Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Cognition 106 (2008) 325–344
COGNITION
www.elsevier.com/locate/COGNIT

Children's understanding of posterior probability ☆

Vittorio Girotto [a,b,*], Michel Gonzalez [b]

[a] *Department of Arts and Design, University IUAV of Venice, Convento delle Terese, DD 2206, 30123 Venice, Italy*
[b] *CNRS-University of Provence, France*

Received 29 May 2006; revised 8 February 2007; accepted 12 February 2007

Abstract

Do young children have a basic intuition of posterior probability? Do they update their decisions and judgments in the light of new evidence? We hypothesized that they can do so extensionally, by considering and counting the various ways in which an event may or may not occur. The results reported in this paper showed that from the age of five, children's decisions under uncertainty (Study 1) and judgments about random outcomes (Study 2) are correctly affected by posterior information. From the same age, children correctly revise their decisions in situations in which they face a single, uncertain event, produced by an intentional agent (Study 3). The finding that young children have some understanding of posterior probability supports the theory of naive extensional reasoning, and contravenes some pessimistic views of probabilistic reasoning, in particular the evolutionary claim that the human mind cannot deal with single-case probability.

Keywords: Probabilistic reasoning; Posterior evaluation; Children's intuitions; Evolutionary psychology; Choice updating

☆ This manuscript was accepted under the editorship of Jacques Mehler.
* Corresponding author. Tel.: +39 041 2571 351.
E-mail address: vittorio.girotto@iuav.it (V. Girotto).

0010-0277/$ - see front matter
doi:10.1016/j.cognition.2007.02.005

1. Introduction

Do young children have a basic intuition of posterior probability? Are they able to update their decisions and judgments about uncertain events in the light of new evidence? These questions have never been addressed. In fact, readers may find them rather odd given the well-documented difficulties of adults in probability problems that require them to combine different pieces of information (e.g., Gilovich, Griffin, & Kahneman, 2002; Kahneman, Slovic, & Tversky, 1982). The aim of this paper is to answer these questions, by examining preschool and school children's ability to make use of new evidence for predicting an uncertain event. The paper defends the thesis that children are able to take into account new evidence for predicting an uncertain event when they can consider the various possibilities of its occurrence, that is, when they can reason in an extensional way. The paper begins with an analysis of the difficulties that individuals encounter when they have to assess probability by integrating multiple pieces of evidence, and it indicates the conditions in which they can correctly integrate new evidence. Then, the paper shows that children as young as five use posterior evidence for updating their predictions of a random event (Studies 1 and 2) or of a unique event produced by an intentional agent (Study 3). Finally, it discusses the implications of the reported findings for the current views of probabilistic reasoning, in particular the evolutionary thesis according to which the human mind is blind to single event probability.

1.1. Prior and posterior judgments under uncertainty

Current literature on the development of probabilistic reasoning offers sparse evidence of children's ability to reason about uncertain events. The only indication of an early emergence of probabilistic intuitions concerns situations in which the occurrence of an event exclusively depends on the initial distribution of a population. For instance, 5-year-olds predict that one will get a monkey rather than a bird, if one takes one chip at random from a box containing seven chips representing monkeys and three chips representing birds (Brainerd, 1981; see also Acredolo, O'Connor, Banks, & Horobin, 1989; Yost, Siegel, & Andrews, 1962). In other words, when given information about a future process (e.g., taking a chip blindly from a box the content of which is known), preschool children are able to take into account prior possibilities in order to predict an uncertain event. This finding is important, because it challenges the traditional Piagetian view, according to which young children lack the logical abilities to take into account probability ratios (Piaget & Inhelder, 1975). However, it does not establish whether young children have an intuition of posterior probability. In the above situation, one can assess prior probability on the basis of general information about a process. Unlike prior probability, assessing posterior probability requires the integration of prior information with more specific information about the outcome's process. For example, in a well-known problem used to investigate adults' probabilistic intuitions, participants are asked to estimate the probability that a randomly chosen member of a group of people is an engineer (Kahneman & Tversky, 1973). Participants know the proportion of engineers in the group. Hence, they can evaluate the prior probability that the individual is an engineer. But participants also read a short description of the individual. If they combine the two sorts of information (general and specific), they can evaluate the posterior probability that the individual is an engineer. Another well-known problem asks participants to estimate the probability that an individual is affected by a given disease (e.g., Hammerton, 1973). The text of this problem conveys different pieces of information (e.g., the rate of the disease in the population, the result of a diagnostic test undertaken by the individual and its reliability). In this case, combining the successive pieces of information allows participants to estimate the probability that the individual is affected by the disease, if the test is positive. Several decades of research have proven that adults have considerable difficulty in evaluating probability in problems of this kind (e.g., Gilovich et al., 2002; Kahneman et al., 1982). For instance, in the engineer problem, participants base their judgments exclusively on the clues to the person's occupation provided by the individual's description, without taking into account the number of engineers in the group. But when they are given no information whatsoever about the individual, participants correctly estimate the probability that the unspecified individual is an engineer, basing their estimation on the proportion of engineers in the group.

These findings suggest that adults' judgments are based on simple, not probabilistic, heuristics, such as the representativeness of relevant cues. Moreover, they suggest that correct probabilistic intuitions are confined to elementary situations, and that probabilities based on successive pieces of evidence, like posterior probabilities, are at the edge of human competence (e.g., Johnson-Laird, 2006). Consequently, a study investigating whether young children possess an intuition of posterior probability seems doomed to failure. Nevertheless, there is reason to believe that adults are not intrinsically unable to integrate successive pieces of information in order to determine probability. For instance, in the engineer problem, participants can easily represent the information about the number of engineers and non-engineers in the group as a set of possibilities concerning the selected individual (i.e., that individual could be any of the engineers and non-engineers in the group). But they cannot easily relate the specific description of the individual to the information about the proportion of engineers in the group. The difficulty in constructing a homogeneous representation of the two sorts of information can be a major source for erroneous evaluations. Indeed, adults make correct probability judgments in conditions in which they can easily represent all the pieces of information in the *same* set of possibilities (Fox & Levav, 2004; Girotto & Gonzalez, 2001; Sloman, Over, Slovak, & Stibel, 2003). For instance, adults solve the disease problem when the pieces of information relate to different subsets of the same set of chances, that is, when the prevalence of the disease is presented by the number of chances of being infected or not-infected, and the reliability of the test by the number of chances of being tested positive among those of being infected (Girotto & Gonzalez, 2001). This finding proves that individuals are able to accurately evaluate the probability of an event in conditions in which they have to integrate various sources of information. And they do so when they can reason in an *extensional* way, by considering and comparing the possibilities in which the event may or may not occur (Johnson-Laird, Legrenzi, Girotto, Sonino-Legrenzi, & Caverni, 1999).

