Received: 9 November 2022 | Revised: 21 April 2023 | Accepted: 13 June 2023

DOI: 10.1002/dev.22407

RESEARCH ARTICLE

Developmental Psychobiology WILEY

When and how do children solve the Weather Prediction Task?

Emilie Bochud-Fragnière[1] | **Pierre Lavenex**[1] | **Pamela Banta Lavenex**[1,2]

[1]Laboratory of Brain and Cognitive Development, Institute of Psychology, University of Lausanne, Lausanne, Switzerland

[2]Faculty of Psychology, UniDistance Suisse, Brig, Switzerland

Correspondence
Pamela Banta Lavenex, Faculty of Psychology, UniDistance Suisse, Schinerstrasse 18, 3900 Brig, Switzerland.
Email: pamela.banta@unidistance.ch

Funding information
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung; Swiss National Science Foundation, Grant/Award Number: 100019_165481; Faculty of Social and Political Sciences of the University of Lausanne

Abstract

The Weather Prediction Task (WPT) can be solved by adults using several strategies dependent on different memory systems. One developmental study reported that 8–12-year-old children can solve WPT-like tasks but, because of inadequate analyses, the cognitive processes involved in solving the task have not been established. The present study aimed to determine at what age children can first solve the WPT and identify the strategies used by children of different ages. We tested 3–12-year-old typically developing children and 20–30-year-old adults on a modified WPT. We performed detailed analyses of performance for each pattern of cue–outcome associations to decipher the strategies used by participants. None of the 3–5.5-year-old children solved the task. About one third of 5.5–7.5-year-old children performed above chance, relying only on the two most predictive cues. In contrast, about 80% of 7.5–12-year-old children performed above chance, relying on a conditional hierarchical strategy. Similar to 20–30-year-old adults, 7.5–12-year-old children considered the highly predictive cues primarily and the less predictive cues secondarily. These findings indicate that the learning strategies used to solve the WPT evolve from middle to late childhood and reflect an increasing ability to use a conditional strategy concomitant with the development of the hippocampus-dependent memory system.

KEYWORDS
conditional learning, development, hippocampus, multiple memory systems, multiple-cue learning, probabilistic classification learning, striatum

1 | INTRODUCTION

1.1 | Multiple memory systems

It is well established that there are multiple parallel but functionally distinct and dissociable memory systems in the brain (Ashby & Maddox, 2005; Cohen et al., 1985; McDonald & White, 1993; McDonald et al., 2004; White et al., 2013). Amongst these multiple memory systems, the hippocampal and the striatal memory systems are often compared as the types of memory representations that they support are functionally ~~the~~ opposite. Although the hippocampal memory system is fundamental for creating flexible conditional or context-dependent representations, the striatal memory system creates fixed stimulus–response representations of behavioral actions (Chang & Gold, 2003a; Cohen et al., 1985; Hartley et al., 2003; Packard et al., 1989; Squire, 2004; Voermans et al., 2004). Fundamental research in rodents and nonhuman primates and neuropsychological research with patients have been key to elucidating the function of and interactions between these two different memory systems by assessing performance on tasks that specifically implicate one or the other of these memory systems. This approach has helped to define what types of memory are fundamentally subserved by these complementary systems, but also how damage to one of the two systems impacts the functioning of the other system, potentially via the release of competitive inhibition (Chang & Gold, 2003a, 2003b; Hartley & Burgess, 2005; Lee et al., 2008; Poldrack & Packard, 2003; Poldrack et al., 2001). Although much of our understanding of the function of these different memory systems and their interaction comes from investigations employing spatial cognitive tasks, studies have also investigated the role of different memory systems in nonspatial cognitive tasks, in particular with probabilistic classification learning tasks.

1.2 | The Weather Prediction Task

The Weather Prediction Task (WPT; Knowlton et al., 1994) is a probabilistic classification learning task widely used to assess striatum-dependent learning and memory processes. The WPT requires participants to predict the weather based on different combinations of four tarot cards. On each trial between one and three cards, resulting in 14 distinct patterns or combinations of cues, are displayed with two weather outcomes and participants must predict whether the outcome is more likely to be sunshine or rain. Participants receive feedback as to the correct outcome. However, just like the weather, cue–outcome associations are probabilistic and thus not perfectly correlated. Although two highly predictive cues are associated with rain or sunshine on 75% of trials and with the opposite outcome on the remaining 25% of trials, two less predictive cues are associated with rain or sunshine on 57% of trials and with the opposite outcome on the remaining 43% of trials. Importantly, participants are not explicitly informed about the probabilistic nature of the task. Due to the probabilistic nature of the WPT, participants were originally considered to gradually acquire implicit knowledge about the cue–outcome associations across successive trials via striatum-dependent learning processes (e.g., Brown et al., 2010; Fernandez et al., 2021; Knowlton et al., 1994; but see Lagnado et al., 2006). Behavioral and neuroimaging studies, however, have provided evidence that the WPT implicates both hippocampus-dependent declarative processes and striatum-dependent non-declarative processes (Foerde et al., 2006, 2007; Gluck et al., 2002; Hopkins et al., 2004; Lagnado et al., 2006; Meeter et al., 2006; Newell et al., 2007; Poldrack et al., 2001; Price, 2009), with hippocampus-dependent processes being particularly prevalent in the early stages of learning. Nevertheless, unduly influenced by the early reports purporting that the WPT specifically assessed striatal-dependent procedural learning, many researchers have persisted in using the WPT and WPT-like tasks for the explicit purposes of characterizing procedural learning and memory in typical and atypical development (e.g., Gabay & Goldfarb, 2017; Kemeny & Lukacs, 2010; Mayor-Dubois et al., 2010, 2016), and the behavioral consequences of known or presumed structural or functional basal ganglia deficiencies in children and adults with, for example, attention-deficit hyperactivity disorder (Gabay & Goldfarb, 2017), bulimia nervosa (Labouliere et al., 2016), schizophrenia (Fernandez et al., 2021; Foerde et al., 2008; Gomar et al., 2011; Horan et al., 2008; Karcher et al., 2019; Keri et al., 2000; Weickert et al., 2002), Parkinson and Huntington diseases (Knowlton, Mangels, et al., 1996; Knowlton, Squire, et al., 1996; Shohamy, Myers, Grossman, et al., 2004; Shohamy, Myers, Onlaor, et al., 2004), and Tourette syndrome (Keri et al., 2002; Marsh et al., 2004).

Consistent with the neuroimaging data showing that the WPT also implicates hippocampus-dependent declarative memory processes (Foerde et al., 2006; Poldrack et al., 2001), Bochud-Fragnière et al. (2022) have recently shown that typically developed 20–30-year-old adults solve an abridged version of the WPT (with only 100 trials) using a flexible cognitive strategy that is likely dependent on the hippocampus. In their study, they introduced new “pattern analyses” as a more appropriate manner of characterizing the learning strategies used to solve the WPT. First, they computed the performance of every single participant on each of the 14 individual patterns of cues. Second, they compared the overall performance of young adults for each individual pattern to the performance that would be expected if participants were using a particular learning strategy. Finally, they compared the different levels of performance between the 14 patterns to characterize how each individual cue influenced the participants’ choices when presented alone or in combination with other cues. They found, in accordance with Gluck et al. (2002) and Meeter et al. (2006), that most young adults do not use a multicue strategy to solve the WPT (i.e., predicting the correct outcome based on their implicit knowledge about the probabilities of cue–outcome associations acquired across successive trials), as initially proposed (Knowlton et al., 1994). Instead, their analyses demonstrated that the responses of young adults matched the use of a hierarchical strategy, a flexible cognitive strategy in which participants primarily base their responses on the presence or absence of the two most predictive cues and consider the two less predictive cues secondarily, except when the two most predictive cues are both present . This strategy, by which the participant determines the correct outcome following a series of hierarchical comparisons of the cues on each trial, is akin to a context-dependent or conditional learning strategy that has been shown to be dependent on the hippocampus (Andersen et al., 2006; Greene et al., 2006; Ridley et al., 1995). These findings, together with previous studies cited above, suggest that the WPT does not necessarily assess implicit procedural learning, and that detailed pattern analyses are required to infer the learning strategy used to solve the task.

1.3 | When and how do children solve the WPT?

To our knowledge, only one study has assessed how typically developing (TD) children solve WPT-like tasks (Mayor-Dubois et al., 2016). Employing the Ice Cream Task (ICT; a modified version of the WPT; Shohamy, Myers, Grossman, et al., 2004), they found that at 8 years of age, children were already able to solve the task. They also showed that although 8-year-old children performed as well as 10- and 12-year-old children during the first 100 trials of the task, 12-year-old children made more correct choices than 8-year-old children in the last 100 trials (101–200). In order to determine whether children were using an implicit procedural learning strategy (i.e., the multicue strategy as originally presumed in Knowlton et al. (1994)), they applied the same types of strategy analyses proposed by Gluck et al. (2002). Their strategy analyses suggested that the number of children who used the multicue strategy increased with age, reaching around 40% in 12-year-old children. However, these strategy analyses are unable to statistically differentiate among several of the most likely used strategies and are thus inadequate to define the learning strategies used by participants to solve the WPT (Bochud-Fragnière et al., 2022). Moreover, the youngest participant tested in the study by Mayor-Dubois et al. (2016) was 7.9 years old. Thus, it is currently unknown at what age children are first able to solve tasks, such as the WPT or ICT, whether children use a similar strategy across development, and if not when they begin using a hierarchical strategy like adults.

Understanding when and how children solve the WPT is particularly important from a theoretical viewpoint and enables comparisons between the function and interaction of these two different memory systems in spatial and nonspatial contexts across development. In the spatial domain, for example, striatum-dependent spatial memory processes have been shown to emerge early (Acredolo, 1978; Acredolo & Evans, 1980; Cornell & Heth, 1979), whereas hippocampus-dependent spatial memory processes exhibit a later emergence and a protracted maturation (Newcombe et al., 1998; Ribordy Lambert et al., 2013, 2015, 2017). If findings in the spatial domain are extended to the nonspatial domain, it is reasonable to hypothesize that if the youngest children solve the WPT that they do so using a striatum-dependent procedural learning strategy, whereas older children may use a hippocampus-dependent declarative learning strategy. In contrast, if younger children are incapable of solving the WPT, this would suggest that like adults, young children do not use striatum-dependent strategies to learn how to solve the WPT.

1.4 | Aim of the study

The current study was designed to determine at what age children first demonstrate the ability to solve the WPT and to identify the strategies used by children of different ages. The identification of different strategies at different stages of development would suggest that different memory systems are implicated in solving the WPT at these ages. We tested children from 3 to 12 years of age on an adapted, abridged version of the WPT. To characterize the learning strategies supporting WPT performance in children, we performed the pattern analyses designed by Bochud-Fragnière et al. (2022) and compared the results obtained for children with those previously published for young adults (Bochud-Fragnière et al., 2022).

2 | MATERIALS AND METHODS

2.1 | Participants

Participants were 71 French-speaking TD children (31 girls; average age: 6.50 years; range: 3.03–11.85) and 22 French-speaking young adults (11 females; average age: 26.29 years; range: 21.72–30.82). Of the 71 children, one 3.59-year-old boy and one 3.70-year-old boy were not included in the analyses because they did not finish the task.

