CHILDREN'S PROBABILITY INTUITIONS: UNDERSTANDING THE EXPECTED VALUE OF COMPLEX GAMBLERS

Anne Schlottmann


ABSTRACT

Two experiments used Information Integration Theory to study how children judge expected value of complex gambles in which alternative outcomes have different prizes. Six-year-olds, 9-year-olds and adults saw chance games that involved shaking a marble in a bicolored tube. One prize was won if the marble stopped on blue, another if it stopped on yellow. Children judged how happy a puppet playing the game would be, with the prizes and probability of the blue and yellow outcomes varied factorially.

Three main results appeared in both studies: First, all ages used the normatively prescribed multiplication rule for integrating probability and value of each individual outcome -- a striking finding because multiplicative reasoning does not usually appear before 8 years in other domains. Second, all ages based judgment of overall expected value meaningfully on both alternative outcomes, but there were individual differences -- many participants deviated from the normative addition rule, showing risk seeking and risk averse patterns of judgment similar to the risk attitudes often found with adults. Third, even the youngest children took probability to be an abstract rather than physical property of the game. Overall, the present results demonstrate functional understanding of probability and expected value in children as young as 5 or 6, in contrast to the traditional view. These results contribute to the growing evidence on children's intuitive reasoning competence. This intuition can, on the one hand, support surprisingly precocious performance in young children, but it may also contribute to the biases still evident in adults' judgment and decision.
INTRODUCTION

Even very young children know that people act to fulfill their desires (e.g., Wellman, 1990). But people do not act blindly to get what they want, their action also depends on how likely the desired outcome is. To understand how humans evaluate their options for action, both outcome value and probability of goal attainment must be considered. These two factors are conjoined in the notion of expected value (EV), a concept basic to understanding goal-directed action in an uncertain world. It is reflected in the expectancy x valence models of human motivation (Tolman, 1932; Lewin, Dembo, Festinger & Sears, 1944; for review see Feather, 1982), as well as in normative approaches to human judgment and decision which hold that rational choice maximizes expected value (Edwards, 1954; von Neumann & Morgenstern, 1944; reviewed in Wright, 1984). The focus here is on an intuitive rather than formal EV concept. This is used in everyday reasoning, by adults as well as young children.

Expected value -- and judgment/decision more generally -- are relatively neglected in developmental research. However, detailed study in this area is vital to better understand and more effectively change adolescents’ risk taking behaviors, such as drug use or sexual activity (Fischhoff, 1992; Loewenstein & Fuerstenberg, 1991). Furby & Beyth-Marom (1992), for instance, pointed to children and adolescents’ seemingly more risk seeking behavior than adults’, and the need to know whether this appears because children differ from adults in how they see the desirability or probability of their options, or because they do not combine these features into a notion of expected value in a mature way. The Information Integration approach employed here allows separation of these aspects (Anderson, 1981, 1982, 1991, 1996). The two experiments reported in this paper focus on the latter issue -- the combinatorial structure of children's EV understanding.

The Development of Relational Reasoning

Emphasizing the structure of children's EV concept links the topic to an issue of broad concern in cognitive development: the growth of children's ability to reason about multiple variables and their interrelationships. Understanding of multidimensional concepts depends on this ability -- expected value, for instance, involves a compensatory, multiplicative relation between outcome probability and value; thus two options may have equal EV, one with high probability and small intrinsic value, the other with larger value, but lower probability. In the traditional view (e.g., Inhelder & Piaget, 1958), the ability to consider such relationships is late emerging. Preschoolers oscillate between undifferentiated judgments and centration on a single dimension, without ability to relate parts and whole systematically. This is taken to require computational and logical strategies that emerge only in adolescence with the acquisition of operational structures. Information Processing approaches essentially agree with the Piagetian emphasis on late developing computational strategies, but appeal to processing factors, such as limits on working memory (e.g., Brainerd, 1981), to account for younger children's difficulties with computation (see Kerkman & Wright, 1988; Siegler, 1998).