1.2. Children's extensional intuitions

Given the evidence that adults are able to integrate successive pieces of information in probabilistic reasoning, we hypothesize that children, too, may exhibit such an ability in posterior probability problems. As indicated above, preschool children possess a basic intuition of prior probability, according to which one future outcome appears more likely than another, if the former can be produced in more ways than the latter. This finding indicates that young children make correct judgments by reasoning extensionally about an initial set of possibilities. Are they also able to make correct judgments when they must integrate specific information about a given outcome? We posit that they would do so in the same situations in which adults perform correctly, that is, in situations in which they can relate specific information to a subset of the set of initial possibilities.

Consider the chips presented in Fig. 1. Suppose you take a chip at random, and ask children to predict whether it is black or white. As suggested by previous results, they are likely to answer "black" because they can observe that there are five black and only three white chips. Alternatively, suppose you take a chip, keep it in your hand, and inform children that it is square. To answer correctly, they should consider the subset of possibilities compatible with the new piece of information (i.e., the four squares) and observe that it contains more white than black chips. Are they able to do so, integrating prior and new information? Finally, suppose you ask children to make one prediction before and one prediction after you inform them that your chip is square. They should answer "black" first, but then "white". Are they able to do so by updating their judgment?

Individuating information about a specific event modifies the possibilities to consider, but asks for the same extensional treatment as prior information does.

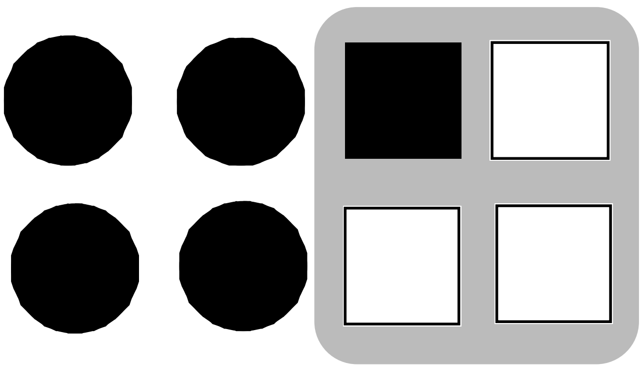

Fig. 1. A prior sample space and the subset (indicated by the grey area) determined by the posterior piece of evidence: "It is a square."

Therefore, children should be able to make correct posterior evaluations, provided that they are able to integrate data from two sources of information. Indeed, evidence exists that 5 and 6-year-olds are able to integrate information over two successive messages and to revise their initial interpretation (Beck & Robinson, 2001). Consequently, children of this age should be able to take into account the chance of an event even in situations in which both sources of information (general and specific) have to be considered. In other words, young children should be able to correctly revise decisions and judgments in the light of specific evidence.

We tested our predictions in three studies, in which preschool and school children dealt with uncertain events, given two successive states of information. In the first two studies, they had to make and revise bets (Study 1) or judgments (Study 2) about random, repeatable outcomes (i.e., drawing a chip from a box). In the final study, they had to make and revise bets about a single event produced by an intentional agent. The final study was aimed to test the diverging predictions derived from the extensional competence view and an alternative view. According to the evolutionary-frequentist hypothesis, the human mind is blind to single-event probability, but contains a module for reasoning about frequencies of multiple, repeatable events (e.g., Cosmides & Tooby, 1996). Hence, naive individuals fail to make correct posterior evaluations in problems concerning a single event because they do not tap into this module; whereas the same individuals do make correct posterior evaluations in problems concerning frequencies of events because they indeed tap into such a module (e.g., Gigerenzer & Hoffrage, 1995). By contrast, the extensional competence view posits that naive individuals are able to make correct posterior evaluations regardless of the type of information about which they have to reason. If they can apply intuitive extensional procedures, naive individuals solve problems which require them to reason about finite sets of cases, possibilities, or frequencies of observations (Girotto & Gonzalez, 2001, 2002; Johnson-Laird et al., 1999). Consequently, investigating young children's ability to evaluate the posterior probability of a single event represents a crucial test of the two hypotheses.

2. Study 1: Updating decisions

In Study 1, we evaluated children's probabilistic intuitions by asking them to bet on one of two possible random outcomes. If children were able to correctly make an implicit assessment of chance, they would prefer the more likely outcome. In some tasks, similar to those used in previous research, children could make a correct bet just on the basis of prior information. In other previously untested tasks, they could do so only by integrating prior and posterior information.

2.1. Method

2.1.1. Participants

In all the reported studies, the participants were native Italian speakers from northern Italian public kindergartens and schools. Signed consent was obtained from parents. In each study, gender was balanced. Each child was tested individually, in a quiet room. In Study 1, the participants were 63 preschoolers (mean age = 4;5 years, range = 3;8–4;11), 38 kindergarteners (mean age = 5;5 years, range = 5;0–5;10), 38 first graders (mean age = 6;4 years, range = 5;10–6;10), 37 second graders (mean age = 7;11 years, range = 7;4–8;3), and 33 fourth graders (mean age = 9;10 years, range = 9;4–10;6).