Both children and adult participants were recruited in Switzerland via personal connections. Consistent with the demographics of the local population, participants were generally of Northern European decent and of middle socioeconomic status. Exclusion criteria were atypical development (e.g., parent- or participant-reported developmental delay or high potential), neurological disorders (e.g., epilepsy), or chronic substance use (drugs or medications). Testing took place Monday through Saturday, between 8:00 a.m. and 6:30 p.m. Participants were tested in their homes, or for some adult participants in the home of the experimenter. The only requirement for testing was a quiet room with a table and two chairs. Adult participants and the parents of children gave informed written consent prior to beginning the study and all children gave oral consent and, when possible, their written consent. Child participants received a stuffed animal and a symbolic Junior Researcher diploma for their participation. Adult participants received a gift card to a local retailer. Human subjects research was approved by the Cantonal Ethics Commission for Human Research [Vaud, Switzerland; project no PB_2017-00074 (60/14)] and was in accordance with the code of ethics of the World Medical Association (Declaration of Helsinki) for experiments involving human subjects in research.

2.2 | General procedure

The task used here with children is identical to that used to assess WPT performance in adults in Bochud-Fragnière et al. (2022): There were no differences in any of the parameters of the task or in the administration of the task between children and adults. The task was modified from the original WPT (Knowlton et al., 1994) and the ICT used to test young children (Shohamy, Myers, Grossman, et al., 2004). First, cartoonish sea animals were used as cues instead of tarot cards, and two different smiley faces served as outcomes (Figure 1). Second, to give young participants a better chance to complete the task, we abridged the task and presented only 100 trials and the same associative outcome probabilities as Gluck et al. (2002) and Mayor-Dubois et al. (2016) (i.e., .20, .40, .60, .80). The number of presentations of each pattern was adapted to respect these probabilities (Table 1). The task was programmed with E-Prime 3.0 software (Psychology Software Tools, Sharpsburg, PA, USA) and was presented on a computer tablet with a tactile screen (HP Spectre x360 Convertible 13-w0XX running with Microsoft Windows 10 64-bit; Processor Intel Core i7; 8 GB RAM; 13.3-in. (1920 × 1080); landscape orientation). The participant's choice was recorded automatically by E-prime when they touched one of the two outcome stimuli.

2.3 | Stimuli and detailed procedure

We used the four unique sets of four cues (cartoonish pictures of sea animals) and two outcomes (a smiley face with sunglasses enjoying the sun, hereafter referred to as "sun"; a smiley face wearing a knit

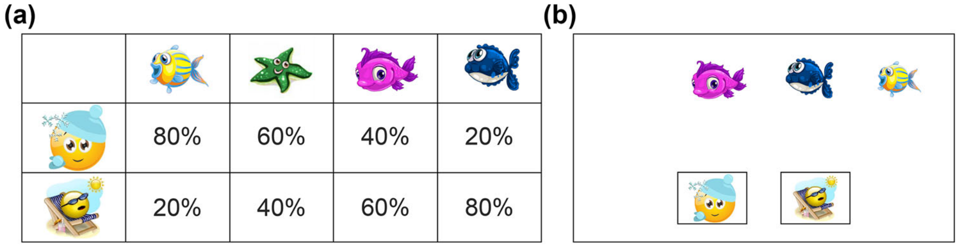

FIGURE 1 (A) Example of one set of four cues (sea animals), which were associated with each outcome (sun or snow) with a fixed level of probability across all patterns. (B) Representation of the computer screen for one trial when three animals were displayed, and the participant had to choose between snow and sun. Note that one, two, or three animals could be displayed on any given trial; all four animals were never displayed simultaneously.

TABLE 1 Number of occurrences of each pattern and their association with each outcome.

	Cues				Outcome A	Outcome B	Total number of
Pattern	1	2	3	4	Snow	Sun	Occurrences
1	0	0	0	1	8	1	9
2	0	0	1	0	4	1	5
3	0	0	1	1	12	1	13
4	0	1	0	0	1	3	4
5	0	1	0	1	5	1	6
6[a]	0	1	1	0	1	2	3
7	0	1	1	1	9	1	10
8	1	0	0	0	1	8	9
9[a]	1	0	0	1	2	2	4
10	1	0	1	0	1	6	7
11	1	0	1	1	2	1	3
12	1	1	0	0	1	12	13
13	1	1	0	1	2	3	5
14	1	1	1	0	1	8	9
Total	7	7	7	7	50	50	100

Note: The probability that each individual cue was associated with each outcome was calculated across all patterns. See main text for details.
[a]Patterns 6 and 9 predict both outcomes equally, because they combine either two cues that are highly predictive for both sun and snow (pattern 9) or two cues that are less predictive for both sun and for snow (pattern 6). They were not included in the analyses, which thus comprised 93 trials instead of 100.

hat enjoying the snow, hereafter referred to as "snow"; Figure 1; see also Figure S1) created by Bochud-Fragnière et al. (2022). One set was pseudo-randomly attributed to each participant. The WPT was divided into two types of trials: learning trials and cue knowledge test trials. For both types, each trial was preceded by a fixation cross displayed for 1 s at the center of the screen. Subsequently, the stimuli appeared. During the learning trials, participants were required to choose which of two outcomes (sun or snow) was predicted by the combination of one to three cues (chosen from amongst the four sea animals that combined to produce 14 unique patterns; Table 1). The instructions were the same for all participants: "You will see some fish, and you will have to discover if they prefer to play in the sun or in the snow." Then, the experimenter or the participant touched the screen to begin the task. The order of the trials and the position of the cues (each cue could appear at one of four possible positions on the screen: left, center left, center right, right) were randomly determined with E-prime. The positions of the sun and snow outcome images were fixed for all participants (Figure 1).

Traditionally, successful learning of the WPT task is assessed by determining whether participants chose the correct outcomes at above chance levels on the training trials (Gluck et al., 2002; Knowlton et al., 1994; Meeter et al., 2006). Accordingly, in our task participants had the opportunity to learn, using trial-by-trial feedback, ~~in~~ which cues were associated probabilistically with which outcome. Once participants had indicated their choice by touching the screen, a congruent or incongruent outcome image appeared, depending on the predetermined probability associated with the cues. (*NB*: The cues remained visible during this time.) The experimenter also provided verbal feedback that was consistent with the visual feedback (when the visual feedback

~~fits~~ the answer of the participant: "yes," "right," when the visual feedback did not fit her or his answer: "no," "not this one," "it's no big deal," "Oops, no"; note that the last two feedbacks were used to lessen the impact of a series of negative feedbacks and to maintain participants' motivation) until participants demonstrated the ability to understand the visual feedback (as evidenced when participants nodded their head or gave themselves verbal feedback). The participants then pressed a red button at the bottom of the screen to advance to the next trial. It is important to understand that due to the probabilistic nature of the task, a participant can make a correct prediction but receive incongruent feedback. For example, even though fish X was correlated 80% of the time with sun, 20% of the time when fish X was shown the snow outcome was displayed on the screen. Therefore, incongruent trials could not be differentiated from congruent trials based on feedback, but rather only by cumulatively considering all the feedback given for each cue or pattern of cues across trials. For scoring, the choice was judged as correct if it corresponded to the most probable outcome for that pattern of cue(s), regardless of the displayed outcome. For example, for pattern 0001, the appropriate prediction judged as correct was always outcome A (snow), even if in one of the trials, the given feedback was outcome B (sun) (Table 1).

After 100 training trials, four cue knowledge test trials without feedback were presented during which each cue was displayed individually (patterns 0001, 0010, 0100, 1000) in the presence of the two possible outcomes, and participants were asked to choose which outcome was associated with each individual cue. Once participants indicated their choice by touching the outcome image on the screen, the next trial began with the fixation cross. The instructions for the cue knowledge test trials were the same for all participants: "You will see a fish on the computer screen and you will have to recall if it preferred to play in the snow or in the sun." Then, the experimenter or the participant touched the screen to begin the cue knowledge test trials. It is important to note that because each cue is assessed only once (for a total of four cue knowledge test trials), and because the probability of being right is 50% (the correct answer is either sun or snow), performance on the cue knowledge test trials cannot be used to statistically affirm learning independently. Moreover, these four cue knowledge test trials are not informative with respect to determining the strategies participants used to solve the WPT but rather inform about a participant's ability to explicitly access knowledge concerning the specific cue–outcome associations. For example, previous studies have shown that although participants perform above chance on the training trials when forced to solve the task using a striatum-based strategy, they exhibit significantly less explicit knowledge of the cue–outcome associations (Foerde et al., 2006, 2007).

For each trial (during learning and cue knowledge test phases), the computer recorded the choice of the participant and the reaction time. Although some previous studies limited response time to 5 s (Knowlton et al., 1994; Knowlton, Mangels, et al., 1996; Knowlton, Squire, et al., 1996), ~~as~~ because reaction time depends on self-confidence and because limiting the time available to respond could induce impulsivity, we imposed no time restriction for our participants to respond. Nevertheless, our participants' average response time (children: mean (M) = 2.59 s; SD = 1.15; adults: M = 2.62 s; SD = .55) was similar to that observed in previous studies, and the large majority of responses were made in under 5 s. Throughout the whole experiment, the experimenter sat next to the participant and wore sunglasses to preclude cueing participants with respect to the correct choice with their eye gaze.

2.4 | Data analysis

2.4.1 | Global performance

We used two-tailed paired samples t-tests to determine whether participants exhibited a bias toward choosing one of the two different outcomes, sun or snow, irrespective of the cues presented. We considered only 93 trials when calculating an overall performance score, as patterns 0110 ($n = 3$) and 1001 ($n = 4$) predicted theoretically both outcomes equally and were thus excluded from the analyses. We normalized the number of times outcome A (snow) and the number of times outcome B (sun) were chosen by dividing them with the number of correct "snow" and "sun" responses (i.e., 46 and 47, respectively; the difference in the number of trials is due to the overall distribution of cue and outcome combinations across all 100 trials). During training trials, as a group children did not choose "sun" ($M = 1.02$, $SD = .15$) more than "snow" ($M = .98$, $SD = .16$; $t_{(68)} = .894$, $p = .374$, $d_z = .108$). During the cue knowledge test trials, there was a tendency for children to choose "sun" ($M = 2.14$, $SD = .65$) more than "snow" ($M = 1.84$, $SD = .66$; $t_{(68)} = 1.947$, $p = .056$, $d_z = .234$). As the comparison for cue knowledge test trials nearly reached significance, we performed the same analyses for children who passed ($n = 24$) or failed ($n = 45$) the task to determine whether one or both groups of children demonstrated a bias toward a specific outcome on either training or cue knowledge test trials. As a group, children who performed the task above chance level did not choose "sun" ($M = .96$, $SD = .12$) more than "snow" during training trials ($M = 1.04$, $SD = .12$; $t_{(23)} = 1.660$, $p = .110$, $d_z = .339$) or cue knowledge test trials ("sun" $M = 2.04$, $SD = .75$; "snow" $M = 1.96$, $SD = .75$; $t_{(23)} = .272$, $p = .788$, $d_z = .056$). In contrast, there was a tendency for children who did not perform above chance level to choose "sun" ($M = 1.05$, $SD = .16$) more than "snow" during training trials ($M = .95$, $SD = .17$; $t_{(44)} = 1.903$, $p = .064$, $d_z = .284$), and during cue knowledge test trials these children chose "sun" ($M = 2.20$, $SD = .59$) more than "snow" ($M = 1.78$, $SD = .60$; $t_{(44)} = 2.406$, $p = .020$, $d_z = .359$). In sum, only participants who failed the task exhibited a slight bias toward choosing "sun." In contrast, participants who performed above chance level did not exhibit any bias toward choosing one of the two different outcomes irrespective of the cues. Because we were interested in characterizing the strategy used by children who solved the WPT, data from participants who failed are not presented except when mentioned.

