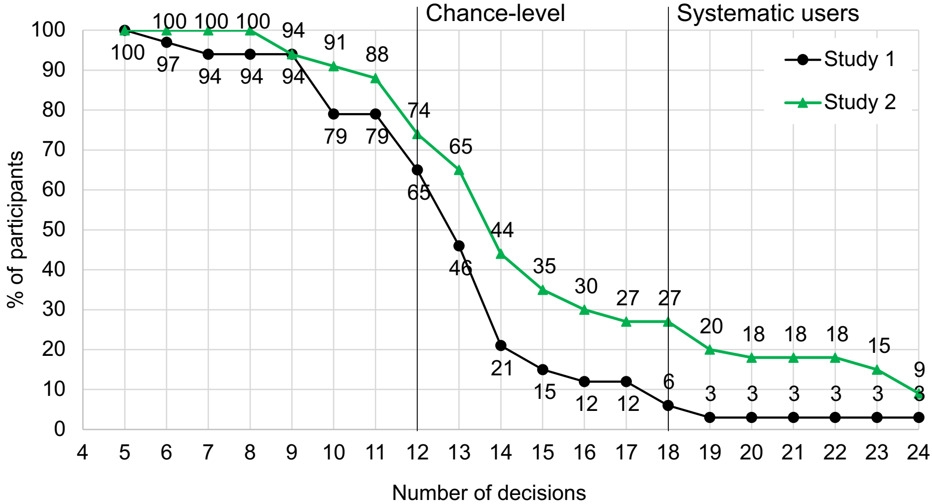

Figure 4. *Percentage of participants who followed the high-validity cue in at least i (5, . . ., 24) decisions (cumulated percentages) in Studies 1 and 2.*

Taken together, preschoolers showed a first tendency to prioritize the high-validity cue in their search. Such a tendency has not been reported for the current paradigm thus far (e.g., Betsch et al., 2020; see Hahne et al., 2015 for a meta-analysis). However, results also indicate preschoolers' difficulty in weighting. Preschoolers also frequently searched less valid cues.

2.2.4. Choice

Although preschoolers did not clearly prioritize the high-validity cue in their information search (e.g., suggested by TTB), they could have still used cue validities for choice (e.g., suggested by a weighted-additive strategy: Payne et al., 1988). Following Betsch et al. (2020), we assessed how often preschoolers followed the prediction of the high-validity cue (by choosing the predicted option). On average, children followed the high-validity cue in 12.3 of their 24 decisions (i.e., 51.3%, SD = 14.5, Bootstrap 95% CI [46.6, 56.7]). This is not significantly above a chance level of 50%[2] (one-sample *t*-test, one-tailed, test-value 50; $t(32) = 0.50$; $p = .310$, $d = 0.087$). Also, it does not significantly differ from how often they followed the predictions of the other 2 cues (low-validity cue: M = 48.7%, SD = 14.5, CI [44.0, 53.9], medium validity cue: M = 49.2%, SD = 20.2, CI [42.1, 56.4], no main effect of cue in a repeated-measures analysis of variance: $F(1.5, 47.3) = 0.17$, $p = .778$, $\eta^2 = .01$). Thus, our findings indicate that preschoolers do not systematically prioritize the high-validity cue in their choices.

2.2.5. Individual-level analysis

2.2.5.1. Systematic cue users

The preceding analysis on the group level may cover developmental differences between children of preschool age. In order to obtain further insights, Figure 4 shows results on the individual level. It displays, for the respective number of decisions, the percentage of participants who followed the high-validity cue in at least that number of decisions. Sixty-five percent of the preschoolers followed the prediction of the high-validity cue in at least 12 decisions, that is, half of their decisions. Above this chance level, the number of participants decreases markedly. That is, only few participants follow the high-validity cue in substantially more than chance level. To identify 'systematic users of cue validity' from this data, Betsch et al., (e.g., 2014, 2020) introduced a threshold: If at least 75% of a participant's

[2]The used variable is which option participants chose. Chance level is set to 50% because participants could choose between two options—that is, the house predicted by the high-validity cue (correct response) or the other house (incorrect response).

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

decisions (i.e., at least 18 decisions) are in line with the high-validity cue, they infer systematicity. As evident from Figure 4, in the current study, only 6.1% of participants ($N = 2$) were systematic users according to that criterion. This is not significantly different from chance-level [binomial test: hypothesized p1 = 16.4%; exact binomial p (one-tailed) = .075]. Taken together, these findings indicate that preschoolers did not prioritize the high-validity cue in their choices.

2.2.5.2. Choice strategies

As the previous analysis indicates, almost no participants in our sample systematically followed the high-validity cue in their decisions. To investigate what other strategies, if any, participants used to make their choices, Betsch et al. (2020) applied an outcome-based method for strategy classification (Bröder and Schiffer, 2003). These analyses should be treated with caution, because the pattern types of the tasks were not designed for identifying strategies. We conducted these exploratory analyses in order to obtain rough insights into the underlying processes of preschoolers' decision-making. Following the procedure of Betsch et al. (2020), we considered 5 potential strategy candidates (other common strategies could not be considered, because tasks did not distinguish between them). The take-the-best strategy (TTB: Gigerenzer and Gaissmaier, 2011) is similar to our prior measure. TTB only considers the predictions of the high-validity cue and follows it. The equal weight (EQW, Payne et al., 1988) strategy ignores cue validities by inspecting all cues and counting positive predictions (i.e., treasures) for each option. The option with the highest number of positive predictions is chosen. In the case of ties, a random choice is made. The take-the-first (TTF) strategy also ignores cue validities but reflects a preference for a selective search aligned with reading habits. Here, only the predictions of the first cue at the top of the information board are considered and followed (i.e., the predictions of the least valid cue). A switching (SW) strategy ignores not only cue validities but also cue predictions. Here, participants simply alternate choices from trial to trial (i.e., choose the left and right house in turns). Lastly, we classified random choices.