An increasingly popular alternative position emphasizes children's intuitive understanding over computational accuracy. Traditionally, childhood intuition is seen as an undifferentiated form of thought, inferior to and gradually replaced by logic and computation. Departures from this view are motivated by the recognition that intuitive, informal reasoning is as frequent in adults as in children, can be highly structured, and is not necessarily inferior to computation (see discussion in Anderson, 1996; Reyna & Brainerd, 1994). Moore and her colleagues, in particular, have explored differences between intuitive and computational reasoning in teens and adults, often showing advanced reasoning in intuitive tasks (Ahl, Moore & Dixon, 1992; Dixon & Moore, 1996; Moore, Dixon, & Haines, 1991; Surber & Haines, 1987). Information Integration studies have shown that even preschoolers relate multiple dimensions in intuitive judgment -- thus centration is not a conceptual but a performance limitation (see overview in Anderson, 1996). Fuzzy Trace Theory, a general account of the developmental relations between memory and reasoning, holds quite generally that intuitive reasoning is preferred to computation at all ages (e.g., Brainerd & Reyna, 1993; Reyna & Brainerd, 1994, 1995).

These contrasting approaches to reasoning can be illustrated with recent changes in views on probability understanding (see reviews in Anderson, 1991; Reyna & Brainerd, 1994; Wilkening & Anderson, 1991). Traditionally, probability is seen as a late acquisition (e.g., Piaget & Inhelder, 1975). Children have well-documented difficulties with the computation of frequency ratios, used to define probability formally. Without this ability, they are denied an appreciation for the relation between different possibilities and the totality of events. Consistent with this view, children use simplified strategies in standard probability choice tasks, e.g., they center on target outcomes but ignore alternative possibilities (e.g., Hoemann & Ross, 1971; Siegler, 1981; Falk & Wilkening, 1998).

Fischbein (1975; also see Yost, Siegel & Andrews, 1962), on the other hand, argued that intuitions of chance and probability are basic adaptive tools for organisms living in an uncertain world and that such intuitions appear early. This was consistent with findings that children solve some
probability tasks without computation and that they are sometimes sensitive to relative frequencies. What was not clear from such data was whether children had an intuitive probability understanding, or whether they were using nonprobabilistic strategies (see review in Hoorman & Ross, 1982).

The intuitive position gained strength when it became clear more recently that children’s intuitions have similar structure to the formal probability concept: When the task does not require children to calculate but merely to estimate even 4-year-olds demonstrate some understanding of the relation between target and alternative outcomes (Acredolo, O’Connor, Banks & Horobin, 1989; Anderson & Schollmann, 1991; Anderson & Wilkening, 1991). In a task that did not even require estimation, Huber & Huber (1987) could also show that 4-year-olds honed six axioms of subjective probability, and Kuzmak & Gelman (1986) found that they distinguished random from determinate phenomena. Clearly, even preschoolers’ intuitions of uncertainty are differentiated in functionally appropriate ways. This makes it difficult to discount them as nonprobabilistic.

If young children have intuitions of probability, they may also have an idea of expected value. Indeed, rather than being more complicated than probability alone, EV may be more natural to children. In everyday life probabilities typically appear in the context of goal attainment -- children's actions, like adults', involve risk and uncertainty. Even laboratory probability judgments are often made meaningful to children by assigning events the values of winning or losing a game, thus treating probability implicitly as a special case of expected value. Studies with explicit focus on EV (Anderson, 1980; Hommers, 1980) also initially considered how children evaluate chance games with prizes on one outcome and none on the other. But not all uncertain situations are of the all-or-nothing, win-or-lose variety. In the more general -- and more complex -- case, alternatives carry different non-zero outcome values. To assess the limits of young children's EV reasoning, the two experiments reported here consider this complex situation.

**Expected Value Model (Part 1): Multiplication of Probability and Value Within a Single Outcome**

The normative expected value model has two parts. First, for a single outcome, EV is defined as the intrinsic value of the event, \( v \), weighted in proportion to the probability of it occurring, \( p \):

\[
EV = pv. \tag{1}
\]

The multiplicative relation reflects that changes in outcome probability have more impact on EV for more valuable outcomes. Halving the chances of winning a prize halves expected value; this change is £ 5 when the prize is £ 10, but £ 50 when the prize is £ 100.

Adults’ expected value judgments obey the multiplying rule of Equation 1, with objective probabilities and values replaced by subjective estimates (for review see Shanteau, 1975). In children, intuitions of expected value can be studied by having them judge how happy they or a puppet would be to play chance games, e.g., involving roulette-type spinners or blind draws of a marble, in which a prize is won if the right color comes up. In studies using this approach, even the youngest children, 4 and 5 years old, typically do not centerate, but know that both value and probability of the winning outcome matter (Anderson, 1980; Hommers, 1980; Schollmann & Anderson, 1994).