2.1.2. Material, procedure and design

Each child received three tasks. In each task, the experimenter informed the children that they could win a chocolate by playing with a set of chips of two different colors and two puppets, each colored similarly to the chips. The experimenter named each puppet on the basis of its color, and explained that it owned the equally colored chips. For instance, if the chips were black and white, the black puppet was named "Mr. Black, the owner of the black chips", and the white one was named "Mr. White, the owner of the white chips". The experimenter checked whether the children understood the instructions by asking them to indicate one chip belonging to each puppet. If children failed, they were corrected.

The chips were put in an opaque bag. As a mnemonic aid about the contents of the bag, children were given a card (30 × 15 cm) depicting the chips. In each task, the experimenter drew a chip from the bag. The puppet owning the drawn chip won the chocolate. Children had to indicate which puppet they wanted to be in order to win the chocolate. In each task, the correct answer was to indicate the puppet who was more likely to own the selected chip.

In the *prior* task, children were presented with four triangular chips (2 cm per side) of yellow and blue color. Three chips had one color (e.g., yellow), one chip another color (e.g., blue). The experimenter said:

> I'm going to put all the chips in this bag, but you can remember how they are because they are copied on this card. I will shake the bag and I will take a chip from it without looking. If the chip is ... [the experimenter named one color], then Mr. ... [the experimenter named and indicated the puppet of that color] wins a chocolate. If the chip is ... [the experimenter named the other color], then Mr. ... [the experimenter named and indicated the puppet of that color] wins a chocolate. You have to choose which puppet you would like to be in order to win the chocolate. Show me which puppet you would like to be.

The odds were 3:1 in favor of the prevalent color. Hence, the correct choice was to indicate the puppet owning the three chips of that color.

In the *updating* task, children were presented with a set of eight chips, four square chips (2 cm per side) and four round chips (1.5 cm in diameter) of red and green color or else of black and white color. Five chips had one color (e.g., red), three chips another color (e.g., green). As in the prior task, the chips were put in a bag. Children had to make two successive choices. First, as in the prior task, they had to a make *a prior* choice, that is, they had to indicate the puppet that they would like to be in order to win the game, *before* the experimenter actually drew a chip from the bag. The odds were 5:3 in favor of the prevalent color. Hence, the correct choice was to indicate the puppet owning the five chips of that color. Second, children had to a make *a posterior* choice, that is, they had to opt again for one of the puppets, *after* the experimenter actually drew the chip from the bag and, without revealing it, informed them of its shape. The experimenter said:

> Ah, listen. I'm touching the chip that I have drawn and now I know something that might help you to win the game. I'm touching the chip that I have in my hand and I feel that it is ... [the experimenter named a shape: either round or square]. Now, show me which puppet you would like to be.

Unknown to the children, the experimenter touched the chips and chose a chip having the shape determined by the task version. In one version (*confirmation*), most chips (three out of four) having the selected shape had the same color as most chips (five out of eight) in the initial set. For instance, consider a case in which most chips were red (i.e., five red vs. three green), and most round chips were equally red (i.e., three red vs. one green). Informing children that the drawn shape was round, increased the odds in favor of the red color from 5:3 to 3:1. Hence, in both prior and posterior task, the correct answer was to select the same puppet, the one owning the greatest number of chips in the initial set and in the subset of chips having the selected shape. In another version (*disconfirmation*), the correct choice was not the same in the two tasks, because the prevalent color in the initial set of chips was not the same as the prevalent color in the subset of chips having the selected shape: only one of the latter had the initially predominant color. For instance, consider a case in which most chips were black (i.e., five black vs. three white), but most square chips were white (i.e., three white and one black; see Fig. 1). Informing children that the drawn shape was square, changed the odds from 5:3 in favor of the black color to 3:1 against it. Hence, in a disconfirmation version, in order to make optimal bets, children had to change their choice. They had to choose the puppet owning the greatest number of chips in the initial set. But, in the second task, they had to choose the other puppet.

The *posterior* task was identical to the updating one, except that children were not asked for a prior choice: they had to make only a posterior choice after the experimenter indicated the shape of the drawn chip. The chips were identical to those used in the updating task, except that they had different colors (e.g., if there were chips of red and green color in the updating task, there were chips of black and white color in the posterior one). Like the updating task, the posterior task occurred in a confirmation or a disconfirmation version.

For all children, the prior task was the initial one. Half of the children received a version of the task in which the prevalent color was yellow. For the other half, the opposite held. Presentation of the other two tasks was counterbalanced across children. In one condition, children received a confirmation version of the updating task, and a disconfirmation version of the posterior task. In another condition, the opposite held. The shape of the drawn chip (square vs. round) was counterbalanced across versions. Half of the children received the tasks with a given allotment of colors (e.g., in the disconfirmation version, five black chips and three white chips). For the other half, the opposite held (e.g., in the disconfirmation version, 3 black chips and 5 white chips).

2.2. Results and discussion

In each task, we scored as correct the choice of the more likely color. Table 1 reports the correct response rate in each task, collapsing data across versions. All but the youngest children performed better than chance in the prior task (in which correct performance by chance was .50): first graders, $\chi^2(1, n = 38) = 3.79$, $p = .05$; second graders, $\chi^2(1, n = 37) = 7.81$, $p = .005$; fourth graders, $\chi^2(1, n = 33) = 22.09$, $p = 3 \times 10^{-6}$; as well as in the posterior test: first graders, $\chi^2(1, n = 38) = 6.74$, $p = .009$; second graders, $\chi^2(1, n = 37) = 9.76$, $p = .002$; fourth graders, $\chi^2(1, n = 33) = 10.94$, $p = 9 \times 10^{-4}$. Performance in the prior task did not differ from performance in the posterior one (66% of correct bets in each task, collapsing data across age groups). If children had neglected posterior information, they would have performed better in the tasks in which one should not change one's bet after receiving the new piece of information. In fact, in the confirmation versions, in which the information about the chip's shape did not change the more likely color, posterior bets were slightly but not significantly more correct than in the disconfirmation versions (average correct rate: 69% and 65%, respectively, collapsing data across age groups).