As there was no consistent effect of gender (data not shown), data from boys and girls were combined for analysis and presentation. To determine whether the entire group of participants performed above chance level during training and cue knowledge test trials, we performed a one-sample t-test comparing the mean performance of the entire group of participants to the theoretical mean. Individual above chance performance during training was defined statistically using a chi-square test as 56 correct choices out of 93 ($\chi^2_{(1)} = 3.882$, $p = .049$). Because children's individual performance varied with age (see Section 3), we separated the children into three age groups (3–5.5 years ($n = 25$), 5.5–7.5 years ($n = 25$), and 7.5–12 years ($n = 19$)) to characterize how the performance and the learning strategies varied with age. We used chi-square tests, performed on the raw number of participants, to determine whether the number of participants who performed above chance level differed between age groups (i.e., 3–5.5 years, 5.5–7.5 years, 7.5–12 years, and adults).

2.4.2 | Pattern analyses

As discussed in Section 1, the analysis methods developed by Gluck et al. (2002) and Meeter et al. (2006) to determine which learning strategy may be used to perform the WPT are inadequate and potentially biased (Bochud-Fragnière et al., 2022). Nevertheless, we included such strategy analyses in the Supporting Information section as they replicated prior findings showing that they do not allow statistical differentiation of the goodness of fit of the different strategies (Bochud-Fragnière et al., 2022). We report the results of pattern analyses in the main text as they are both necessary and sufficient to determine the learning strategy used to solve the WPT.

To perform pattern analyses, we first constructed ideal response profiles for each strategy (Table 2). These profiles were calculated by assuming that participants adhere strictly to this strategy throughout the entire task or for each block of 50 trials, respectively. To compare the performance for each pattern, we first normalized the number of correct choices for each pattern by dividing it by the number of times it was presented. We then performed two-tailed one-sample *t*-tests to determine whether participants performed better than chance for each pattern across the entire task. We then took into consideration the congruence and the level of predictability of the cues and grouped together the following patterns: 0001 and 1000 (one highly predictive cue); 0010 and 0100 (one less predictive cue); 1100 and 0011 (two congruent cues); 0101 and 1010 (one highly predictive cue and one incongruent less predictive cue); 0111 and 1110 (two congruent cues and one incongruent less predictive cue); and 1011 and 1101 (two congruent cues and one incongruent highly predictive cue). The normalized number of correct choices for each category of patterns was compared with a repeated-measures general linear model analysis. Post hoc analyses were performed with paired-sample *t*-tests when F ratios were significant, thus controlling for type I error rate (Carmer & Swanson, 1973). Although we normalized the performance for each pattern, participants had an equal number of chances to learn the influence of each individual cue (Table 1). These comparisons thus reflect how participants processed each cue, whether presented alone or in combination, and not how they processed each of the 14 patterns.

For the cue knowledge test trials, we used one-sample *t*-tests to determine whether participants could identify which outcome was associated with each cue. We also performed a paired samples *t*-test to compare the participants' performance when presented with highly predictive cues versus less predictive cues.

We did not use corrections for multiple comparisons for individual *t*-tests because in our study, the risk of reporting a difference that may not exist (type I error) is not worse than the risk of missing a difference that may exist (type II error). Accordingly, we followed the recommendations of Rothman (1990), who argued that "not making adjustments for multiple comparisons is preferable because it will lead to fewer errors of interpretation when the data under evaluation are not random numbers but actual observations on nature," and Saville (1990), who also argued that a procedure without correction is preferable because it provides greater consistency to compare results between studies. Although some comparisons might be more critical than others, considering all comparisons is essential to characterize the behavior of participants. The most important indicator of the way participants considered each cue is the direction of the numerical difference for each individual comparison.

All statistical analyses were performed with IBM SPSS Statistics (version 27.0). Significance level was set at $p < .05$ for all analyses. We reported effect size with partial eta squared (η^2_p) for ANOVAs, Cohen's d_s for independent samples *t*-tests, and Cohen's d_z for paired or one-sample *t*-tests. Data can be requested from the corresponding author.

3 | RESULTS

3.1 | Individual and group performance

As a group, children (including children that did not perform above chance level) made more correct choices than if they had chosen randomly across training trials (Figure 2A; $M = 54.06$, $SD = 12.49$; $t_{(68)} = 5.028$, $p < .001$, $d_z = .605$) and the four cue knowledge test trials (Figure 2B; $M = 2.49$, $SD = 1.22$, $t_{(68)} = 3.354$, $p = .001$, $d_z = .404$). However, as compared to all adults (training: $M = 68.95$, $SD = 9.40$; test: $M = 3.19$, $SD = .75$), children made fewer correct choices across training trials ($t_{(89)} = 5.142$, $p < .001$, $d_s = 1.259$) and test trials ($t_{(88)} = 2.476$, $p = .015$, $d_s = .617$; *NB*: One adult did not perform the test trials).

We then evaluated each child's ability to perform the WPT above chance level (≥56 correct choices: all $\chi^2_{(1)} \geq 3.882$, $p \leq .049$). Twenty-four children between 3 and 12 years of age (34.8%) performed the WPT above chance, and thus, reciprocally, 45 children (65.2%) failed the task (Figure 2A). The overall performance of the 24 children who performed the training phase above chance level was 74% of the optimal score across training trials ($M = 69.08$ correct choices out of 93, $SD = 7.49$; above chance; $t_{(23)} = 14.774$, $p < .001$, $d_z = 3.016$); their performance was also above chance level for the four cue knowledge test trials ($M = 3.13$, $SD = .80$, $t_{(23)} = 6.912$, $p < .001$, $d_z = 1.411$). Moreover, the performance of children who performed the task above chance level did not differ from the performance of adults who performed the task above chance level across training trials (adult $M = 70.65$, $SD = 8.00$; $t_{(42)} = .670$, $p = .507$, $d_s = .203$) and cue knowledge test trials (adult $M = 3.26$, $SD = .73$; $t_{(41)} = .584$, $p = .562$, $d_s = .179$).

TABLE 2 Strategies, percentage of optimal responses for the entire task and each pattern.

Strategies	Choices based on	%	Patterns													
			0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110
		Entire task	*n* = 9	*n* = 5	*n* = 13	*n* = 4	*n* = 6	*n* = 3	*n* = 10	*n* = 9	*n* = 4	*n* = 7	*n* = 3	*n* = 13	*n* = 5	*n* = 9
Multicue	All combinations of cues	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Hierarchical	Primarily on highly predictive cues Secondarily on less predictive cues Except in presence of both highly predictive cues	95.7	100	100	100	100	100	100	100	100	100	100	50	100	50	100
Equal weight	Combination of cues without hierarchy (i.e., without probabilistic value)	93.0	100	100	100	100	50	100	100	100	100	50	100	100	100	100
One-cue-1000	Highly predictive cue associated with sun	92.5	100	100	100	0	100	100	100	100	100	100	0	100	100	100
Two most predictive cues	Presence or absence of the two most predictive cues	90.9	100	50	100	50	100	100	100	100	100	100	50	100	50	100
One-cue-0001	Highly predictive cue associated with snow	89.2	100	0	100	100	100	100	100	100	100	100	100	100	0	100
Congruent cues	Patterns containing only one cue and their congruent combinations	78.5	100	100	100	100	50	100	50	100	100	50	50	100	50	50
One-cue-0010	Less predictive cue associated with snow	66.7	0	100	100	100	0	100	100	100	100	0	100	100	100	0
One-cue-0100	Less predictive cue associated with sun	65.6	100	100	100	100	0	100	0	0	100	0	100	100	100	100
Singleton	Patterns containing only one cue	64.5	100	100	50	100	50	100	50	100	100	50	50	50	50	50
Singleton strong	Patterns containing only one of the most predictive cues	59.7	100	50	50	50	50	100	50	100	100	50	50	50	50	50
Undetermined	Random or undetermined strategy	50.0	50	50	50	50	50	100	50	50	100	50	50	50	50	50

Note. The percentage of correct responses for each pattern is based on the answers that should be given by a participant strictly following the strategy. For example, a participant following the one-cue-1000 strategy will answer outcome B (sun) in presence of the cue number 1000 (i.e., 1000, 1001, 1010, 1011, 1100, 1101, and 1110), but outcome A (snow) in its absence (i.e., 0001, 0010, 0011, 0100, 0101, 0110, and 0111). It is essentially impossible for a participant to adhere strictly to a strategy across the entire task as participants obviously need at least a few trials to learn any rule and be able to reliably implement any given strategy, and they may also switch from one strategy to another during the task. Nevertheless, when the model predicted 100% correct responses for a particular pattern, participants should exhibit at least above chance performance. In contrast, when the model predicted 0% correct responses for a particular pattern, participants should exhibit a performance that is below chance level.

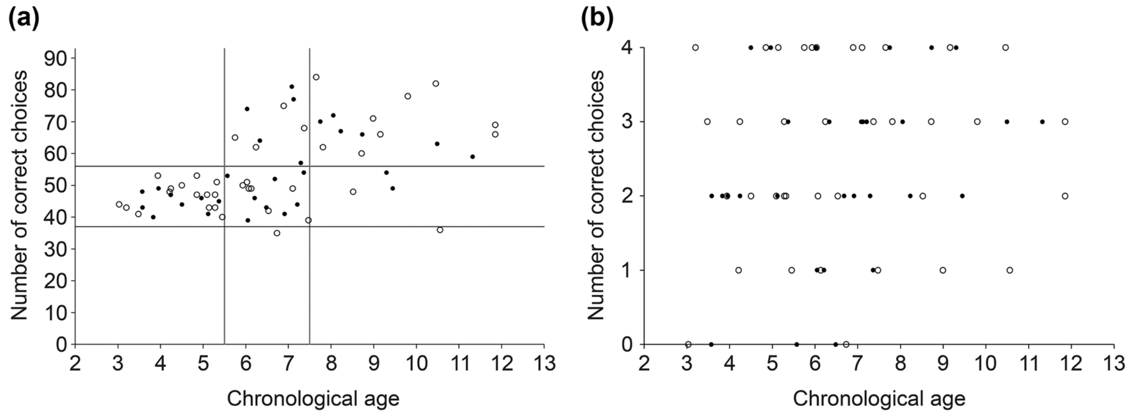

FIGURE 2 Individual performance of children between 3 and 12 years of age in the Weather Prediction Task (WPT) (closed circles: girls; open circles: boys). (A) Number of correct choices during training trials with various combinations of cues and feedback. The upper black line represents the number of correct choices (56/93) defined as performance above chance at the individual level, whereas the lower line represents the number of correct choices (37/93) defined as performance below chance at the individual level (both $\chi^2_{(1)} = 3.882$, $p = .049$). The two vertical lines separate the three different age groups (3–5.5, 5.5–7.5, and 7.5–12 years) (B) number of correct choices during the four cue knowledge test trials with individual cues and no feedback.