For classification, first, we checked whether participants' choices perfectly fit 1 of the 5 strategies. Participants were classified accordingly. Under the assumption that application errors might occur, we determined the likelihood of the observed choice pattern for each strategy for the remaining participants. We classified participants as a user of the strategy with the highest likelihood if the associated error probability was less than 0.5 (i.e., above chance level). Otherwise, the classification was considered unreliable; and participants remained unclassified. Error rates for the strategies are shown in Table 4. The classification procedure permits variation between participants but assumes a uniform probability distribution of making errors within participants. Note, that we applied loose classification criteria. Others have, for example, used stricter error rates or applied additional criteria such as considering the sum of the likelihoods of the remaining strategies (e.g., Glöckner and Betsch, 2008). Thus, again, our classification results only provide a rough direction of what children used to guide their choices in order to explanatorily inspire future research endeavors.

Table 4 shows the classification results. The least applied strategy was random choice, which indicates that preschoolers did not make their choices in random fashion. TTB, EQW, and TTF, were applied by only few participants (Table 4). Concerning TTB, this finding matches our previously reported results on systematic cue users. Similar to the findings of Betsch et al. (2020), the dominant strategy was SW. Along the lines of 'each one gets his turn' (Piaget and Inhelder, 1951/1975), around 30 % of the participants chose each option in turn regardless of the cue predictions. This can, for example, be exploratory behavior and originate from the consideration of option reinforcement. In our tasks, options were equally reinforced across trials. That is, overall, the houses contained an equal number of treasures. Preschoolers might have tracked the distribution of the treasures in the houses, which, in turn, could have invited SW. Thus, the application of this strategy may indicate the consideration of option probabilities. Such a procedure, however, is maladaptive, because it neglected the fact that the high-validity cue made certain predictions regarding the location of the treasure. Taken together, the results of strategy classification indicate that preschoolers do not systematically use cue validities as weights in decision-making. However, the results also suggest that their decision-making is likely not

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

Table 4. *Results of strategy classification of Study 1.*

	n	%	Error
TTB	3	9.1	.31
EQW	3	9.1	.40
TTF	3	9.1	.36
SW	10	30.3	.30
Random	1	3.0	
Unclass.	13	39.4	
Total	33		

Note: Number of participants (*n*), percentage of participants, and classification error (i.e., mean proportion of trials in which strategy inconsistent choices were observed) for the strategies. EQW, equal weight; TTB, take-the-best; TTF, take-the-first; Random, random choice; SW, switching between options; Unclass., participants who could not be classified to one of the considered strategies.

random. Note, that a substantial portion of participants remained unclassified (Table 4). This could happen, for example, when strategy application was too imprecise or when participants applied a different strategy that we did not consider in our preselection. In addition, the method we applied for strategy classification cannot detect when participants change their procedure across trials. That is, we are unable to classify the combination of multiple strategies. Betsch et al. (2018) showed that children used a number of strategies (toolboxes) and switched between them over a series of trials. That is, considering the test phase as a whole, they tried out different procedures (systematically) instead of consistently following the high-validity cue. Such an approach could be the cause of our unclassified results.

2.2.6. Pattern effect

In previous studies, the so-called pattern effect was analyzed as an additional indicator of weighting competence. Normatively, the specific pattern of predictions (Figure 3) should have no effect on choice. However, previous studies repeatedly demonstrated a pattern effect in participants' decision behavior—that is, participants had difficulties with pattern Type 3 when both less valid cues consistently contradicted the high-validity cue (e.g., Betsch et al., 2020; Betsch and Lang, 2013). In contrast, in the current study no such pattern effect was found (repeated-measures analyses of variance for the frequency of choices that follow the high-validity cue with pattern type as a within-subjects factor: $F(2, 64) = 1.43$, $p = .248$, $\eta^2 = .04$). As would be normatively expected, children's performance was comparable for the 3 types of prediction patterns. Specifically, of the 8 repetitions of each pattern type, decisions followed the high-validity cue in an average of 4.4 decisions (SD = 1.7) in Type 1 patterns, 3.8 decisions (SD = 1.7) in Type 2 patterns, and 4.1 decisions (SD = 1.6) in Type 3 patterns.[3] Because the pattern effect is a robust finding with the applied paradigm (see Hahne et al., 2015 for a meta-analytic review), results of the current study may suggest an easier handling of the choice situation in the current study. When a nonprobabilistic cue is involved, the intrusion of less valid cues appears to be independent of the specific prediction pattern.

2.2.7. Learning effect

To investigate learning effects, we inspected potential changes across trials. Replicating prior studies (e.g., Betsch et al., 2018, 2020, 2021), we found only slight changes between the first and second half of

[3]To assess the pattern effect in the original study, Betsch et al. (2020) calculated the rate of systematic cue users separately for each pattern type. The analysis of this measure yields comparable results. For easier comprehension, we report the frequency of choices that follow the high-validity cue.