As a general index of performance in the updating task, we considered the number of correct bets (0, 1, or 2; see Table 1). In all age groups, the mean number of correct bets out of two (m) was greater than 1, that is, the expected number of bets correctly produced by chance: preschoolers, $m = 1.24$, $t(62) = 3.07$, $p = .002$, one tailed; kindergarteners, $m = 1.18$, $t(37) = 1.56$, $p = .06$, one tailed; first graders, $m = 1.26$, $t(37) = 2.24$, $p = .02$, one tailed; second graders, $m = 1.32$, $t(36) = 2.52$, $p = .008$, one tailed; fourth graders, $m = 1.52$, $t(32) = 4.15$, $p = .0001$, one tailed. Evidence exists that questioning procedures may affect children's inferences. For instance, if

Table 1
Rate of correct decision in the three tasks, by age group (Study 1)

Group	Task				
			Updating[a]		
	Prior	Posterior	Prior	Posterior	Both
Preschool	.52	.51	.59	.65	.33
Kindergarten	.61	.63	.53	.66	.37
1st Grade	.66	.71	.63	.63	.42
2nd Grade	.73	.76	.59	.73	.51
4th Grade	.91	.79	.76	.76	.64

[a] The mean number of correct bets (0, 1, or 2) in the updating task (see the text) can be derived from *Prior, Posterior*, and *Both* entries. *Both* entry indicates the rate of two correct bets. The rate of only one correct bet equals the sum of the rates of only prior bet and only posterior bet correct, that is, $(Prior - Both) + (Posterior - Both)$. The rate of no correct bet equals the rate of prior incorrect minus the rate of only posterior correct, that is, $(1 - Prior) - (Posterior - Both)$.

they are asked the same question twice, 6-year-olds tend to change their initial answer, even in conditions in which they should not do so (e.g., Rose & Blank, 1974). Therefore, one might argue that the request to make two consecutive bets about the same set of objects could have affected children's performance in the updating task. The results, however, suggest that the first bet of the updating task did not affect the second one. Indeed, the latter turned out to be as correct as the bet choice in the posterior task (average correct rate: 68% and 66%, respectively, collapsing data across age groups). In conclusion, from the age of about five, children are able to integrate prior and individuating information and to update their bets.

3. Study 2: Updating judgments

Are children's intuitions of posterior probability limited to situations in which they have to make a bet? To answer this question, in Study 2 we presented the same situations as in Study 1, but children were asked to make a judgment closer to a proper probability evaluation. In other words, children were not asked to bet on a puppet in order to win a chocolate, but to indicate which puppet was more likely to win the game.

3.1. Method

The participants were 29 preschoolers (mean age = 4;2 years, range = 3;9–4;7), 32 kindergarteners (mean age = 5;8 years, range = 5;0–6;5), 33 second graders (mean age = 7;10 years, range = 7;4–8;5), and 31 fourth graders (mean age = 9;11 years, range = 9;5–10;4).

Children received an updating and a posterior task. The material and the procedure were the same as in Study 1, except that children did not receive any reward and had to answer a different sort of question. They did not have to indicate the puppet that they wanted to be in order to win a chocolate, but rather the puppet they judged more likely to win. In the posterior task and in the posterior judgment of the updating task, children had to answer the question: "According to you, who is the puppet favored by the chip I have drawn, Mr. X or Mr. Y, or neither of them?", where X and Y were the names of the colors involved in the version. In the prior judgment of the updating task, children had to answer the question: "According to you, who is the puppet favored by the chip I will draw, Mr. X or Mr. Y, or neither of them?"

3.2. Results and discussion

In each task, we scored as correct the indication of the puppet associated with the more likely color. Table 2 reports the correct response rate in the two tasks. As in Study 1, children did not neglect posterior information. In the disconfirmation versions (see Study 1), in which the information about the chip's shape changed the more likely color, there was even a marginally significant tendency to produce better posterior judgments than in the confirmation versions (average correct rate: 80% and

Table 2
Rate of correct judgment in the two tasks, by age group (Study 2)

Group	Task			
		Updating[a]		
	Posterior	Prior	Posterior	Both
Preschool	.48	.45	.59	.24
Kindergarten	.84	.66	.75	.53
2nd Grade	.76	.85	.76	.73
4th Grade	.97	.90	.90	.87

[a] The mean number of correct judgments (0, 1, or 2) in the updating task (see the text) can be derived from *Prior, Posterior*, and *Both* entries. *Both* entry indicates the rate of two correct judgments. The rate of only one correct judgment equals the sum of the rates of only prior judgment and only posterior judgment correct, that is, (*Prior* – *Both*) + (*Posterior* – *Both*). The rate of no correct judgment equals the rate of prior incorrect minus the rate of only posterior correct, that is, (1 – *Prior*) – (*Posterior* – *Both*).

72%, respectively, collapsing data across age groups), $\chi^2(1, N = 125) = 2.70$, $p = .10$. In the posterior task, all but the youngest children performed better than chance: kindergartners, $\chi^2(1, n = 32) = 15.13$, $p = 1 \times 10^{-4}$; second graders, $\chi^2(1, n = 33) = 8.76$, $p = .003$; fourth graders, $\chi^2(1, n = 31) = 27.13$, $p = 2 \times 10^{-7}$. As a general index of performance in the updating task, we considered the number of correct judgments (0, 1, or 2; see Table 2). In all but the youngest age group, the mean number of correct judgments out of two (m) was greater than 1, that is, the expected number of judgments correctly produced by chance: kindergartners, $m = 1.41$, $t(32) = 3.23$, $p = .001$, one tailed; second graders, $m = 1.61$, $t(32) = 4.94$, $p = 1 \times 10^{-5}$, one tailed; fourth graders, $m = 1.81$, $t(30) = 8.27$, $p = 2 \times 10^{-9}$, one tailed. Children's posterior judgment in the updating task did not significantly differ from their unique judgment in the posterior task (average correct rate: 75% and 77%, respectively, collapsing data across age groups). This result suggests that the first judgment of the updating task did not affect the second one.