Of the 25 children between 3 and 5.5 years of age who completed the 100 training trials, none (0%) performed above chance level (Figure 2A; see also Table S1). The youngest child who performed the task above chance level was a 5.75-year-old boy. Of the 25 children between 5.5 and 7.5 years of age, nine (36%) performed above chance level. Finally, of the 19 children between 7.5 and 12 years of age, 15 (79%) performed above chance level. Accordingly, the proportion of participants who performed above chance level differed between age groups (3–5.5 years (0/25), 5.5–7.5 years (9/25), 7.5–12 years (15/19) and 20–30 years (20/22): $\chi^2_{(3)} = 48.009$, $p < .001$). More specifically, there were differences in the number of participants who performed above chance level between the following groups: 3–5.5 and 5.5–7.5 years, 3–5.5 and 7.5–12 years, 3–5.5 and 20–30 years, 5.5–7.5 and 7.5–12 years, 5.5–7.5 and 20–30 years (all $\chi^2_{(1)} > 7.808$, $p < .006$), but not between 7.5–12 and 20–30 years ($\chi^2_{(1)} = 1.167$, $p = .279$). Note that one 6.73-year-old boy and one 10.56-year-old boy who failed the task did not reach the level of performance expected by chance (Figure 2A; see also Table S1). Instead, detailed choice analyses revealed that both participants alternated their responses between the two outcomes: sun and snow. The number of choices of the outcome that was opposite to the immediately preceding outcome was higher than chance ($\chi^2_{(1)} = 7.860$, $p = .005$ and $\chi^2_{(1)} = 17.483$, $p < .001$, respectively).

In sum, although the youngest child who performed above chance level was 5.75 years old, not all participants older than 5.75 years solved the task. Indeed, only 36% of children between 5.5 and 7.5 years of age performed above chance, and 79% of children between 7.5 and 12 years of age performed above chance, as compared to 91% of 20–30-year-old adults. Nevertheless, the performance of children who performed above chance level was similar to that of young adults, and the proportion of the oldest children who performed above chance level was not different from that of adults.

3.2 | Individual patterns

Children ($n = 45$) and adults ($n = 2$) who did not perform above chance were excluded from all subsequent analyses designed to characterize the strategies used to solve the WPT, except when mentioned specifically. We therefore analyzed the performance of two groups of successful children (i.e., one group between 5.5 and 7.5 years of age and the other group between 7.5 and 12 years of age) and one group of young adults for each pattern of cues. (*NB*: We did not analyze patterns 0110 and 1001 as they both theoretically predicted each outcome equally; Table S2 for the analysis of the performance of children who did not perform above chance level.) As previously shown (Bochud-Fragnière et al., 2022), this analysis results in distinct performance profiles that can be compared to the theoretical profile illustrating the expected percentage of correct responses for each pattern of cues when using a given strategy (Table 2). Thus, if an observed profile corresponds to a theoretical profile, this indicates that this specific strategy was used by this group of participants.

3.2.1 | 5.5–7.5-year-old children

Considering all training trials, the overall group performance of children between 5.5 and 7.5 years of age was above chance for all the patterns, except for the two patterns that included one less predictive cue (patterns 0010 and 0100) and the two patterns that included the two highly predictive cues and one of the two less predictive cues (patterns 1011 and 1101; Table 3). The results of the 5.5–7.5-year-old children were thus consistent with the use of the two most predictive cues strategy, which enables above chance performance for all patterns (i.e., 0001, 0011, 0101, 0111, 1000, 1010, 1100, and 1110), except when the less predictive cues are necessary to determine the correct

TABLE 3 Comparison of the performance of each pattern with chance level.

		5.5–7.5-year-old ($n = 9$)					7.5–12-year-old ($n = 15$)					20–30-year-old ($n = 20$)				
Patterns	n	M	SD	$t_{(8)}$	p	Cohen's d_z	M	SD	$t_{(14)}$	p	Cohen's d_z	M	SD	$t_{(19)}$	p	Cohen's d_z
0001	9	7.67	1.73	5.485	<.001**	1.828	7.13	1.81	5.643	<.001**	1.457	7.20	1.67	7.216	<.001**	1.614
0010	5	2.89	1.54	.759	.469	.253	3.47	1.41	2.660	.019**	.687	3.30	1.66	2.158	.044**	.483
0011	13	10.78	1.20	10.678	<.001**	3.559	10.60	1.50	10.569	<.001**	2.729	11.20	1.40	15.022	<.001**	3.359
0100	4	2.44	1.13	1.180	.272	.393	2.73	1.53	1.852	.085*	.478	2.90	1.12	3.596	.002**	.804
0101	6	4.78	1.09	4.880	.001**	1.627	4.33	1.23	4.183	<.001**	1.080	3.95	1.54	2.762	.012**	.618
0110[a]	3	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
0111	10	7.33	1.58	4.427	.002**	1.476	7.47	1.64	5.819	<.001**	1.502	6.70	2.11	3.611	.002**	.808
1000	9	7.00	1.87	4.009	.004**	1.336	7.53	1.41	8.347	<.001**	2.155	7.70	1.38	10.368	<.001**	2.318
1001[a]	4	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1010	7	5.56	1.01	6.083	<.001**	2.028	4.53	1.81	2.214	.044**	.572	4.70	1.59	3.369	.003**	.753
1011	3	1.78	.83	1.000	.347	.333	1.80	1.08	1.074	.301	.277	1.55	.89	.252	.804	.056
1100	13	10.22	1.92	5.810	<.001**	1.937	10.40	1.55	9.750	<.001**	2.517	11.60	1.31	17.359	<.001**	3.882
1101	5	2.33	1.73	.289	.780	.096	2.80	1.21	.963	.352	.249	2.80	1.24	1.082	.293	.242
1110	9	6.44	2.01	2.907	.020**	.969	6.20	1.70	3.876	.002**	1.001	7.05	1.50	7.585	<.001**	1.696

Note: Two-tailed Student *t*-tests comparing the performance for each individual pattern with chance level for each age group of participants who performed the task above chance level based on the total number of correct choices (Figure 2). n = number of presentations of each pattern.
Abbreviation: M, mean.
[a]The patterns including either the two highly predictive cues (1001) or the two less predictive cues (0110) predicted both outcomes equally but are nonetheless listed for the sake of completeness.
**$p < .05$ for two-tailed, one-sample Student's *t*-test comparisons.
* $p < .10$ for two-tailed, one-sample Student's *t*-test comparisons.

answer (i.e., 0010, 0100, 1011, and 1101; Table 2; see also Table S3 for a correspondence between pattern performance and strategies).

3.2.2 | 7.5–12-year-old children

Considering all training trials, the overall group performance of children between 7.5 and 12 years of age was above chance for all the patterns, except for the two patterns that included the two highly predictive cues and one of the two less predictive cues (patterns 1011 and 1101) and pattern 0100 for which the *p* value just failed to reach the predefined level of statistical significance ($p = .085$ for two-tailed, paired Student's *t*-test comparison; Table 3). These results are more consistent with the use of the hierarchical strategy than any other strategy. Indeed, the hierarchical strategy should enable above chance performance for all patterns (i.e., 0001, 0010, 0011, 0100, 0101, 0111, 1000, 1010, 1100, and 1110), except when the two most predictive cues are present simultaneously (i.e., 1011 and 1101; Table 2; see also Table S4 for a correspondence between pattern performance and strategies).

3.2.3 | Adults

As discussed in Bochud-Fragnière et al. (2022), considering all training trials (Table 3), the overall group performance was above chance for all the patterns except for the two patterns that included the two highly predictive cues and one of the two less predictive cues (patterns 1011 and 1101). These results are consistent with the use of the hierarchical strategy, which enables above chance performance for all patterns (i.e., 0001, 0010, 0011, 0100, 0101, 0111, 1000, 1010, 1100, and 1110), except when the two most predictive cues are present simultaneously (i.e., 1011 and 1101; Table 2; see also Table S5 for a correspondence between pattern performance and strategies).

In sum, the fact that neither of the two groups of children nor the group of adults scored above chance on all the patterns indicates that they did not use the multicue strategy to solve the task as originally hypothesized (Knowlton et al., 1994). Moreover, the fact that 5.5–7.5-year-old children's choices were consistent with the two most predictive cues strategy and that 7.5–12-year-old children's choices were consistent with the hierarchical strategy (similar to adults) suggests that the ability to rely on a more effective strategy evolves gradually with age. The performance of 7.5–12-year-old children was similar to that of 20–30-year-old adults for each pattern of cues.

3.3 | Categories of patterns

As previously demonstrated (Bochud-Fragnière et al., 2022), individual pattern analyses can reveal how well participants perform for each pattern of cues, but not how the different categories of cues (highly vs. less predictive) may have influenced their choices. We therefore performed an analysis of the different groups of patterns based on their congruence and level of predictability for each age group. Note that we did not include patterns 0110 and 1001 in the following analyses as they theoretically predicted both outcomes equally.

3.3.1 | 5.5–7.5-year-old children

Considering the entire block of training trials, we found differences in the normalized number of correct choices for the different categories of patterns (Figure 3A; $F_{(5,40)} = 6.474, p < .001, \eta^2_p = .447$). Group performance was above chance for all categories of patterns except for the two patterns that consisted of only one less predictive cue (patterns 0010 and 0100) and the two patterns that included the two highly predictive cues and one of the two less predictive cues (patterns 1011 and 1101): 0001 and 1000 (one highly predictive cue; $M = .81, SD = .18, t_{(8)} = 5.246, p = .001, d_z = 1.749$), 0011 and 1100 (two congruent cues; $M = .81, SD = .10, t_{(8)} = 9.071, p < .001, d_z = 3.024$), 0101 and 1010 (one highly predictive cue and one incongruent less predictive cue; $M = .79, SD = .13, t_{(8)} = 6.565, p < .001, d_z = 2.188$), 0111 and 1110 (two congruent cues and one incongruent less predictive cue; $M = .72, SD = .14, t_{(8)} = 4.872, p = .001, d_z = 1.624$), 0010 and 0100 (one less predictive cue; $M = .59, SD = .17, t_{(8)} = 1.654, p = .137, d_z = .551$), 1011 and 1101 (two congruent cues and one incongruent highly predictive cue; $M = .53, SD = .18, t_{(8)} = .503, p = .629, d_z = .168$).

Importantly, post hoc analyses revealed that 5.5–7.5-year-old children who performed the task above chance had a higher normalized number of correct choices for the categories that included patterns 0001 and 1000, 0011 and 1100, and 0101 and 1010 than the categories that included patterns 0010 and 0100, and 1011 and 1101 (all $p < .05$; Table 4). They also had a higher normalized number of correct choices for the category that included patterns 0011 and 1100 than the category that included patterns 0111 and 1110 ($p < .05$). These children also exhibited a higher normalized number of correct choices for 0001 and 1000 than 0111 and 1110 and for 0111 and 1110 than 1011 and 1101 (although these comparisons just failed to reach the predefined level of statistical significance; $p = .074$ and $p = .056$, for two-tailed, paired Student's *t*-test comparison, respectively). No other comparisons were statistically significant (Table 4).