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

Table 5. *Number of choices in which children followed the high-validity cue in Study 1.*

	M (SD)	CI	Range
First half	5.8 (1.8)	[5.2, 6.4]	1–12
Second half	6.5 (2.2)	[5.7, 7.4]	1–12
Block 1 (trials 1–6)	2.9 (1.1)	[2.5, 3.3]	1–6
Block 2 (trials 7–12)	2.9 (1.2)	[2.5, 3.3]	1–6
Block 3 (trials 13–18)	3.5 (1.6)	[2.9, 4.0]	1–6
Block 4 (trials 19–24)	3.1 (1.2)	[2.6, 3.5]	1–6

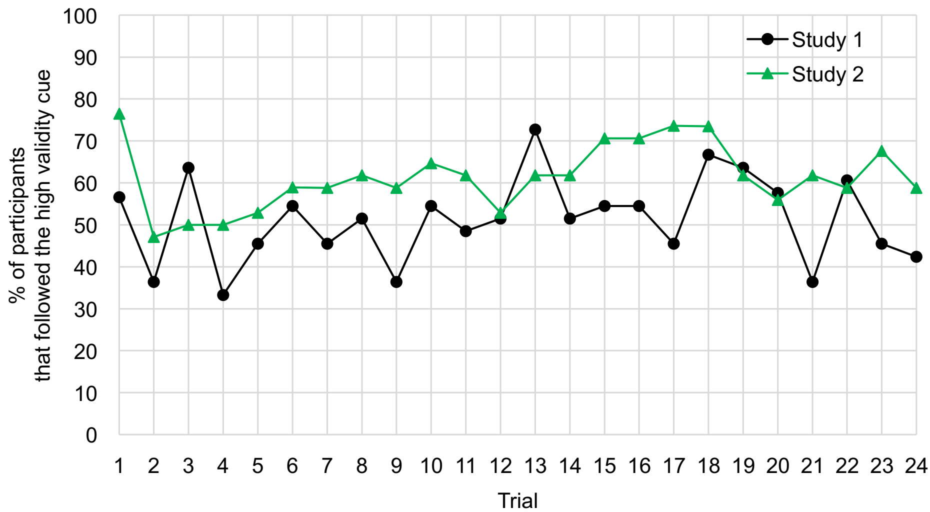

Figure 5. *Percentage of participants that followed the high-validity cue in each of the 24 trials in Studies 1 and 2.*

trials in terms of the number of choices in which children followed the high-validity cue (Table 5). We found a slight tendency toward improvement (paired-sample t-test, one-tailed: $t(32) = 2.00$, $p = .029$, $d = 0.34$). Following up on this, one may argue that our 2 practice trials were not sufficient for preschoolers to become accustomed to the game. Their first test trials may represent them navigating through the game rather than real choices. Our preceding analyses that collapsed all trials would then underestimate their performance. If this were the case, one would expect their performance to increase between their first and last choices in the study. In order to test for such a learning effect, we split trials in 4 blocks of 6 trials each (including 2 trials of each pattern type, Table 5). Performance in the first and last block did not differ substantially (paired-sample t-test, one-tailed: $t(32) = 0.67$, $p = .25$, $d = 0.12$). Figure 5 that shows the rate of preschoolers (y-axis) that followed the high-validity cue on each of the 24 trials (x-axis) yields the same conclusion. If preschoolers profited from experience in the game, one would expect an increase over trials. However, the average rate of preschoolers who followed the high-validity cue obviously remained quite stable. In sum, preschoolers did not appear to systematically profit from feedback after choice. That is, although feedback reinforced the validities of the cues, preschoolers did not learn to follow the predictions of the high-validity cue over trials (note that Betsch et al., 2018 report similar results for a test session with 80 instead of 24 test trials). Taken together, our findings do not indicate a reliable learning effect.

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

3. Study 2

Contrary to our expectation, in Study 1 we found only an initial tendency of preschoolers to prioritize the high-validity cue in their decision-making. This may be due to the involvement of probabilistic cues. In Study 1, although the high-validity cue was clearly superior, the other, less valid cues were still occasionally correct. This might have led preschoolers to perceive the overall decision environment as probabilistic, which, in turn, might have triggered a low sensitivity for validity differences in general. Thus, in Study 2, we assessed decisions under certainty. That is, one cue was always correct, whereas the other 2 cues were always incorrect ($p_1 = 1$, $p_2 = .0$, $p_3 = .0$). This structure represents the most extreme form of dispersion in cue validities and thus should promote the psychological processes of prioritization and inhibition.

3.1. Method

3.1.1. Participants

Thirty-five 5- to 6-year-olds participated. One child did not finish the test session and was excluded from data analysis ($N = 34$; 47% female; age: Mdn = 74.5 months, M = 73.4, SD = 6.33). Sample size (aimed $N > 30$) allows for detecting moderate effects with $\alpha = .05$ and at least $1 - \beta \geq .85$ in the conducted tests (i.e., repeated-measures analysis of variance, one-sample and one-tailed *t*-test, paired-sample and one-tailed *t*-test, one-sample and two-tailed binomial test; a priori power analysis conducted with G*Power, Faul et al., 2007). Children (all native German speakers) were recruited in 4 daycare centers located in middle-class areas of 2 moderately large cities in Germany. Parents had previously provided written consent. Participants received a personalized certificate with a photo for participation. This research was approved by the research ethics board of the University of Erfurt; project title is 'Probabilistic decisions in children'.

3.1.2. Material and procedure

We used the same material and procedure as in Study 1. All modifications concern the change of the less valid cues into nonprobabilistic cues whose predictions were always incorrect (cue validities: 100%, 0%, and 0%). Specifically, the video material of the learning session was rearranged to increase the rate of incorrect predictions. Additionally, only pattern Type 3 could be used in the test session. Pattern Types 1 and 2 could not be applied since, in decisions between 2 options, it is logically impossible that the incorrect cues simultaneously are both incorrect and contradict each other.