Of major interest is how young children integrate these dimensions: As discussed above, they are traditionally denied the ability to relate two dimensions systematically. Computational multiplication strategies are usually not found before adolescence (e.g., Inhelder & Piaget, 1958; Siegler, 1981). Intuitive multiplication appears earlier, but still not below 8 or so. Younger children tend to use simpler addition rules instead, perhaps because they understand that two variables are relevant, but not how they are related (e.g., Anderson & Cuneo, 1978; Wilkening, 1981; Wolf & Algern, 1987).

From the limited evidence available it appears that expected value intuitions may be an exception to this. Anderson (1980) found multiplication already among 5-year-olds. Schollmann & Anderson (1994, session 2) confirmed this in one of two tasks, but their other task showed clear signs of multiplication only by age 8. Hommers (1980) also did not find multiplication in children up to 13 years, although this null result may mainly reflect low power single subject statistics. On balance, it appears that young children may already use the normative EV rule, indicating some understanding that probability acts as a modifier of value. Nevertheless, more evidence is desirable.

**Expected Value Model (Part 2): Additive Integration Across Multiple Outcomes**

The expected value model has a second part, relevant to situations in which more than one non-zero outcome is possible. In this case, overall expected value is the sum of the component expected values. With two alternatives of value \( v_1 \) and \( v_2 \):

\[
EV = pv_1 + (1-p)v_2. \tag{2}
\]
The complex expected value model thus includes additive across-outcome integration together with multiplicative within-outcome integration. Although in this case addition is the normative rule, and although addition is generally considered simpler than multiplication, adults do not obey this model exactly: Judgments reveal persistent nonadditivities of unclear origin, with the value of one option apparently influenced by the values of the other options (for review see Shanteau, 1975).

One study with children interspersed standard 1-prize games with games in which alternative outcomes both carried prizes (Schlottmann & Anderson, 1994, session 2). Surprisingly, all ages made similar judgments, with little sign that the task exceeded younger children’s abilities. Instead of simplifying the judgment by ignoring the lower value or lower probability outcome, even 5-year-olds distinguished “win-or-lose” 1-prize-games from 2-prize games, and riskless 2-prize-games with the same prize for both alternatives from risky games with two different prizes.

Unfortunately, this study did not allow determination of children's strategy. Several options were considered. First, young children may use a serial outcome-by-outcome addition strategy: They first compare values, then assess the higher value option using an additive rule, then make an equally additive adjustment for the lower value option. This approach is parsimonious because it requires children to be capable of nothing but additive integration for both within- and across-outcome integration. However, data patterns at all ages were equally consistent with use of a multiplicative EV strategy for the within-outcome integration. To complicate matters further, the data suggested some small deviation from normative across-outcome additivity – as with adults – inconsistent with both the EV and addition-only strategy. In the present experiments these possibilities can be distinguished.

The Present Study

The two experiments reported here aim to determine children's strategy for evaluating complex gambles with alternative prizes for alternative outcomes. To find the winning outcome, a marble is shaken in a clear tube inset with a bicolored strip; probability is manipulated by varying the number of equal sized segments of each color (see Figure 1). In this tube game, probability is not confounded with the absolute area for winning. Instead, a more sophisticated judgment of the relative proportion of each color is required. As discussed above, even young children can make intuitive judgments of this kind.

Each child evaluated multiple gambles with different outcome probabilities and values. Either a very small or very large prize could be won on one outcome (1 or 10 crayons on yellow), while the other outcome carried intermediate prizes (3 or 6 crayons on blue). Children judged how happy a puppet would be to play the game, this judgment taken as a measure of expected value.

The normative model for the present design is shown in Figure 2a: The bottom pair of curves is for games in which the blue outcome carries a larger prize than yellow. EV increases with the probability of the blue outcome and the curves have positive slope. The top pair of curves is for games with the same blue outcome, but now yellow carries a larger prize than blue. EV decreases with the probability of the blue outcome and the curves have negative slope. The curve reversal thus reflects that EV always increases with the probability of the larger prize outcome.
Despite their opposite slopes, however, both the top and bottom pair of curves diverge towards the right, and these fan patterns reflect the multiplicative within-outcome integration of Equation 1. That the top and bottom fans show the same amount of divergence, in contrast, reflects the additive across-outcome integration of Equation 2.