In accordance with the analyses of training trials, overall group performance on the four cue knowledge test trials was above chance for the category including the highly predictive cues: 0001 and 1000 ($M = 1.89, SD = .33, t_{(8)} = 8.000, p < .001, d_z = 2.667$), but it failed to reach the predefined level of statistical significance for the category including the less predictive cues: 0010 and 0100 ($M = 1.33, SD = .50, t_{(8)} = 2.000, p = .081, d_z = .667$). Moreover, 5.5–7.5-year-old children had a higher performance for highly predictive cues than for less predictive cues (Figure 3B; $t_{(8)} = 3.162, p = .013, d_z = 1.054$).

In sum, consistent with individual pattern analyses, these results indicate that 5.5–7.5-year-old children knew the outcomes associated with the two highly predictive cues (0001 and 1000), but not the outcomes associated with the two less predictive cues (0010 and 0100), and thus can be described as using the two most predictive cues strategy.

TABLE 4 Post hoc comparisons among the different categories of patterns.

	Values	0001 and 1000	0011 and 1100	0101 and 1010	0111 and 1110	0010 and 0100	1011 and 1101
5.5–7.5-year-old children ($n = 9$)							
0001 and 1000	p	–					
	d_z	–					
0011 and 1100	p	.898	–				
	d_z	.044	–				
0101 and 1010	p	.727	.778	–			
	d_z	.121	.097	–			
0111 and 1110	p	.074*	.048**	.185	–		
	d_z	.686	.778	.484	–		
0010 and 0100	p	.038**	.016**	.043**	.153	–	
	d_z	.829	1.011	.801	.527	–	
1011 and 1101	p	.003**	.008**	.009**	.056*	.328	–
	d_z	1.369	1.175	1.146	.743	.348	–
7.5–12-year-old children ($n = 15$)							
0001 and 1000	p	–					
	d_z	–					
0011 and 1100	p	.863	–				
	d_z	.045	–				
0101 and 1010	p	.013**	.022**	–			
	d_z	.738	.668	–			
0111 and 1110	p	.008**	.042**	.522	–		
	d_z	.796	.579	.170	–		
0010 and 0100	p	.039**	.025**	.964	.581	–	
	d_z	.587	.649	.012	.146	–	
1011 and 1101	p	.003**	<.001**	.181	.027**	.018**	–
	d_z	.927	1.347	.364	.640	.688	–
20–30-year-old adults ($n = 20$)							
0001 and 1000	p	–					
	d_z	–					
0011 and 1100	p	.106	–				
	d_z	.380	–				
0101 and 1010	p	<.001**	<.001**	–			
	d_z	.905	1.036	–			
0111 and 1110	p	<.001**	<.001**	.222	–		
	d_z	.921	1.451	.282	–		
0010 and 0100	p	.026**	<.001**	.685	.498	–	
	d_z	.539	.993	.092	.154	–	
1011 and 1101	p	<.001**	<.001**	.059*	<.001**	.004**	–
	d_z	1.496	2.263	.449	1.194	.736	–

Note: p Values and Cohen's d_z for the post hoc paired t-test comparisons among the different categories of patterns for the three different age groups.
**$p < .05$ for two-tailed, paired Student's t-test comparisons.
*$p < .10$ for two-tailed, paired Student's t-test comparisons, which would be *equivalent* to p values < .05 for one-tailed comparisons.

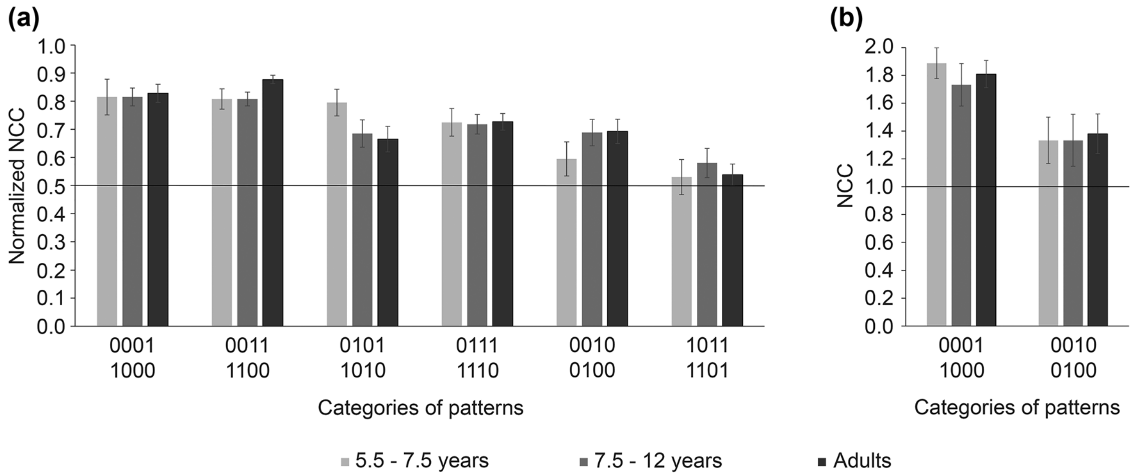

FIGURE 3 (A) Normalized number of correct choices (NCC; mean ± SE) during the training phase for each category of patterns for the participants who performed the Weather Prediction Task (WPT) above chance level; (B) NCC (mean ± SE) during the cue knowledge test phase for each category of patterns for the participants who performed the WPT above chance level.

3.3.2 | 7.5–12-year-old children

Considering the entire block of training trials, we found differences in the normalized number of correct choices for the different categories of patterns (Figure 3A; $F_{(5,70)} = 5.620$, $p < .001$, $\eta^2_p = .286$). Group performance was above chance for all categories of patterns except when the two highly predictive cues were present simultaneously: 0001 and 1000 (one highly predictive cue; $M = .81$, $SD = .12$, $t_{(14)} = 9.862$, $p < .001$, $d_z = 2.546$), 0011 and 1100 (two congruent cues; $M = .81$, $SD = .10$, $t_{(14)} = 12.501$, $p < .001$, $d_z = 3.228$), 0101 and 1010 (one highly predictive cue and one incongruent less predictive cue; $M = .68$, $SD = .19$, $t_{(14)} = 3.818$, $p = .002$, $d_z = .986$), 0111 and 1110 (two congruent cues and one incongruent less predictive cue; $M = .72$, $SD = .13$, $t_{(14)} = 6.310$, $p < .001$, $d_z = 1.629$), 0010 and 0100 (one less predictive cue; $M = .69$, $SD = .18$, $t_{(14)} = 4.008$, $p = .001$, $d_z = 1.035$), 1011 and 1101 (two congruent cues and one incongruent highly predictive cue; $M = .58$, $SD = .20$, $t_{(14)} = 1.532$, $p = .148$, $d_z = .395$).

Importantly, 7.5–12-year-old children who performed the task above chance had a higher normalized number of correct choices for the category that included patterns 0011 and 1100 than the categories that included patterns 0101 and 1010, 0111 and 1110, 0010 and 0100, and 1011 and 1101 (all $p < .05$; Table 4). These results suggested that congruent less predictive cues acted synergistically and that incongruent less predictive cues acted antagonistically to influence the participants' choices in presence of one highly predictive cue. Although the presence of a less predictive cue congruent with a highly predictive cue did not improve performance, as compared to when only one highly predictive cue was present (0011 and 1100 vs. 0001 and 1000, $p = .863$), this was likely due to a ceiling effect. The predominant influence of the two highly predictive cues was further supported by the findings that participants exhibited a higher normalized number of correct choices for patterns 0001 and 1000 than for patterns 0101 and 1010, 0111 and 1110, 0010 and 0100, and 1011 and 1101 (all $p < .05$) and a higher normalized number of correct choices for 0111 and 1110 than 1011 and 1101, and for 0010 and 0100 than 1011 and 1101 (both $p < .05$). The other comparisons were not statistically significant (Table 4). Thus, again, the performance of 7.5–12-year-old children for each pattern was similar to that of adults (Bochud-Fragnière et al., 2022; Figure 3A; see also Table S4).

In accordance with the analyses of training trials, overall group performance on the four cue knowledge test trials was above chance for the category including the highly predictive cues: 0001 and 1000 ($M = 1.73$, $SD = .59$, $t_{(14)} = 4.785$, $p < .001$, $d_z = 1.235$), but it failed to reach the predefined level of statistical significance for the category including the less predictive cues: 0010 and 0100 ($M = 1.33$, $SD = .72$, $t_{(14)} = 1.784$, $p = .096$, $d_z = .461$). However, a comparison of the participants' performance between the two different categories of patterns did not show a statistically significant difference between highly predictive cues and less predictive cues (Figure 3B; $t_{(14)} = 1.572$, $p = .138$, $d_z = .406$).

In sum, although 7.5–12-year-old children learned all the individual cue–outcome associations (i.e., for both highly and less predictive cues), the impact of the two less predictive cues was weaker than expected if they would have followed the multicue strategy. Indeed, the two less predictive cues only influenced performance when one or no highly predictive cue was present. In contrast, they did not influence performance when the two opposite highly predictive cues were present simultaneously. This pattern of results is thus more consistent with the hierarchical strategy: 7.5–12-year-old children based their choices primarily on the presence or absence of the highly predictive cues and considered the two less predictive cues secondarily.

3.3.3 | Adults

As shown in Bochud-Fragnière et al. (2022), considering the entire block of training trials, we found differences in the normalized number of correct choices for the different categories of patterns (Figure 3A; $F_{(5,95)} = 14.472$, $p < .001$, $\eta^2_p = .432$). Group performance was above chance for all categories of patterns except when the two highly predictive cues were present simultaneously: 0011 and 1100 (two congruent cues; $M = .88$, $SD = .07$, $t_{(19)} = 25.259$, $p < .001$, $d_z = 5.648$), 0001 and 1000 (one highly predictive cue; $M = .83$, $SD = .14$, $t_{(19)} = 10.174$, $p < .001$, $d_z = 2.275$), 0111 and 1110 (two congruent cues and one incongruent less predictive cue; $M = .73$, $SD = .13$, $t_{(19)} = 7.650$, $p < .001$, $d_z = 1.711$), 0010 and 0100 (one less predictive cue; $M = .69$, $SD = .19$, $t_{(19)} = 4.457$, $p < .001$, $d_z = .997$), 0101 and 1010 (one highly predictive cue and one incongruent less predictive cue; $M = .66$, $SD = .20$, $t_{(19)} = 3.638$, $p = .002$, $d_z = .813$), 1011 and 1101 (two congruent cues and one incongruent highly predictive cue; $M = .54$, $SD = .17$, $t_{(19)} = 1.010$, $p = .325$, $d_z = .226$).

Importantly, post hoc analyses revealed that the participants who performed the task above chance had a higher normalized number of correct choices for the category that included patterns 0011 and 1100 than the categories that included patterns 0111 and 1110, 0010 and 0100, 0101 and 1010, and 1011 and 1101 (all $p < .001$), suggesting that congruent less predictive cues acted synergistically and that incongruent less predictive cues acted antagonistically (Table 4). Although the presence of a less predictive cue congruent with a highly predictive cue did not improve performance, as compared to when only one highly predictive cue was present (0011 and 1100 vs. 0001 and 1000, $p = .106$), this was likely due to a ceiling effect. The predominant influence of the two highly predictive cues was further supported by the findings that participants exhibited a higher normalized number of correct choices for patterns 0001 and 1000 than for patterns 0111 and 1110, 0101 and 1010, 1011 and 1101, 0010 and 0100 (all $p < .05$), and a higher normalized number of correct choices for 0111 and 1110 than 1011 and 1101, and for 0010 and 0100 than 1011 and 1101 (both $p < .05$). Although participants had a higher normalized number of correct choices for 0101 and 1010 than 1011 and 1101, this comparison failed to reach the predefined level of statistical significance ($p = .059$, two-tailed).