3.2. Results and discussion

Anonymized data are publicly available at the OSF project page: https://doi.org/10.17605/OSF.IO/HE7ZR; https://osf.io/he7zr/?view_only=184f9b892ea74db89b9b5d914fde9dd1.

3.2.1. Manipulation check of cue validity learning

All participants passed the manipulation check by responding that the high-validity cue was correct most often in the learning session (i.e., no data exclusion, $N = 34$). The data set we used for analysis included all 34 participants.

3.2.2. Data analysis

Data analysis follows Study 1 with the exception that results on the pattern effect cannot be calculated. In Study 2, only one pattern type was applied due to logical constraints in decisions under certainty (i.e., with all cues, either completely correct or incorrect). Due to this reason, we also could not conduct the strategy classification of Study 1, which relied on these pattern types.

3.2.3. Cue evaluations

Nearly 75% of preschoolers preferred the high-validity cue in their evaluations in different domains—even more than in Study 1 (see average across domains in Table 6; nonoverlapping confidence intervals

Table 6. *Cue evaluations in Study 2.*

Domain	High-validity cue	Incorrect middle*	Incorrect top**	All cues	None	Other
Labeling	82.4 (28)	8.8 (3)	8.8 (3)	0	0	0
Knowledge	76.5 (26)	8.8 (3)	14.7 (5)	0	0	0
Creativity	70.6 (24)	17.6 (6)	8.8 (3)	2.9 (1)	0	0
Prosocial behavior	70.6 (24)	5.9 (2)	17.6 (6)	2.9 (1)	2.9 (1)	0
Friendship	67.6 (23)	8.8 (3)	2.9 (1)	5.9 (2)	14.7 (5)	0
Treasure	79.4 (27)	14.7 (5)	2.9 (1)	0	0	2.9 (1)
Averaged across domains	74.5	10.7	9.3	2.0	2.9	0.5

Note: The table shows the percentage (number) of participants who responded with the respective category. The response category '2 cues' did not occur and is thus not included in the table.

*Incorrect cue, middle row;

**Incorrect cue, top row.

in Table 3). This is substantiated by a large main effect of response category on frequencies of cue evaluations in an analysis of variance with repeated measurements, $F(3, 99) = 81.85$, $p < .001$, $\eta^2 = .71$. Thus, this is further evidence suggesting that preschoolers consider cue validity when evaluating cues.

3.2.4. Information search

Replicating results from Study 1, preschoolers frequently searched the high-validity cue first (M = 40.0% of decisions, (CI [27.3, 52.8], SD = 38.2) and rarely searched the cue positioned at the middle first (M = 7.2% of decisions, CI [3.8, 10.9], SD = 11.0) or no cue (M = 3.2% of decisions, CI [0.5, 7.6], SD = 11.1). Their frequent search starts with the high-validity cue is significantly above a chance level of 25%[4] (one-sample *t*-test, one-tailed, test-value 25; $t(33) = 2.28$; $p = .015$, $d = 0.391$). However, preschoolers most frequently began search with the incorrect cue positioned at the top of the information board (M = 50.0% of decisions, Bootstrap 95% CI [37.5, 61.5], SD = 36.4). In addition, although preschoolers' search was not as exhaustive as in Study 1, they still frequently searched the incorrect cues. On average, they searched 4.5 (CI [3.9, 5.2], SD = 1.9) pieces of information per decision. Taken together, results are comparable to Study 1, with preschoolers showing only an initial tendency to prioritize the high-validity cue in search.

3.2.5. Choice and individual-level analysis

In contrast to Study 1, on average, preschoolers followed the high-validity cue in 14.7 of their 24 decisions (i.e., in 61.3%, SD = 19.9, Bootstrap 95% CI [54.9, 68.0]). This is significantly above a chance level of 50% (one-sample *t*-test, one-tailed, test-value 50%; $t(33) = 3.31$; $p = .001$, $d = 0.568$). Also, with this, preschoolers followed the high-validity cue more often than the prediction of the other 2 cues (M = 38.8%, SD = 19.9, CI [31.8, 45.7], paired-sample *t*-test, one-tailed: $t(33) = 3.31$, $p = .001$, $d = 0.57$).

Figure 4 shows the results on the individual level. As shown, more preschoolers in Study 2 performed above chance-level than in Study 1. Also, with 27% ($N = 9$), the rate of systematic cue users is above that of Study 1. This is by far the highest user rate of preschoolers reported for the applied paradigm thus far (e.g., meta-analytic review by Hahne et al., 2015: 4% of preschoolers, $N = 826$; 21% of elementary schoolers, $N = 689$). Although this rate substantially exceeds results from prior studies, the majority of preschoolers still failed to systematically use cue validities for choice. The rate of systematic cue users is only marginally above chance level and is not statistically significant (according to a binomial test: hypothesized $p_1 = 16.4\%$; exact binomial p (one-tailed) = .093).

[4]For explanation of chance levels, see footnotes in the results section of Study 1.

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

Table 7. *Number of choices in which children followed the high-validity cue in Study 2.*

	M (SD)	CI	Range
First half	7.0 (2.6)	[6.4, 7.8]	1–12
Second half	7.8 (3.1)	[6.8, 8.8]	1–12
Block 1 (trials 1–6)	3.4 (1.5)	[2.9, 3.9]	1–6
Block 2 (trials 7–12)	3.6 (1.6)	[3.1, 4.1]	1–6
Block 3 (trials 13–18)	4.1 (2.0)	[3.5, 4.4]	1–6
Block 4 (trials 19–24)	3.6 (1.9)	[3.0, 4.3]	1–6

Taken together, even in a certain choice environment involving always correct and always incorrect cues, preschoolers show an initial tendency to prioritize high-validity cues for choice but also difficulties in doing so.