4. Study 3: Updating decisions with respect to single events

According to the evolutionary-frequentist view, the human mind is able to reason about uncertain events only when problems activate a module for reasoning about frequencies of multiple, repeatable events (e.g., Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995). Contrary to this view, our results show that children do not need to observe a series of outcomes in order to reason correctly about a single event. In Study 1 and 2, however, the single event was the drawing of a chip from a box, that is a random, repeatable event. In Study 3, we tested our hypothesis asking children to reason about a single, not repeatable event, produced by an intentional agent, rather than about the random outcome of a chance device. No previous studies have investigated young children's ability to reason about this sort of uncertain, single event.

Children were presented with two boxes, each containing a group of animal-toys, and told that a troll had secretly placed a chocolate in the bag of one animal.

Children had to select one box. If it contained the animal carrying the chocolate, children were allowed to keep it. The task consisted in two phases. In the first phase, children chose one box, but prior evidence did not favor one box over the other. In the second phase, children received a new piece of information, evidencing that one box had twice as many possibilities as the other box of containing the chocolate. Given the new evidence, children could either stay with their initial choice, or else switch to the other box.

Following our hypothesis, children should pass this sort of task, because it requires them to consider and compare possibilities, even if these possibilities concern a single, not repeatable event. In particular, children who make an incorrect choice in the first phase should be able to switch from the initially chosen box to the one favored by posterior evidence. By contrast, following the frequentist-evolutionary hypothesis, children should fail this sort of task, given that it required them to reason about a single, not repeatable event.

4.1. Method

The participants were 33 preschoolers (mean age = 4;5 years, range = 3;9–4;l 1), 19 kindergarteners (mean age = 5;10 years, range = 5;5–6;5), 13 first graders (mean age = 6;4 years, range = 5;11–6;8), 66 second graders (mean age = 7;4 years, range = 6;10–7;l 1), 63 third graders (mean age = 8;7 years, range = 8;0–9;2), and 71 fifth graders (mean age = 10;6 years, range = 9;10–11;3).

Each child received three tasks. In each task, children were presented with two boxes (20 × 10 × 7 cm each) and told that one of them contained a chocolate. They had to make two successive choices in order to find the chocolate. The experimenter said:

> Some animals live in this house [the experimenter pointed at a box], some others in this house [the experimenter pointed at another box]. While the animals were outside, a troll put one chocolate in the bag of one of the animals. In the evening, the animals went home without checking what they had in their bags. Thus, now there is a chocolate in one house, but nobody knows in which one. Only the troll knows, and he has written it in a letter that he has put in this sealed envelope [the experimenter showed a sealed envelope]. If you find the animal that has the bag with the chocolate, you win the chocolate. You can choose one of the two houses in order to find the animal that has the bag with the chocolate. Which house do you choose? [Children indicated one box.]

Each task conveyed a different type of information. In one task, both boxes were sealed. The experimenter presented a card (20 × 10 cm) depicting six animals (three sheep and three goats), without informing children of the house in which each animal lived. After children made their first choice, a *base rate* information was provided. The experimenter said: “Before I open the troll’s letter and read the name of the animal with the chocolate, do you want to see the animals living in each of the two houses?” If children accepted, the experimenter removed the cover of each box and showed them the animal-toys it contained, that is, two goats and two sheep, versus one goat and one sheep. "Now, which house do you choose?" In the first phase, there was no correct choice, since no information about the content of the two boxes was provided. In the second phase, the new piece of information about the base rate favored the box with four animals over the one with two animals. Indeed, the troll chose to place the chocolate in the bag of one of the six animals, so that each animal had one chance in six of carrying the chocolate. Hence, the correct choice was to opt for the box with four animals, given that it had four possibilities in six of containing the chocolate.

In another task, posterior evidence conveyed information about a *category*. The procedure was the same as in the previous task, except that in the first phase children had to choose between two open boxes, each containing three animal-toys (two cats and one dog vs. two dogs and one cat). There was no correct choice, given that both boxes contained the same number of animals. Notice that the troll and his secret action were introduced in this scenario in order to avoid a potentially misleading cue. If one animal had deliberately taken the chocolate, children could make their choices judging, for instance, that dogs like chocolate more than cats. Because the troll had placed the chocolate in the bag of an animal, and had done so secretly, it was difficult to associate the chocolate with a specific kind of animal. In the second phase, after children chose one box, the experimenter silently read the troll's letter and told them: "In this letter, the troll has written the kind of animal carrying the chocolate. Do you want to know which kind of animal this is?" If they accepted, half of the children were told that a cat (the other half, a dog) carried the chocolate, and were then asked to make a new choice. The new piece of information favored the box with more cats (or dogs) over the other one. Hence, the correct choice was to opt for the box with two cats (or two dogs), given that it had twice as many possibilities of containing the animal carrying the chocolate than the box with only one cat (or dog).

In still another task, posterior evidence conveyed information about a *sub-category*. The procedure was the same as in the previous tasks, except that in the first phase children saw two open boxes, each containing six animal-toys (two white bears, one black bear and three rabbits vs. two black bears, one white bear and three rabbits). Moreover, before asking children for a first choice, the experimenter silently read the troll's letter, and informed them that a bear carried the chocolate. There was no correct choice, given that both boxes contained the same number of bears. In the second phase, the experimenter read again the troll's letter and said: "In this letter, the troll has also written which kind of bear carries the chocolate. Do you want to know which kind of bear this is?" If they accepted, half of the children were told that a white (the other half, a black) bear carried the chocolate, and were then asked to make a new choice. The new piece of information favored the box with more white bears (or black bears) over the other one. Hence, the correct choice was to opt for the box with two white bears (or two black bears), given that it had twice as many possibilities of containing the bear carrying the chocolate than the box with only one white bear (or one black bear).

The task order and the position (right or left) of the winning box (which was always the one favored by evidence) were balanced across participants.