In accordance with training trials analyses, overall group performance on the four cue knowledge test trials was above chance for the two categories of patterns: 0001 and 1000 ($M = 1.84$, $SD = .38$, $t_{(18)} = 9.798$, $p < .001$, $d_z = 2.248$) and 0010 and 0100 ($M = 1.42$, $SD = .61$, $t_{(18)} = 3.024$, $p = .007$, $d_z = .694$), confirming that participants knew the outcomes associated with the two highly predictive cues (0001 and 1000) and the two less predictive cues (0010 and 0100). Moreover, a comparison of the participants' performance on the cue knowledge test trials on these two different categories of patterns showed that participants had superior performance for highly predictive cues than for less predictive cues (Figure 3B; $t_{(18)} = 2.650$, $p = .016$, $d_z = .608$).

Altogether, these data show that young healthy adults learned the relationships between individual cues and outcomes, for both highly and less predictive cues. However, the influence of the two less predictive cues was weaker than is predicted by the multicue strategy. The two less predictive cues influenced performance when only one or no highly predictive cue was present, but not when the two highly predictive cues associated with the opposite outcomes were present simultaneously. This pattern of results is thus consistent with the hierarchical strategy: adults based their choices primarily on the presence or absence of the highly predictive cues and considered the two less predictive cues secondarily.

4 | DISCUSSION

The present study was designed to characterize the abilities and learning strategies of children between 3 and 12 years of age to solve the WPT. As different learning strategies can enable above chance performance (Ashby & Maddox, 2005; Bochud-Fragnière et al., 2022; Gluck et al., 2002; Meeter et al., 2006), we implemented new pattern analyses introduced by Bochud-Fragnière et al. (2022) to characterize the learning strategies involved in the resolution of the task throughout childhood. None of the 3–5.5-year-old children performed above chance. About one third of 5.5–7.5-year-old children performed above chance, relying on the two most predictive cues and neglecting the two less predictive cues. In contrast, about 80% of 7.5–12-year-old children performed above chance, relying on a hierarchical strategy. Similar to adults, 7.5–12-year-old children considered the highly predictive cues primarily and the less predictive cues secondarily. Although both the highly and the less predictive cues influenced their choices, the highly predictive cues had a greater influence. Indeed, when the two highly predictive cues were present simultaneously, the presence of one of the less predictive cues did not influence their choices. In sum, we found no evidence that any of the three groups of children relied on striatal-dependent learning and memory strategies to solve the task. Instead, our findings indicate that the evolution from middle to late childhood of the learning strategies used by TD children to solve the WPT may reflect the increasing ability to use a conditional strategy associated with the development of the hippocampus-dependent learning and memory system.

4.1 | Comparison with previous findings

Our present findings expand upon previous findings reporting that 8–12-year-old children are able to solve the ICT (Mayor-Dubois et al., 2016), by showing that children as young as 5.5 years of age are able to solve the WPT. More importantly, our pattern analyses revealed age-related differences in learning strategies. Specifically, we showed that as a group, 5.5–7.5-year-old children relied primarily on the two most predictive cues strategy. In contrast, as a group 7.5–12-year-old children relied primarily on the hierarchical strategy, similar to young adults. Mayor-Dubois et al. (2016) used the strategy analyses designed by Gluck et al. (2002) and reported changes in learning strategy from 8 to 12 years of age. However, as previously shown by Bochud-Fragnière et al. (2022), the present findings further confirm that the strategy analyses developed by Gluck et al. (2002) and Meeter et al. (2006) are inadequate to infer the learning strategies in the WPT, because the goodness of fit of different strategies cannot be distinguished statistically. Moreover, as strategy analyses require the a priori definition of all possible strategies to avoid misclassification

(Meeter et al., 2006), and as they only considered three possible learning strategies initially defined by Gluck et al. (2002), neither the two most predictive cues strategy nor the hierarchical strategy (Bochud-Fragnière et al., 2022) was considered by Mayor-Dubois et al. (2016). Although about 80% of children between 7.5 and 12 years of age we tested performed above chance level, it is possible that more subtle differences in learning strategies exist in children of this age, as suggested by Mayor-Dubois et al. (2016). However, this seems unlikely as the learning strategy identified for this group of children in the present study was the same as that of the young adult group. Unfortunately, as neither the strategy analyses (Gluck et al., 2002), nor the pattern analyses (Bochud-Fragnière et al., 2022), can be used to determine the learning strategy at the individual level, more in-depth analyses of individual strategies cannot be performed. Nevertheless, it would be interesting to reanalyze the data published by Mayor-Dubois et al. (2016) with the new pattern analyses to confirm our current findings with a larger cohort of children between 8 and 12 years of age.

4.2 | Age-related changes in performance and strategy

We showed that the youngest child who solved the WPT was 5.75 years old, and that there was an increase in the number of participants performing above chance level with age. Importantly, the proportion of successful 7.5–12-year-old children was not significantly lower than that of adults. More importantly, as a group, the performance of all children who scored above chance level was similar to the performance of adults who scored above chance. This can likely be explained by the fact that no participant ever performs without making any errors across the 100 training trials, and that overall successful performance is very similar irrespective of the specific learning strategy that participants may follow (Table 2 and Supporting Information section). Thus, as previously discussed (Bochud-Fragnière et al., 2022; Gluck et al., 2002; Meeter et al., 2006), global performance cannot be used to infer the cognitive processes involved in the resolution of the WPT. In contrast, the pattern analyses designed by Bochud-Fragnière et al. (2022) can be used to infer the learning strategy at the group level.

4.2.1 | 5.5–7.5-year-old children used the two most predictive cues strategy

Our analyses demonstrated that, as a group, successful 5.5–7.5-year-old children learned the correct outcomes associated with all different patterns except 0010 and 0100 and 1011 and 1101 (i.e., the patterns for which the associations between the less predictive cues and the outcomes are necessary to answer correctly). This indicates that 5.5–7.5-year-old children learned the correct outcomes associated with the highly predictive cues, but not with the less predictive cues. Their performance on the cue knowledge test trials supports this interpretation.

The reliance on the two most predictive cues strategy is also supported by comparison of the performance between the different categories of patterns. First, 5.5–7.5-year-old children performed better on patterns that contained only one highly predictive cue (0001 and 1000, 0011 and 1100, and 0101 and 1010) than patterns that contained no highly predictive cue or both contradictory highly predictive cues (0010 and 0100 and 1011 and 1101). The fact that 5.5–7.5-year-old children performed quantitatively better at 0111 and 1110 than 0010 and 0100 and 1011 and 1101 also supports this view. Second, there was no beneficial effect of a congruent less predictive cue as 5.5–7.5-year-old children did not perform better with patterns 0011 and 1100 than patterns 0001 and 1000 (although this could be linked to a ceiling effect), or with patterns 0111 and 1110 than patterns 0101 and 1010. Third, there was no negative impact of an incongruent less predictive cue on performance as they did not perform better with patterns 0001 and 1000 or 0011 and 1100 than patterns 0101 and 1010. Nevertheless, 5.5–7.5-year-old children performed better with patterns 0011 and 1100 than 0111 and 1110, and with a trend for patterns 0001 and 1000 than 0111 and 1110 ($p = .074$). Although this finding could suggest that incongruent less predictive cues might have had some influence, it could also be due to the presence of three cues, instead of one or two, which might have acted as distractors or increased cognitive load. This latter explanation is supported by the fact that there was no negative or positive impact of less predictive cues in any other pattern that included only two cues. Finally, the fact that performance on 0010 and 0100 did not differ from 1011 and 1101 indicates that the less predictive cue did not influence performance even when presented alone.

Altogether, the above-described analyses show that 5.5–7.5-year-old children relied on only the two highly predictive cues to solve the WPT, neglecting the two less predictive cues. This two most predictive cue strategy was previously identified by Bochud-Fragnière et al. (2022) but was not identified in the strategy analyses designed by Gluck et al. (2002)or Meeter et al. (2006), or used by Mayor-Dubois et al. (2016).

4.2.2 | 7.5–12-years-old children and adults used the hierarchical strategy

Our analyses demonstrated that most 7.5–12-year-old children and 20–30-year-old adults chose the correct outcomes associated with all patterns except for the two patterns containing the two highly predictive cues and one less predictive cue (1011 and 1101). Their performance with patterns 0010 and 0100 across training trials showed that they knew which outcomes were associated with the less predictive cues. The fact that the performance of 7.5–12-year-old children on pattern 0100 was only tendentially associated with the correct outcome suggests a progressive acquisition of the association between the two less predictive cues and their respective outcomes. An alternative explanation may be that they considered the less predictive cues less consistently than adults. Nevertheless, support for their progressive acquisition is provided by the fact that performance with the pattern

0010 (i.e., the other pattern that included only one less predictive cue) was above chance level, and the fact that incongruent less predictive cues had a negative impact on performance. Note that we were unable to perform our analyses on blocks of 50 trials because there were insufficient numbers of presentations of the different types of patterns for statistical comparison. We were thus unable to determine whether there may be a switch of strategies from the two most predictive cues to the hierarchical strategy over the course of the 100 training trials for 7.5–12-year-old children.

Altogether, all our analyses suggest that both 7.5–12-year-old children and 20–30-year-old adults employed a hierarchical strategy. Indeed, the use of a hierarchical strategy is supported by the fact that both the highly and the less predictive cues had an impact on performance, but the impact of the highly predictive cues was stronger than that of the less predictive cues. First, there was a beneficial effect of a congruent highly predictive cue as 7.5–12-year-old children and adults performed better on 0011 and 1100 than on 0010 and 0100. In contrast, there was a detrimental effect of an incongruent highly predictive cue as both groups performed better on 0011 and 1100 than on 1011 and 1101. Second, there was a detrimental effect of incongruent less predictive cues as 7.5–12-year-old children and adults performed better on 0001 and 1000 than on 0101 and 1010, and better on 0011 and 1100 than on 0111 and 1110. Moreover, our data also suggest a beneficial effect of a congruent less predictive cue in adults, because they performed better on 0011 and 1100 than on 0001 and 1000, although the difference was only tendential likely due to a ceiling effect. Third, we showed that a highly predictive cue had a stronger impact than a less predictive cue on participants' choices as they performed better on 0111 and 1110 than on 1011 and 1101. Importantly, the less predictive cues did not influence performance in a manner predicted by the multicue strategy, as the less predictive cues failed to guide choices in the presence of the two highly predictive cues. Nonetheless, the less predictive cues were not treated as simple distractors, but rather cues influencing choices, as evidenced by the fact that 7.5–12-year-old children and 20–30-year-old adults performed better with patterns 0001 and 1000 and 0011 and 1100 than with patterns 0101 and 1010.