3.2.6. Learning effects

Similar to Study 1, there were only slight changes between the first and second half of trials in terms of the number of choices in which children followed the high-validity cue. According to the descriptives in Table 7, we found a slight tendency toward improvement that is not significant (paired-sample t-test, one-tailed: $t(33) = 1.51$, $p = .069$, $d = 0.26$). In line with this finding, preschoolers' performance did not substantially increase between their first and last block of 6 trials (Table 7, paired-sample t-test, one-tailed: $t(33) = 0.82$, $p = .21$, $d = 0.14$). Also, the rate of preschoolers who followed the high-validity cue on each of the 24 trials remained quite stable (Figure 5). Comparable to Study 1, preschoolers thus did not learn to follow the predictions of the high-validity cue over trials.

4. General discussion

A child's world is full of cues that can be used to enhance decision-making. For instance, children can ask parents, friends, or caregivers for predictions concerning the weather or the location of misplaced items. An adaptive decision-maker should consider the fact that such cues may differ in terms of how often they are correct. The current research was conducted to take a closer look at preschoolers' ability to use cue validities as weights in decision-making, resulting in prioritization and inhibition of cues according to their validity. It was based on frequent findings with JDM paradigms that highlight children's difficulties (e.g., Betsch et al., 2014, 2016, 2020). Our studies investigated the idea that the fallibility of the high-validity cue may trigger exploration approaches. The involvement of probabilistic cues might go along with equalized 'unsure' weights for all cues in order to further explore the novel environments used in JDM paradigms. We applied a noncompensatory environment in which prioritization of the high-validity cue and inhibition of the less valid cues is highly adaptive (e.g., by applying TTB). However, if children applied equalized weights, this would change the decision environment (subjective cue validities, e.g., lens model: Brunswik, 1952). On the behavioral level, children's decision-making would appear to be unweighted. However, this conclusion might be an underestimation of children's weighting competency because this may correctly reflect their subjective mental representation of the decision environment. To tackle this issue, in our 2 conceptual replications of the probabilistic inference study by Betsch et al. (2020), we used a nonprobabilistic, high-validity cue and accompanied it with less valid, probabilistic cues (Study 1). In Study 2, we applied the most extreme form of dispersion in cue validities (i.e., 0 and 1) in order to promote the psychological processes of prioritization and inhibition.

Contrary to our expectations, we found that even with our nonprobabilistic, high-validity cue, preschoolers had difficulties using cue validities for decision-making. Although in our studies the high-validity cue was clearly superior because it was consistently correct in the learning session, preschoolers frequently also searched and followed the predictions of the less valid cues. In line with these findings, our exploratory strategy classification results suggested that preschoolers refrained from using the TTB strategy, which would be an appropriate heuristic in our noncompensatory environment. We also found this insufficient prioritization in search and choice in completely certain decision situations in which the less valid cues had proven themselves to be completely incorrect (i.e., in Study 2). Thus, preschoolers' difficulty does not appear to be limited to probabilistic situations in which the fallibility and occasional correctness of cues might trigger exploration by equal weighting. This finding is remarkable because our results on cue evaluations show a different picture. Replicating Betsch et al.'s (2020) findings concerning this measure, our participants more frequently trusted the high-validity cue as opposed to the less valid cues. That is, in their cue evaluations, participants preferred the high-validity cue. This indicates proper learning and discrimination of cue validities. These findings are underpinned by research suggesting that the frequency and cooccurrence of events are encoded automatically with little effort (overview: Sedlmeier, 2002). Moreover, probabilistic thinking was shown to be well developed at an early age (e.g., Denison and Xu, 2014; Koenig et al., 2019; Sternberg, 2009; Zmyj et al., 2010). However, when it comes to decision-making, cue validities appear to be lost. Although, our findings indicate that the development toward a decision-maker who can prioritize high-validity cues is well on its way at preschool age, preschoolers considered cue validities in decision-making in a rudimentary form only. Still, it should be noted that our studies are the first in the study battery with the applied paradigm to show initial prioritization in preschoolers. Comparable tendencies of prioritization have been reported to not occur before school age (i.e., around 9 years, e.g., Betsch et al., 2020; overview: Hahne et al., 2015). Future research should further investigate developmental trajectories. In our studies, the sample was restricted to preschoolers. For school-aged children, who have been shown to begin weighting with cue validities (e.g., Betsch et al., 2020), it might indeed be beneficial if the decision involves nonprobabilistic cues. Older children with more mature cognitive abilities might show more matured weighting in a certain world.

Concerning our current studies, one could argue that a lack of game understanding caused our findings. We consider this unlikely based on an extensive comprehension study that demonstrated children's game understanding (see the discussion section in Betsch et al., 2020 for results). In this context, note that the applied information board differed slightly from adult research. In research with adults, cues are commonly ordered with decreasing validity from the top to bottom row (e.g., Söllner et al., 2014). We deviated from this practice in order to avoid a confound of reading habits and cue validity, which would prevent any conclusions on children's weighting with cue validities. Pretests had shown that a task-by-task randomization of the cue order hampered encoding of the information board. To enhance the overall handling of the game, we decided to fix the ordering, with the high-validity cue positioned at the bottom. This is the strongest test of our research question, since reading habits invite participants to open the upper rows (Klayman, 1985). Our finding that children so frequently considered the inviting top of the information board implies difficulty in systematically using cue validities. At the same time, our findings also suggest that the display of information might be an environmental condition with the potential to facilitate the use of the high-validity cue. However, note that the mere use of the high-validity cue is different from a purposeful consideration of cue validity in decision-making.