4.2. Results and discussion

We scored as correct the choice of the box favored by posterior evidence. In the initial phase, no information favored one box over the other one, and the irrelevant features of the situation (e.g., the kind of animal carrying the chocolate and the position of the box containing it) were balanced across participants. Hence, the expected rate of a prior correct choice was 1/2. In the three tasks, the rate of prior correct choice did not differ significantly from 1/2 (47% in the base-rate information task, 54% in the category information task, and 49% in the sub-category information task, collapsing data across age groups). If children did not use the additional piece of information, their final performance would not differ from the initial one. But, if they did, their final choice would outperform the initial one. Consequently, difference in performance between the initial and the final choices, that is, the proportion of correct posterior choices minus the proportion of correct prior choices, indicates the *gain*, that is, the mean difference in performance in each task, produced by posterior evidence. The rate of correct final choices and the gain were calculated considering only the tasks in which children accepted to receive the new piece of evidence. In 96% of the cases, children accepted it. There was no noticeable effect of task type. The average gain (collapsing data across age groups) was 17% in the base-rate information task, 15% in the category information task, and 14% in the sub-category information task. Given that the gain did not differ significantly in the three tasks, this variable was not considered in the following analysis. Table 3 reports the rate of correct final choices and the gain (collapsing data across tasks), by age group. The sample of kindergartners and first graders was small in size. Consequently, we analyzed their answers as a group.

The gain was only marginally significant for the youngest children, $t(32) = 1.42$, $p = .08$, one-tailed, but it was greatly significant in all the other age groups: kindergartners and first graders, $t(32) = 2.55$, $p = .008$, one-tailed; second graders, $t(64) = 3.55$, $p = 2 \times 10^{-4}$, one-tailed; third graders, $t(63) = 2.51$, $p = .007$, one-tailed; fifth graders, $t(69) = 4.01$, $p = 6 \times 10^{-5}$, one-tailed. In conclusion, from the age of five, children are able to update their choices about single events, in the light of new evidence.

Table 3
Mean rate of correct final choice and gain in performance from the initial to the final choice, by age group (Study 3)

Group	Final choice	Gain
Preschool	.62	.11
Kindergarten-1st Grade	.65	.17
2nd Grade	.61	.17
3rd Grade	.65	.12
4th Grade	.72	.19

5. General discussion

Our results show that, when faced with uncertain events, children are able to integrate a new piece of information into prior information. From the age of about five, they are able to use posterior information and to update their bets (Study 1) and their judgments (Study 2) about random outcomes. They are able to do so even when they face a single event, produced by an intentional agent (Study 3). The moral of these results is that preschool children correctly revise their decisions and judgments, on the basis of inferences that have been considered on the outer edge of human competence. And they do so in the same situations in which adults perform correctly, that is, when they can associate specific information with a subset of the possibilities evoked by prior information.

Skeptics might argue that our experimental settings are too elementary to grant general conclusions and appropriate comparisons with adults' performance. In point of fact, our settings are closer to an ideal situation in which individuals update their decisions and judgments than those typically used in adult literature. In the latter case, participants have to make a *single* probability judgment, rather than two successive judgments, the first one based on prior and the second one on posterior evidence. Typical tasks ask participants to evaluate the probability of a given hypothesis (e.g., one individual has a given disease) on the basis of a text conveying various pieces of evidence (e.g., prevalence of the disease, positive result of the individual in a screening test, reliability of the test). In such tasks, participants did not have to revise any initial evaluation. They had only to combine the values provided by the pieces of evidence. By contrast, in our case, children had to provide a prior answer and then a subsequent one, in the light of posterior information, that is, they were actually required to update their answers. Therefore, their correct decisions and judgments can be used to support the hypothesis that naive individuals possess an intuition of the basic principles that link prior and posterior probability.

What could be the source of this intuition? The finding that such intuition emerges around the age of five or six may suggest that schooling shapes it. In fact, most primary schools (including those in which we have conducted our studies) do not train children in probabilistic reasoning, at least during the first years of primary school. Yet, the present results show that even kindergartners succeed on posterior probability tasks. Primary schools provide formal arithmetic instruction. However, the arithmetic skills that children need to solve our probabilistic problems do not depend on formal instruction. Basic numerical abilities, like comparing and adding quantities, are acquired before (e.g., Barth, Le Mont, Lipton, & Spelke, 2005) and regardless of schooling (e.g., Pica, Lerner, Izard, & Dehaene, 2004). Thus, the available evidence does not favor the hypothesis that school facilitates the ability to take into account posterior probability.

The present results extend those obtained in studies in which children had to use only prior information (e.g., Brainerd, 1981), and complement those obtained in studies in which children had to consider the result of a combination of possibilities (Girotto & Gonzalez, 2005). Taken together, they corroborate the extensional reasoning hypothesis (Johnson-Laird et al., 1999): children, like adults, base their decisions and judgments under uncertainty on an extensional evaluation of possibilities, considering and comparing the various ways in which an outcome may or may not occur.

5.1. Children's posterior evaluations of single events

The findings that preschool children draw correct posterior inferences is of significance to the current debate on whether the human mind can deal with single-event probability. According to an influential view, the human mind is blind to single-case probability, but tuned to frequencies of repeatable events: "Humans seem to be developmentally and evolutionarily prepared to handle natural frequencies" (Gigerenzer & Hoffrage, 1999, p. 430). The results of a recent study appear to corroborate the evolutionary-frequentist hypothesis. Zhu and Gigerenzer (2006) found that 11-year-olds solve problems in which they have to reason about frequencies. By contrast, they fail on problems in which they have to evaluate the probability of a single event. The trouble with these results is that they have been obtained with verbal problems that convey probability information by means of numerical symbols, and ask for a numerical judgment, that is, with verbal problems that even undergraduate students fail. Given that verbal problems are likely to underestimate children's thinking abilities (e.g., Mehler & Bever, 1967), one should rather investigate children's probabilistic competence using non-verbal tasks that ask for a decision, a comparison, or a non-numerical judgment. Indeed, by using problems of this sort, we proved the existence of an early intuition of chance. In particular, contrary to the evolutionary-frequentist hypothesis, Study 3 showed that even 5-year-olds are able to base their responses on the posterior chance of a single event.