In sum, although both 7.5–12-year-old children and 20–30-year-old adults paid attention to all four cues when responding, they processed the cues hierarchically. They primarily based their responses on the presence or absence of the two highly predictive cues and considered the two less predictive cues secondarily, except when both highly predictive cues were present simultaneously. Importantly, as discussed in Bochud-Fragnière et al. (2022), although the present data suggest a hierarchical treatment of the cues, neither our experimental design nor our analyses allowed us to determine whether this hierarchical treatment of information follows a sequential temporal organization with highly predictive cues being considered first, and less predictive cues being considered second.

4.2.3 | Children who did not solve the task

No child younger than 5.5 years of age and numerous older children were unable to solve the WPT. This shows that these participants were unable to reliably apply any of the response strategies that can be used to perform above chance level (Table 2). There may be several explanations for this inability. First, we showed that two participants, who performed statistically worse than would be expected if they chose the outcome randomly, alternated systematically between the two outcomes (sun and snow). Second, the probabilistic nature of the task, and thus the lack of absolute consistency in the given feedback, may negatively impact learning in certain participants who doubt their understanding of the cue–outcome associations or the task. Third, the presence of multiple cues and increased cognitive load may impact the performance of some participants who may have started learning with a simple strategy that focused on single cues. Fourth, a tendency to choose one particular outcome, for example, sun, irrespective of the cues present on the screen may also lead some participants to fail the task. In addition, as the outcome sun was always present on the right side of the screen, this may have led to a preference to choose this outcome because most participants (all adults and about 90% of the children) used their right hand to perform the task. Although we could have counterbalanced the position of the outcome to avoid this behavior, this might have increased errors due to inattention when choosing the outcome.

In sum, about one third of 5.5–7.5-year-old children solved the task using the two most predictive cues strategy, whereas most 7.5–12-year-old children and 20–30-year-old adults solved the task using a hierarchical strategy. Although our experimental design and analyses enabled us to define the learning strategies used by specific groups of participants to solve the WPT, we did not assess which neural substrates are associated with the different strategies. Nevertheless, knowing the characteristics of different learning and memory systems allows us to make assumptions regarding the neural substrates that may contribute to successful performance on the WPT.

4.3 | Hippocampus-dependent memory can support WPT performance

If, similar to spatial response learning, the WPT specifically assessed striatum-dependent "implicit" or procedural memory as originally presumed (Knowlton et al., 1994), young children between 3 and 5.5 years of age should be able to solve the WPT using a stimulus-response strategy, such as the multicue strategy. Instead, we found that the youngest individual who solved the task was 5.75 years old, and as a group children of this age used the two most predictive cues strategy and not the striatum-dependent multicue strategy. The reasons for this discrepancy may be manifold. First, it is possible that very young children can use response learning in spatial and nonspatial contexts, but that other variables such as cognitive load and probabilistic feedback may negatively influence this type of learning in the WPT. Second, the WPT may assess learning and memory capacities that depend on other striatal circuits exhibiting a different developmental profile than those implicated in spatial response learning. Third, other strategies, such as hippocampus-dependent learning strategies, may underlie above chance performance on the WPT (Ashby & O'Brien, 2005; Foerde et al., 2006, 2007; Newell et al., 2007; Poldrack & Rodriguez, 2004; Price, 2009). Our pattern analyses and previous work on other types of hippocampus-dependent memory support this third hypothesis (Bochud-Fragnière et al., 2022; Bostelmann et al., 2017, 2018; Bostelmann et al., 2020).

Adults and most children 7.5 years of age and older solved the WPT using the hierarchical strategy. To successfully implement this strategy, participants must be able to learn and remember which of four relatively similar cues is associated with which outcome, even though there is not a one-to-one association between cues and outcomes as two cues are associated with one outcome, and the other two cues with the other outcome. Moreover, participants must be able to construct a conditional hierarchy between the items associated with the same outcome, recognizing that the predictive value of cues can change depending on what other cues are present simultaneously. Finally, they must be able to learn, or at least not be negatively impacted by the fact that the cues are associated probabilistically, such that sometimes the correct feedback is not given when the most likely outcome is chosen.

Numerous studies have revealed the role of the hippocampus in conditional learning (Andersen et al., 2006; Greene et al., 2006; Ridley et al., 1995; Smith & Squire, 2005; Zeithamova et al., 2012). Conditional learning is the ability to learn which response is appropriate based on context and independent of its intrinsic value, for example, that response C is correct in the context of B but incorrect in the context of D (Hirsh, 1980; Ridley et al., 1995). In these relational representations, items can be encoded both as distinct objects and as related in a larger schema (e.g., in scenes and events) (Cohen & Eichenbaum, 1993; Eichenbaum & Cohen, 2014). Conditional learning thus allows the differentiation of similar situations or items. For example, Zeithamova et al. (2012) showed that participants who made successful inferential judgments had greater initial engagement of the hippocampus during the training of overlapping pairs of items. Correct inferential judgments revealed that participants had learned conditionally the hierarchical relationships between the items and thus demonstrated the role of the hippocampus in conditional learning in humans.

In sum, the fact that the hierarchical strategy depends on conditional learning suggests that the hippocampus was implicated in solving our abridged WPT in most TD children older than 7.5 years of age and young adults. Age-related improvements in hippocampus-dependent memory processes may thus underlie the development of the ability of children to solve the WPT. Future experiments should employ functional brain imaging to confirm the role of the hippocampus in probabilistic classification learning in TD children.

5 | CONCLUSION

Our present findings expand upon previous findings reporting that 8–12-year-old children are able to solve a modified version of the WPT (Mayor-Dubois et al., 2016), by showing that children as young as 5.75 years of age are able to solve the WPT. More importantly, as different learning strategies enable above chance performance, we implemented new pattern analyses designed by Bochud-Fragnière et al. (2022) to characterize the learning strategies involved in the resolution of the task throughout childhood. We showed that 5.5–7.5-year-old children relied on the two most predictive cues strategy, whereas 7.5–12-year-old children and adults relied on the hierarchical strategy. Altogether, these findings indicate that the learning strategies used by TD children to solve the WPT evolve from middle to late childhood and may reflect improvements in hippocampus-dependent processes, such as conditional learning.

ACKNOWLEDGMENTS

The authors would like to thank all the participants for making this study possible. This work was supported by grant 100019_165481 from the Swiss National Science Foundation to PBL and PL, and the Faculty of Social and Political Sciences of the University of Lausanne. The funders had no role in the design of the study, in the collection, analysis, or interpretation of the data, in the preparation of this manuscript nor in the decision to submit it for publication.

Open access funding provided by Universite de Lausanne.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ORCID

Emilie Bochud-Fragnière https://orcid.org/0000-0002-7453-9071
Pierre Lavenex https://orcid.org/0000-0002-9278-1312
Pamela Banta Lavenex https://orcid.org/0000-0001-8868-2912

REFERENCES

Acredolo, L. P. (1978). Development of spatial orientation in infancy. *Developmental Psychology*, *14*(3), 224–234. https://doi.org/10.1037/0012-1649.14.3.224

Acredolo, L. P., & Evans, D. (1980). Developmental changes in the effects of landmarks on infant spatial behavior. *Developmental Psychology*, *16*(4), 312–318. https://doi.org/10.1037/0012-1649.16.4.312

Andersen, P., Morris, R., Amaral, D., Bliss, T., & O'Keefe, J. (2006). *The hippocampus book*. Oxford University Press.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, *56*, 149–178. https://doi.org/10.1146/annurev.psych.56.091103.070217

Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Sciences*, *9*(2), 83–89.

Bochud-Fragnière, E., Banta Lavenex, P., & Lavenex, P. (2022). What is the Weather Prediction Task good for? A new analysis of learning strategies reveals how young adults solve the task. *Frontiers in Psychology*, *13*, 886339. https://doi.org/10.3389/fpsyg.2022.886339

Bostelmann, M., Bochud-Fragniere, E., Costanzo, F., Di Vara, S., Menghini, D., Vicari, S., Lavenex, P., & Banta Lavenex, P. (2017). Dissociation of spatial memory systems in Williams syndrome. *Hippocampus*, *27*(11), 1192–1203. https://doi.org/10.1002/hipo.22764

Bostelmann, M., Costanzo, F., Martorana, L., Menghini, D., Vicari, S., Banta Lavenex, P., & Lavenex, P. (2018). Low-resolution place and response learning capacities in Down syndrome. *Frontiers in Psychology*, *9*, 1–17. https://doi.org/10.3389/fpsyg.2018.02049

Bostelmann, M., Lavenex, P., & Banta Lavenex, P. (2020). Children five-to-nine years old can use path integration to build a cognitive map without vision. *Cognitive Psychology*, *121*, 101307. https://doi.org/10.1016/j.cogpsych.2020.101307

Brown, J., Aczel, B., Jimenez, L., Kaufman, S. B., & Grant, K. P. (2010). Intact implicit learning in autism spectrum conditions. *Quarterly Journal of Experimental Psychology*, *63*(9), 1789–1812. https://doi.org/10.1080/17470210903536910

Carmer, S. G., & Swanson, M. R. (1973). Evaluation of 10 pairwise multiple comparison procedures by Monte-Carlo methods. *Journal of the American Statistical Association*, *68*(341), 66–74. https://doi.org/10.2307/2284140

Chang, Q., & Gold, P. E. (2003a). Intra-hippocampal lidocaine injections impair acquisition of a place task and facilitate acquisition of a response task in rats. *Behavioral Brain Research*, *144*(1–2), 19–24.

Chang, Q., & Gold, P. E. (2003b). Switching memory systems during learning: Changes in patterns of brain acetylcholine release in the hippocampus and striatum in rats. *Journal of Neuroscience*, *23*(7), 3001–3005.

Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia, and the hippocampal system*. MIT Press.

Cohen, N. J., Eichenbaum, H., Deacedo, B. S., & Corkin, S. (1985). Different memory systems underlying acquisition of procedural and declarative knowledge. *Annals of the New York Academy of Sciences*, *444*(1), 54–71.

Cornell, E. H., & Heth, C. D. (1979). Response versus place learning by human infants. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(2), 188.

Eichenbaum, H., & Cohen, N. J. (2014). Can we reconcile the declarative memory and spatial navigation views on hippocampal function? *Neuron*, *83*(4), 764–770. https://doi.org/10.1016/j.neuron.2014.07.032

Fernandez, V. G., Asarnow, R., Hodges, M., & Nuechterlein, K. H. (2021). Linguistic and neurocognitive correlates of probabilistic classification learning in schizophrenia. *Schizophrenia Research: Cognition*, *26*, 100209. https://doi.org/10.1016/j.scog.2021.100209

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(31), 11778–11783. https://doi.org/10.1073/pnas.0602659103

Foerde, K., Poldrack, R. A., Khan, B. J., Sabb, F. W., Bookheimer, S. Y., Bilder, R. M., Guthrie, D., Granholm, E., Nuechterlein, K. H., Marder, S. R., & Asarnow, R. F. (2008). Selective corticostriatal dysfunction in schizophrenia: Examination of motor and cognitive skill learning. *Neuropsychology*, *22*(1), 100–109. https://doi.org/10.1037/0894-4105.22.1.100

Foerde, K., Poldrack, R. A., & Knowlton, B. J. (2007). Secondary-task effects on classification learning. *Memory and Cognition*, *35*(5), 864–874. https://doi.org/10.3758/bf03193461

Gabay, Y., & Goldfarb, L. (2017). Feedback-based probabilistic category learning is selectively impaired in attention/hyperactivity deficit disorder. *Neurobiology of Learning and Memory*, *142*, 200–208. https://doi.org/10.1016/j.nlm.2017.04.012

Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do people solve the "weather prediction" task?: Individual variability in strategies for probabilistic category learning. *Learning and Memory*, *9*(6), 408–418. https://doi.org/10.1101/lm.45202

Gomar, J. J., Pomarol-Clotet, E., Sarro, S., Salvador, R., Myers, C. E., & McKenna, P. J. (2011). Procedural learning in schizophrenia: Reconciling the discrepant findings. *Biological Psychiatry*, *69*(1), 49–54. https://doi.org/10.1016/j.biopsych.2010.07.013

Greene, A. J., Gross, W. L., Elsinger, C. L., & Rao, S. M. (2006). An fMRI analysis of the human hippocampus: Inference, context, and task awareness. *Journal of Cognitive Neuroscience*, *18*(7), 1156–1173. https://doi.org/10.1162/jocn.2006.18.7.1156

Hartley, T., & Burgess, N. (2005). Complementary memory systems: Competition, cooperation and compensation. *Trends in Neurosciences*, *28*(4), 169–170.