As an alternative explanation for our findings, one might also argue that preschoolers did not generalize the cue validity they had learned to the test session and, thus, did not have applicable cue validities available for choice. After all, a labeling task differs from a treasure hunt. In a decision environment with unknown cue validities, the SW strategy that we found in Study 1 may be an appropriate way to explore the environment. Such a procedure, however, was maladaptive in our studies, because it neglected the fact that the high-validity cue made certain predictions regarding the treasure location. Future research should further investigate whether a failed generalization from the learning to the test session drives preschoolers' failure to use cue validities. Indeed, research suggests that children distinguish between domain-relevant and domain-irrelevant expertise (overview: Sobel

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

and Kushnir, 2013). However, this research also demonstrates that children generalize the allocation of expertise to conceptually clustered domains. The instructions of the current study were carefully designed and pretested to make use of this. The objects in the labeling task were children's games, in order to conceptually match the treasure hunt game. The learning instances in the labeling task were called 'being correct points' ('Rechthabepunkt' in the German instructions). This refers to a general trait that covers the test instances equally well. In addition, the experimenter directly associated the results of the learning session with the test session by explanation and, thus, set a strong demand (e.g., by repeatedly explaining that, since the children know how often the cues were correct when labeling toys, they also know how well the cues can help in the treasure hunt game). Finally, the association was visualized by the 'being correct points' that appeared in both sessions alike next to the cues. Our result that preschoolers consider cue validities in their cue evaluations indicates that our attempts to foster generalization were successful. It appears rather unlikely that children showed a 'halo effect' (i.e., an impression that unjustifiably affects the overall conception: Buttelmann et al., 2013) for domains such as creativity, knowledge, and prosocial behavior but excluded specifically the treasure hunt from this generalization.

The study of Mata et al. (2011), in which a similar information-board decision game to ours is applied, suggests the presence of longer-lasting difficulties in children. The authors showed that, even for 9- to 10-year-olds, it is not trivial to systematically apply decision strategies that rely on weighting. Similar difficulties with probabilistic weights were, for example, obtained in studies on risk learning (e.g., Beitz et al., 2014) and expected value integration (e.g., Levin et al., 2007; Levin and Hart, 2003; Schlottmann and Wilkening, 2012). Our findings are also consistent with findings using preferential choice tasks that investigate the use of nonprobabilistic, relevance weights in decision-making. For example, in information-board decision games that involve preferential choices between bicycles or piggy banks, children repeatedly showed rather unsystematic decision-making that was not systematically guided by the relevance of information (e.g., Davidson, 1991; Lindow and Betsch, 2018). Following the assumptions of the influential adaptive decision-making framework (e.g., Gigerenzer and Gaissmaier, 2011), Mata et al. (2011) described this phenomenon as 'when easy comes hard'. Accordingly, heuristic strategies that decision-makers possess to simplify the decision may be a challenge, even in late childhood, when they draw on the ability to selectively attend to relevant information. To not 'simply' reduce decision-making to the most relevant information, for example, by prioritizing the high-validity cue, and instead showing broader exploration in search and choice could well be reasonable for children. According to adaptive decision-making models, experience and learning should be of great importance for the development of decision-making. For example, it is assumed that strategy knowledge differs between individuals (Beach and Mitchell, 1978) and heuristics are acquired with increasing age (Gigerenzer, 2003). Proposed mechanisms are learning instances where decision-makers imitate others, design new strategies from scratch, or construct strategies from preexisting strategies (Gigerenzer, 2003). Moreover, Rieskamp and Otto (2006) propose a learning theory as an answer to the strategy selection problem. Through reinforcement learning of the strategies' past performances, decision-makers learn to assess which strategy will perform well. Such strategy expectancies are assumed to enable successful adaptive decision-making later on. Thus, from an adaptive decision-making perspective, one could argue that trying out different strategies is advantageous in order to develop a better understanding of those strategies and their fit to specific task structures. Even in alternative decision models that assume a one-purpose rule instead of multiple strategies for decision-making, learning is no less important (e.g., Glöckner and Betsch, 2008; Jekel et al., 2018). These models are based on the idea that different decision strategies are not necessary because automatic processes effortlessly provide weighted integration in an information network. However, the quality of this integration process depends on the fit of the mental network with the actual decision situation. Appropriate automatic weighting is only possible if the decision-maker has sufficient and adequate knowledge about the world. The observation that children decide 'without any systematic plan' (7- and 8-year-olds: Davidson, 1996, p. 264) could be due to the higher order plan to learn about the structure of a, for children, still widely novel world. Thus, following evolutionary developmental

https://doi.org/10.1017/jdm.2023.36 Published online by Cambridge University Press

psychology (Bjorklund, 2018), children's immaturity (according to normative standards) might not be a cognitive limitation but, rather, a developmental necessity. Accordingly, trying out and experiencing the consequences is reasonable for maturation (see, e.g., Bjorklund, 2018, for the functionality of play for development). This could explain our findings that the effective learning and discrimination of cue validities that obviously go along with adequate cue evaluations are insufficient for their efficient use as weights in decision-making. This is an important finding, because it suggests that weighting—as one of the core processes of decision-making (He et al., 2022)—is a competence that decision-makers develop. With this, our research suggests that preschoolers' weighting abilities reported in previous work using familiar, experience-based contexts should not be mistaken as a generalizable competence (e.g., Koenig et al., 2019; Koenig and Harris, 2005; Lindow and Betsch, 2021; Pasquini et al., 2007). Using the same words as the best speaker (e.g., Pasquini et al., 2007) or accepting help from the best problem-solver at daycare (Lindow and Betsch, 2021) might not be the same as systematically weighting cues by their validity in decision-making. Everyday functionality could be caused by other, experienced-based mechanisms that are frequently subsumed under the label 'intuitive' processes (e.g., Löckenhoff, 2018; Schlottmann and Wilkening, 2012). Thus, valuing specific cue characteristics such as friendship (e.g., Harris, 2007) to tackle the obstacles of everyday life may be a skill available to children at an early age—but not necessarily a demonstration of their competence to systematically use cue validities in decision-making.