Besides methodological concerns, there is a theoretical reason to use non-verbal tasks in investigating children's reasoning about uncertain events. According to the evolutionary argument, in the environment in which they evolved, our ancestors encountered frequencies of actual events, so that natural selection has shaped the human mind to extract frequencies from series of observations (e.g., Cosmides & Tooby, 1996). Therefore, if one wants to seriously test the hypothesis of an early intuition of frequencies, one should ask children to make judgments in problems closer to the natural settings in which sequential information is acquired. In other words, children should face situations requiring them to extract frequencies from *actual* observations, rather than problems in which they have to process numerical symbols presented in a verbal text. Rather surprisingly, the advocates of the evolutionary-frequentist hypothesis have never used the former sort of problem in their investigations. But there is a study in which children had to make judgments on the basis of information acquired through a sequence of actual observations. As indicated above, Brainerd's (1981) preschool children had to predict the result of a random draw from a set of chips. They performed very well in the initial trial, that is, in the trial in which they had to make a prediction only on the basis of the initial set of possibilities determined by prior information. But these children were also asked to predict the results of'the *following* random draws from the same set. In other words, they were asked to make predictions on the basis of the encountered frequencies of actual drawings. Far from being accurate, their later responses were based on erroneous strategies like "choose the kind of chip that was not chosen on the preceding trial". This finding has been neglected in the current debate on naive probabilistic reasoning. But it shows that preschool children make erroneous predictions in situations in which they make actual observations of events.

One might argue that reasoning about prior possibilities before experiencing actual outcomes does not actually differ from reasoning about empirical frequencies, given that, in both cases, individuals compare sets of countable elements. For instance, consider the chips presented in Fig. 1. Suppose you put them in a bag. If one takes a chip at random, you can infer that it is more likely to be black because there are five black and three white chips in the bag. In this case, you base your reasoning on prior possibilities. Consider another situation, in which you do not know the content of the bag and you can only observe one person taking a chip at random and replacing it in the bag eight times: the sample of draws shows five black and three white chips. On the basis of this frequency, you can infer that, if one takes again one chip at random, it is more likely to be black because black chips were more frequent in the sample. In sum, in both situations you draw an inference by simply considering and comparing sets of countable elements (either prior possibilities or actual frequencies). Both inferences are based on a similar comparison: the set of black chips is larger than the set of white chips, and the set of draws of a black chip is larger than the set of draws of a white chip. However, the two inferences cannot be assimilated. On the normative side, the former yields an exact probability, based on the assumption that all of the prior possibilities are equiprobable. The latter is a probability estimate, based on the assumption that each event, past and future, obeys the same probability law. More importantly, the two inferences ask for a different cognitive treatment. In the former case, you have to reason about a set of prior possibilities *before* making any actual experience. In the latter case, you have to reason about a sample of actual observations.

In conclusion, our results show that children can solve problems that do *not* convey frequency information and do *not* ask for a frequency prediction. The finding that children form correct expectations about uncertain events before experiencing their actual frequency contradicts the evolutionary-frequentist view that the human mind is shaped to deal only with frequencies.

5.2. Normative vs. intuitive probabilistic reasoning

The finding that young children possess a basic intuition of probability contradicts the idea that probabilistic reasoning is not intuitive. Several authors, including those who do not adhere to the evolutionary-frequentist view (e.g., Nisbett, Krantz, Jepson, & Kunda, 1983; Piatelli-Palmarini, 1994), defend this idea, which appears to be sustained by two main arguments. On the one hand, most adults fail on simple probabilistic problems. On the other hand, the probability calculus (i.e., a normative system for measuring chance) emerged only in the 17th century. We consider these arguments in turn.

5.2.1. The limits of the extensional evaluation of chance

Skeptics of the extensional reasoning hypothesis might point out that children are unlikely to possess correct probabilistic intuitions given that intelligent adults are unable to solve very elementary probabilistic problems. One problem with this criticism is that it equals *elementary* intuitions and flawless reasoning. But children's extensional intuition about probability does not imply that they or even adults will always draw correct probabilistic inferences. Similarly, young children's arithmetic intuition (e.g., Barth et al., 2005) does not imply that adults' arithmetic is flawless. Now, adults err even in problems that in principle could be solved by means of an extensional treatment of very few possibilities, like the notorious Monty Hall problem (for a review, see Nickerson, 1999). This finding, however, does not contravene the extensional competence hypothesis. In fact, the hypothesis assumes that the extensional assessment of chance is constrained by the limited capacity of working memory and the difficulty of working out all relevant possibilities (Johnson-Laird et al., 1999). Accordingly, the hypothesis has predicted and shown that individuals produce illusory probabilistic evaluations in problems that evoke an incomplete representation of possibilities (Girotto & Gonzalez, 2005; Johnson-Laird et al., 1999; Johnson-Laird & Savary, 1996). In sum, extensional procedures do not always yield accurate probability evaluations. Inferential errors, however, indicate specific difficulties in constructing a correct representation of the suitable possibilities, rather than a generic inability to reason about possibilities.

5.2.2. The late emergency of probabilistic norms

The recent invention of the probability calculus has been used to maintain the thesis that probabilistic reasoning is not intuitive (e.g., Gigerenzer & Hoffrage, 1995). There are two assumptions at the basis of this argument. First, there is something special in the late emergence of the theory of probability. Second, before the seventeenth century there was no intuitive evaluation of chance. Both assumptions are questionable. On the one hand, basic mathematical tools (e.g., decimal notation, algebraic symbolism, combinatorics) and shared mathematical knowledge were lacking before the seventeenth century. These factors are likely to have hindered the emergence of a theory of probability, as well as the emergence of other fields of mathematics (Franklin, 2001). On the other hand, individuals living in pre-modern times had some notions of aleatory probability, as proved by their conceptualization of chance devices such as the throwing of coins or dice (e.g., Franklin, 2001; Girotto & Gonzalez, 2006). They had some notions of logical probability too, as proved by medieval jurists' ability to distinguish conclusive from inconclusive evidence (Franklin, 2001). Such an ability contradicts the popular idea that before the Renaissance individuals lacked even the basic concept of evidence, that is the concept that one thing can indicate, contingently, the state of another thing (Hacking, 1975). This idea is also in conflict with the reported finding that even young children have some notions of evidential support, given that they update their judgments and decisions in the light of posterior evidence.