Hartley, T., Maguire, E. A., Spiers, H. J., & Burgess, N. (2003). The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron*, *37*(5), 877–888. https://doi.org/10.1016/S0896-6273(03)00095-3

Hirsh, R. (1980). The hippocampus, conditional operations, and cognition. *Physiological Psychology*, *8*(2), 175–182. https://doi.org/10.3758/BF03332848

Hopkins, R. O., Myers, C. E., Shohamy, D., Grossman, S., & Gluck, M. (2004). Impaired probabilistic category learning in hypoxic subjects with hippocampal damage. *Neuropsychologia*, *42*(4), 524–535. https://doi.org/10.1016/j.neuropsychologia.2003.09.005

Horan, W. P., Green, M. F., Knowlton, B. J., Wynn, J. K., Mintz, J., & Nuechterlein, K. H. (2008). Impaired implicit learning in schizophrenia. *Neuropsychology*, *22*(5), 606–617. https://doi.org/10.1037/a0012602

Karcher, N. R., Hua, J. P. Y., & Kerns, J. G. (2019). Probabilistic category learning and striatal functional activation in psychosis risk. *Schizophrenia Bulletin*, *45*(2), 396–404.

Kemeny, F., & Lukacs, A. (2010). Impaired procedural learning in language impairment: Results from probabilistic categorization. *Journal of Clinical and Experimental Neuropsychology*, *32*(3), 249–258. https://doi.org/10.1080/13803390902971131

Keri, S., Kelemen, O., Szekeres, G., Bagoczky, N., Erdelyi, R., Antal, A., Benedek, G., & Janka, Z. (2000). Schizophrenics know more than they can tell: Probabilistic classification learning in schizophrenia. *Psychological Medicine*, *30*(1), 149–155. https://doi.org/10.1017/s0033291799001403

Keri, S., Szlobodnyik, C., Benedek, G., Janka, Z., & Gadoros, J. (2002). Probabilistic classification learning in Tourette syndrome. *Neuropsychologia*, *40*(8), 1356–1362. https://doi.org/10.1016/s0028-3932(01)00210-x

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science*, *273*(5280), 1399–1402. https://doi.org/10.1126/science.273.5280.1399

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning and Memory*, *1*(2), 106–120. https://doi.org/10.1101/lm.1.2.106

Knowlton, B. J., Squire, L. R., Paulsen, J. S., Swerdlow, N. R., Swenson, M., & Butters, N. (1996). Dissociations within nondeclarative memory in Huntington's disease. *Neuropsychology*, *10*(4), 538–548. https://doi.org/10.1037/0894-4105.10.4.538

Labouliere, C. D., Terranova, K., Steinglass, J., & Marsh, R. (2016). Implicit learning on a probabilistic classification task in adults and adolescents with Bulimia Nervosa. *Journal of Psychiatric Research*, *77*, 35–41. https://doi.org/10.1016/j.jpsychires.2016.02.004

Lagnado, D. A., Newell, B. R., Kahan, S., & Shanks, D. R. (2006). Insight and strategy in multiple-cue learning. *Journal of Experimental Psychology: General*, *135*(2), 162.

Lee, A. S., Duman, R. S., & Pittenger, C. (2008). A double dissociation revealing bidirectional competition between striatum and hippocampus during learning. *Proceedings of the National Academy of Sciences*, *105*(44), 17163–17168.

Marsh, R., Alexander, G. M., Packard, M. G., Zhu, H., Wingard, J. C., Quackenbush, G., & Peterson, B. S. (2004). Habit learning in Tourette syndrome: A translational neuroscience approach to a developmental psychopathology. *Archives of General Psychiatry*, *61*(12), 1259–1268. https://doi.org/10.1001/archpsyc.61.12.1259

Mayor-Dubois, C., Maeder, P., Zesiger, P., & Roulet-Perez, E. (2010). Visuomotor and cognitive procedural learning in children with basal ganglia pathology. *Neuropsychologia*, *48*(7), 2009–2017.

Mayor-Dubois, C., Zesiger, P., Van der Linden, M., & Roulet-Perez, E. (2016). Procedural learning: A developmental study of motor sequence learning and probabilistic classification learning in school-aged children. *Child Neuropsychology*, *22*(6), 718–734. https://doi.org/10.1080/09297049.2015.1058347

McDonald, R. J., Devan, B. D., & Hong, N. S. (2004). Multiple memory systems: The power of interactions. *Neurobiology of Learning and Memory*, *82*(3), 333–346. https://doi.org/10.1016/j.nlm.2004.05.009

McDonald, R. J., & White, N. M. (1993). A triple dissociation of memory systems: Hippocampus, amygdala and dorsal striatum. *Behavioral Neuroscience*, *107*(1), 3–22.

Meeter, M., Myers, C. E., Shohamy, D., Hopkins, R. O., & Gluck, M. A. (2006). Strategies in probabilistic categorization: Results from a new way of analyzing performance. *Learning and Memory*, *13*(2), 230–239. https://doi.org/10.1101/lm.43006

Newcombe, N., Huttenlocher, J., Drummey, A. B., & Wiley, J. G. (1998). The development of spatial location coding: Place learning and dead reckoning in the second and third years. *Cognitive Development*, *13*(2), 185–200. https://doi.org/10.1016/S0885-2014(98)90038-7

Newell, B. R., Lagnado, D. A., & Shanks, D. R. (2007). Challenging the role of implicit processes in probabilistic category learning. *Psychonomic Bulletin and Review*, *14*(3), 505–511. https://doi.org/10.3758/BF03194098

Packard, M. G., Hirsh, R., & White, N. M. (1989). Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: Evidence for multiple memory systems. *Journal of Neuroscience*, *9*(5), 1465–1472.

Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature*, *414*(6863), 546–550. https://doi.org/10.1038/35107080

Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia*, *41*(3), 245–251. Pii S0028-3932(02)00157-4 https://doi.org/10.1016/S0028-3932(02)00157-4

Poldrack, R. A., & Rodriguez, P. (2004). How do memory systems interact? Evidence from human classification learning. *Neurobiology of Learning and Memory*, *82*(3), 324–332. https://doi.org/10.1016/j.nlm.2004.05.003

Price, A. L. (2009). Distinguishing the contributions of implicit and explicit processes to performance of the weather prediction task. *Memory and Cognition*, *37*(2), 210–222. https://doi.org/10.3758/MC.37.2.210

Ribordy Lambert, F., Jabes, A., Banta Lavenex, P., & Lavenex, P. (2013). Development of allocentric spatial memory abilities in children from 18 months to 5 years of age. *Cognitive Psychology*, *66*(1), 1–29. https://doi.org/10.1016/j.cogpsych.2012.08.001

Ribordy Lambert, F., Lavenex, P., & Banta Lavenex, P. (2015). Improvement of allocentric spatial memory resolution in children from 2 to 4 years of age. *International Journal of Behavioral Development*, *39*(4), 318–331. https://doi.org/10.1177/0165025415584808

Ribordy Lambert, F., Lavenex, P., & Banta Lavenex, P. (2017). The "when" and the "where" of single-trial allocentric spatial memory performance in young children: Insights into the development of episodic memory. *Developmental Psychobiology*, *59*, 185–196. https://doi.org/10.1002/dev.21479

Ridley, R. M., Timothy, C. J., Maclean, C. J., & Baker, H. F. (1995). Conditional learning and memory impairments following neurotoxic lesion of the CA1 field of the hippocampus. *Neuroscience*, *67*(2), 263–275. https://doi.org/10.1016/0306-4522(95)00063-O

Rothman, K. (1990). No adjustments are needed for multiple comparisons. *Epidemiology (Cambridge, Mass.)*, *1*, 43–46. https://doi.org/10.1097/00001648-199001000-00010

Saville, D. J. (1990). Multiple comparison procedures – The practical solution. *American Statistician*, *44*(2), 174–180. https://doi.org/10.2307/2684163

Shohamy, D., Myers, C. E., Grossman, S., Sage, J., Gluck, M. A., & Poldrack, R. A. (2004). Cortico-striatal contributions to feedback-based learning: Converging data from neuroimaging and neuropsychology. *Brain*, *127*, 851–859. https://doi.org/10.1093/brain/awh100

Shohamy, D., Myers, C. E., Onlaor, S., & Gluck, M. A. (2004). Role of the basal ganglia in category learning: How do patients with Parkinson's disease learn? *Behavioral Neuroscience*, *118*(4), 676–686. https://doi.org/10.1037/0735-7044.118.4.676

Smith, C., & Squire, L. R. (2005). Declarative memory, awareness, and transitive inference. *The Journal of Neuroscience*, *25*(44), 10138–10146. https://doi.org/10.1523/JNEUROSCI.2731-05.2005

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory*, *82*(3), 171–177. https://doi.org/10.1016/j.nlm.2004.06.005

Voermans, N. C., Petersson, K. M., Daudey, L., Weber, B., van Spaendonck, K. P., Kremer, H. P. H., & Fernandez, G. (2004). Interaction between the human hippocampus and the caudate nucleus during route recognition. *Neuron*, *43*(3), 427–435. https://doi.org/10.1016/j.neuron.2004.07.009

Weickert, T. W., Terrazas, A., Bigelow, L. B., Malley, J. D., Hyde, T., Egan, M. F., Weinberger, D. R., & Goldberg, T. E. (2002). Habit and skill learning in schizophrenia: Evidence of normal striatal processing with abnormal cortical input. *Learning and Memory*, *9*(6), 430–442. https://doi.org/10.1101/lm.49102

White, N. M., Packard, M. G., & McDonald, R. J. (2013). Dissociation of memory systems: The story unfolds. *Behavioral Neuroscience*, *127*(6), 813–834. https://doi.org/10.1037/a0034859

Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron*, *75*(1), 168–179. https://doi.org/10.1016/j.neuron.2012.05.010

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Bochud-Fragnière, E., Banta Lavenex, P., & Lavenex, P. B. (2023). When and how do children solve the Weather Prediction Task?. *Developmental Psychobiology*, *65*, e22407. https://doi.org/10.1002/dev.22407