5. Conclusion and practical implications

Our 2 conceptual replication studies strengthen the current state of decision research that suggests that preschoolers have difficulty in systematically using cue validities in decision-making. Prioritizing and inhibiting cues according to their validity in information search and choice appeared to be a challenge. Our research rules out the possibility that this is restricted to a risky world with probabilistic cues and, instead, suggests a fundamental weighting deficit. However, regarding the rapidly growing amount of information that we are exposed to as digitalization and globalization progresses, this is an urgently needed competence. Decision-making should be guided by valid information. The applied lesson to be learned from our studies is twofold. First, systematically collecting information from cues according to their validity is a competence that decision-makers must develop. Second, our research suggests that an unfiltered flood of information represents a wicked decision environment for preschoolers. If the environment also provides information with low validity, preschoolers may well follow it. Depending on the specific situation, this can be unfavorable and even dangerous, as it makes one susceptible to misinformation. This would imply the need to protect preschoolers from fundamentally severe and threatening consequences of suboptimal decisions. One could argue that this recommendation overstretches the results of the current research. After all, in our studies, disadvantageous decision-making 'only' resulted in finding fewer treasures. Thus, our findings might not generalize to decisions with threatening consequences. However, for one, it is likely that preschoolers do not always anticipate the full range of consequences—that is, they might not be reliably able to recognize such decision situations. Second, comprehension tests of the applied paradigm clearly suggested that preschoolers were highly motivated to find the treasures. Thus, it is quite possible that our findings generalize to more important decision situations.

Data availability statement. Anonymized data and materials have been made publicly available at the OSF and can be accessed at the OSF project page: https://doi.org/10.17605/OSF.IO/HE7ZR.

Acknowledgments. We would like to thank Lisa Büchler, Lisa Schirmner, Verena Sonntag, and Rachel Stenmans for their help with material development and data collection as part of a bachelor seminar. We are grateful to Heather Fiala for copy-editing.

Funding statement. This work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) under Grant BE 2012/11-2 and project number 512467589 with no content-related involvement.

Competing interest. The authors declare none.

References

Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *The Academy of Management Review*, *3*, 439–449. https://doi.org/10.2307/257535

Beitz, K. M., Salthouse, T. A., & Davis, H. P. (2014). Performance on the Iowa Gambling Task: From 5 to 89 years of age. *Journal of Experimental Psychology: General*, *143*, 1677–1689. https://doi.org/10.1037/a0035823

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. https://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. https://doi.org/10.1016/j.actpsy.2014.09.008

Betsch, T., Lehmann, A., Jekel, M., Lindow, S., & Glöckner, A. (2018). Children's application of decision strategies in a compensatory environment. *Judgment and Decision Making*, *13*, 514–528. http://journal.sjdm.org/18/18225/jdm18225.pdf

Betsch, T., Lehmann, A., Lindow, S., & Buttelmann, D. (2020). Children's trust in advice-givers in risky decisions. *Cognitive Development*, *53*, 100846. https://doi.org/10.1016/j.cogdev.2019.100846

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*(2), 311–325. https://doi.org/10.1037/dev0000077

Betsch, T., Lindow, S., Lehmann, A., & Stenmans, R. (2021). From perception to inference: Utilization of probabilities as decision weights in children. *Memory and Cognition*, *49*, 826–842. https://doi.org/10.3758/s13421-020-01127-0

Betsch, T., Wünsche, K., Großkopf, A., Schröder, K., & Stenmans, R. (2018). Sonification and visualization of predecisional information search: Identifying toolboxes in children. *Developmental Psychology*, *54*, 474–481. https://doi.org/10.1037/dev0000447

Bhatia, S., Loomes, G., & Read, D. (2021). Establishing the laws of preferential choice behavior. *Judgment and Decision Making*, *16*, 1324–1369.

Bjorklund, D. F. (2018). A metatheory for cognitive development (or "Piaget is dead" revisited). *Child Development*, *89*, 2288–2302. https://doi.org/10.1111/cdev.13019

Blumberg, F. C., Torenberg, M., & Randall, J. D. (2005). The relationship between preschoolers' selective attention and memory for location strategies. *Cognitive Development*, *20*, 242–255. https://doi.org/10.1016/j.cogdev.2005.04.006

Bröder, A., Dülz, E., Heidecke, D., Wehler, A., & Weimann, F. (2023). Improving carbon footprint estimates of food items with a simple seeding procedure. *Applied Cognitive Psychology*, *37*, 651–659. https://doi.org/10.1002/acp.4060

Bröder, A., & Schiffer, S. (2003). Bayesian strategy assessment in multi-attribute decision making. *Journal of Behavioral Decision Making*, *16*, 193–213. https://doi.org/10.1002/bdm.442

Brunswik, E. (1952). *The conceptual framework of psychology*. Chicago, IL: University of Chicago Press.