In sum, the evidence of an intuition of probability in children and in individuals living before the advent of the theory of probability contradicts the thesis that the acquisition of the probability calculus is the only way to reason about probabilities (pace Gigerenzer & Hoffrage, 1995; Nisbett et al., 1983). More generally, this evidence supports Locke's demonstration that intuitive reasoning does not depend on normative systems of reasoning:

> He that will look into many parts of Asia and America, will find men reason there perhaps as acutely as himself, who yet never heard of a syllogism, nor can reduce any one argument to those forms [...] Syllogism [is] not the great instrument of reason [...] if syllogisms must be taken for the only proper instrument and means of knowledge; it will follow, that before Aristotle there was no man that did or could know anything by reason; and that since the invention of syllogisms there is not one often thousand that doth. But God has not been so sparing to Men to make them barely two-legged creatures, and left to Aristotle to make them rational (Locke, 1690/1975, p. 671).

To paraphrase Locke, just as Aristotle did not make human beings rational by providing them with a normative system of deductive reasoning, Pascal and Fermat did not make human beings rational by providing them with a normative system of probabilistic reasoning.

6. Conclusions

Posterior probability evaluation is sometimes very difficult, even for adults. In problems that evoke an heterogeneous representation of the various pieces of evidence or in problems whose solution requires complex numerical computation, only individuals who are familiar with the probability calculus appear to combine new evidence and prior information in a normative way. The present results, however, license a more optimistic conclusion: in problems whose solution depends on a simple enumeration of possibilities, even young children are able to integrate new information into prior information, by making accurate judgments and optimal decisions.

Acknowledgement

We would like to thank Ira Noveck, Jennifer Knaeble and two anonymous referees for their comments on a previous version of this paper, Claudia Bin, Ilaria Camozzo, Lara Pavan and Stefania Pighin, for their help in collecting data. This research was supported by a COFIN (2005117840_003) grant from the Italian Ministry of Universities. Portions of this research were presented at The Fifth International Conference on Thinking, University of Leuven, July 22–24, 2004.

References

Acredolo, C., O'Connor, J., Banks, L., & Horobin, K. (1989). Children's ability to make probability estimates: Skills revealed through application of Anderson's functional measurement methodology. *Child Development, 60*, 933–945.

Barth, H., Le Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Science, 102*, 14116–14121.

Beck, S. R., & Robinson, E. J. (2001). Children's ability to make tentative interpretations of ambiguous messages. *Journal of Experimental Child Psychology, 79*, 95–114.

Brainerd, C. J. (1981). Working memory and the developmental analysis of probability judgment. *Psychological Review, 88*, 463–502.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58*, 1–73.

Fox, C. R., & Levav, J. (2004). Partition-edit-count: naive extensional reasoning in judgment of conditional probability. *Journal of Experimental Psychology: General, 133*, 626–642.

Franklin, J. (2001). *The science of conjecture. Evidence and probability before Pascal*. Baltimore: John Hopkins University Press.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency format. *Psychological Review, 102*, 684–704.

Gigerenzer, G., & Hoffrage, U. (1999). Overcoming difficulties in Bayesian reasoning: a reply to Lewis and Keren (1999) and Mellers and McGraw (1999). *Psychological Review, 106*, 425–430.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases. The psychology of intuitive judgment*. Cambridge: Cambridge University Press.

Girotto, V., & Gonzalez, M. (2001). Solving probabilistic and statistical problems: a matter of information structure and question form. *Cognition, 78*, 247–276.

Girotto, V., & Gonzalez, M. (2002). Chances and frequencies in probabilistic reasoning: rejoinder to Hoffrage, Gigerenzer, Krauss, and Martignon. *Cognition, 84*, 353–359.

Girotto, V., & Gonzalez, M. (2005). Probabilistic reasoning and combinatorial analysis. In V. Girotto & P. N. Johnson-Laird (Eds.), *The shape of reason* (pp. 161–175). New York: Psychology Press.

Girotto, V., & Gonzalez, M. (2006). Norms and intuitions about chance. In L. Smith & J. Voneche (Eds.), *Norms in human development* (pp. 220–236). Cambridge: Cambridge University Press.

Hacking, I. (1975). *The emergence of probability*. Cambridge: Cambridge University Press.

Hammerton, M. (1973). A case of radical probability estimation. *Journal of Experimental Psychology, 101*, 252–254.

Johnson-Laird, P. N. (2006). *How we reason*. Oxford: Oxford University Press.

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Sonino-Legrenzi, M., & Caverni, J. P. (1999). Naive probability: a model theory of extensional reasoning. *Psychological Review, 106*, 62–88.

Johnson-Laird, P. N., & Savary, F. (1996). Illusory inferences about probabilities. *Acta Psychologica, 93*, 69–90.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*, 237–251.

Locke, J. (1975). *An essay concerning human understanding*. Oxford: Clarendon Press (Original work published 1690).

Mehler, J., & Bever, T. (1967). Cognitive capacities of very young children. *Science, 158*, 141–142.

Nickerson, R. S. (1999). Ambiguities and unstated assumptions in probabilistic reasoning. *Psychological Bulletin, 120*, 410–430.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, D. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review, 90*, 339–363.

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

Piatelli-Palmarini, M. (1994). *Inevitable illusions: How mistakes of reason rule our minds*. New York: Wiley.

Pica, P., Lerner, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306*, 499–503.

Rose, S. A., & Blank, M. (1974). The potency of context in children's cognition: an illustration through conservation. *Child Development, 45*, 499–502.

Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes, 91*, 296–309.

Yost, P. A., Siegel, A. E., & Andrews, J. M. (1962). Nonverbal probability judgments by young children. *Child Development, 33*, 769–780.

Zhu, L., & Gigerenzer, G. (2006). Children can solve Bayesian problems: the role of representation in mental computation. *Cognition, 98*, 287–308.