Buttelmann, D., Zmyj, N., Daum, M., & Carpenter, M. (2013). Selective imitation of in-group over out-group members in 14-month-old infants. *Child Development*, *84*, 422–428. https://doi.org/10.1111/j.1467-8624.2012.01860.x

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

Davidson, D. (1996). The effect of decision characteristics on children's selective search of predecisional information. *Acta Psychologica*, *92*, 263–281. https://doi.org/10.1016/0001-6918(95)00014-3

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

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 1–547. https://doi.org/10.1146/annurev-psych-113011-143750

Ebersbach, M. (2009). Achieving a new dimension: Children integrate three stimulus dimensions in volume estimations. *Developmental Psychology*, *45*, 877–883. https://doi.org/10.1037/a0014616

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Fechner, H. B., Pachur, T., & Schooler, L. J. (2019). How does aging impact decision-making? The contribution of cognitive decline and strategic compensation revealed in a cognitive architecture. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*, 1634–1663. https://doi.org/10.1037/xlm0000661

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.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*, 451–482. https://doi.org/10.1146/annurev-psych-120709-145346

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*(4), 506–528. http://doi.org/10.1037/0033-295X.98.4.506

Glöckner, A., & Betsch, T. (2008). Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision-making. *Judgment and Decision-making*, *3*, 215–228.

Hahne, A. S., Lang, A., & Betsch, T. (2015). *Risky decision-making in children: A meta-analysis* (Unpublished master's thesis). Erfurt, Germany: University of Erfurt.

Harris, P. L. (2007). Trust. *Developmental Science*, *10*, 135–138. https://doi.org/10.1111/j.1467-7687.2007.00575.x

He, L., Zhao, W. J., & Bhatia, S. (2022). An ontology of decision models. *Psychological Review*, *129*, 49–72. https://doi.org/10.1037/rev0000231

Horn, S., Ruggeri, A., & Pachur, T. (2016). The development of adaptive decision making: Recognition-based inference in children and adolescents. *Developmental Psychology*, *52*(9), 1470–1485. https://doi.org/10.1037/dev0000181

Jekel, M., Glöckner, A., & Bröder, A. (2018). A new and unique prediction for cue-search in a parallel-constraint satisfaction network model: The attraction search effect. *Psychological Review*, *125*(5), 744–768. https://doi.org/10.1037/rev0000107

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

Koenig, M., Tiberius, V. & Hamlin, J. K. (2019). Children's judgments of epistemic and moral agents: From situations to intentions. *Perspectives in Psychological Science*, *14*, 344–360. https://doi.org/10.1177/1745691618805452

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, *76*, 1261–1277. https://doi.org/10.1111/j.1467-8624.2005.00849.x

Levin, I. P., & Hart, S. S. (2003). Risk preferences in young children: Early evidence of individual differences in reaction to potential gains and losses. *Journal of Behavioral Decision Making*, *16* (5), 397–413. https://doi.org/10.1002/bdm.453

Levin, I. P., Weller, J. A., Pederson, A. A., & Harshman, L. A. (2007). Age-related differences in adaptive decision making: Sensitivity to expected value in risky choice. *Judgment and Decision Making*, *2*, 225–233. https://doi.org/10.1080/03601277.2014.900263

Lindow, S., & Betsch, T. (2018). Child decision-making: On the burden of pre-decisional information search. *Journal of Cognition and Development*, *19*, 137–164. https://doi.org/10.1080/15248372.2018.1436057

Lindow, S., & Betsch, T. (2021). Preschoolers' competence to use advice in everyday decision contexts. *Journal of Experimental Child Psychology*, *215*, 105311. https://doi.org/10.1016/j.jecp.2021.105311

Löckenhoff, C. E. (2018). Aging and decision-making: A conceptual framework for future research – A mini-review. *Gerontology*, *64*, 140–148. https://doi.org/10.1159/000485247

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*(4), 796–810. https://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. https://doi.org/10.1111/j.1467-8624.2010.01535.x

Pasquini, E. S., Corriveau, K. H., Koenig, M. A., & Harris, P. L., (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology*, *43*, 1216–1226. https://doi.org/10.1037/0012-1649.43.5.1216

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 534–552. https://doi.org/10.1037/0278-7393.14.3.534

Piaget, J., & Inhelder, B. (1951/1975). *The origin of the idea of chance in children*. New York: W.W. Norton & Company.

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*, 207–236. http://doi.org/10.1037/0096-3445.135.2.207

Schlottmann, A., & Wilkening, F. (2012). 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. 55–83). Cambridge: Cambridge University Press.

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

Sobel, D. M., & Kushnir, T. (2013). Knowledge matters: How children evaluate the reliability of testimony as a process of rational inference. *Psychological Review*, *120*, 779–797. https://doi.org/10.1037/a0034191

Söllner, A., Bröder, A., Glöckner, A., & Betsch, T. (2014). Single-process versus multiple-strategy models of decision making: Evidence from an information intrusion paradigm. *Acta Psychologica*, *146*, 84–96. https://doi.org/10.1016/j.actpsy.2013.12.007

Sternberg, G. (2009). Selectivity in infant social referencing. *Infancy*, *14*, 457–473. https://doi.org/10.1080/15250000902994115

Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.

Zmyj, N., Buttelmann, D., Carpenter, M. J. & Daum, M. M. (2010). The reliability of a model influences 14-month-olds' imitation. *Journal of Experimental Child Psychology*, *106*, 208–220. https://doi.org/10.1016/j.jecp.2010.03.002

Cite this article: Lindow, S., Lehmann, A., Buttelmann, D., and Betsch, T. (2023). Preschoolers' use of cue validities as weights in decision-making: Certainty does not substantially change the world. *Judgment and Decision Making*, e38. https://doi.org/10.1017/jdm.2023.36