If the within-outcome integration is not multiplicative, then the curves do not fan. As illustrated in Figure 2b, an addition-only strategy predicts no statistical or visual interaction between probability and value of the blue outcome. In this model, the across-outcome integration is additive as well, so the separation in the top pair of curves mirrors that of the bottom pair. Across-outcome additivity thus appears both in Figures 2a and 2b.

If the across-outcome integration is not additive, then the top and bottom pairs of curves have different amounts of separation. In Figure 2c, for instance, the larger prize has more effect on the judgment than the smaller prize. Thus the bottom curves for games differing in the larger prize lie further apart than the top curves for games sharing the same large prize; if children ignored the smaller prize completely, the top curves would be identical. Because the larger prize is the risky option while the smaller prize is a sure thing, such a pattern may be called “risk seeking”, a term borrowed from the adult literature (e.g., Bromly & Curley, 1992; Lopes, 1987; see discussion below). Different types of deviation from across-outcome additivity are possible, of course. All such configurations, however, produce statistical and visual interactions involving outcome values.

Different strategies thus generate different curve patterns, and this allows diagnosis of how children understand the functional relations between probabilities and values without requiring numerically exact calculations; subjective estimates suffice. (For details on Functional Measurement, see Anderson, 1981, 1982, 1991, 1996). The normative model in Figure 2a was calculated with objective values. With different subjective values, the curve pairs change position or slope, evident in the ANOVA main effects. But the overall pattern remains, evident in the ANOVA interactions: Under the EV model, the curve pairs reverse direction, the two curves in each pair fan and they have identical separation. The integration analysis focuses on these interaction patterns in the means and statistics.

Standard developmental theories predict that with age children’s judgment should conform increasingly to the normative EV model in Figure 2a. From results on intuitive judgment in other domains, we might expect multiplication -- required for the within-outcome integration -- by 8 years or so, but not in younger children. On the other hand, additive reasoning -- required for the across outcome integration -- is often found even with preschoolers. Nevertheless, the limited evidence on EV suggests the opposite: early multiplication, but some deviations from additivity even in mature reasoning.

If the present study obtains further evidence of multiplicative EV judgments even in the 5- and 6-year-olds tested here, this would stand in contrast not only to the traditional position that preadolescents lack the computational tools to reason about uncertainty, but also to the more recent evidence on intuitive multiplication. Such a result would argue against an entirely domain-general account of the development of intuitive reasoning operations. It would also strengthen the view that the evaluation of uncertain goals is developmentally basic.

If the present study finds nonadditivities in children's EV judgments, this would contrast with their additive performance in other intuitive reasoning tasks, but it would suggest some continuity between children and adults in the judgment/decision area. Nonadditivities, such as risk seeking or risk aversion, may appear as group biases or individual differences, and in adult performance often take the form of individual differences (e.g., Bromly & Curley, 1992; Lopes, 1987). Developmental evidence on this is scarce, however, in part perhaps due to the use of relatively simple tasks with children. The more complex EV situation of the present study affords increased opportunity for risk seeking or other deviations from the normative model. To consider the extent of individual differences in this, individual children's performance will be assessed in addition to group performance.
EXPERIMENT 1

Method

Children met a puppet, Lucy Lemur, who likes to gamble for crayon prizes. Lucy shook a marble in a clear plastic tube inset with a bicolored paper strip. If the marble landed on yellow she would win one prize; if it landed on blue another. Probability of landing and value of the prize on each color were varied factorially. Children assessed expected value of each game by judging how happy Lucy would be with each game.

Participants. Fifty-six children and 17 adults participated. The younger age group included 14 girls and 14 boys between 4 years 10 months and 6 years 1 months, with a mean age of 6 years 0 months. The older age group included 17 girls and 11 boys between 7 years 9 months and 11 years 8 months with a mean age of 9 years 3 months. An additional 5-year-old was eliminated for not paying attention. Children were volunteers at a community after-school play program, attended by predominantly caucasian children from middle to lower middle class homes. Adults were UCL undergraduates and professional people educated to at least university entry level, between 17 and 45, with a mean age of 30.

Children were tested individually in sessions of 20 to 30 minutes each. Children did not attend the center regularly, and so only half the children in each age group participated in a second session typically given 1 to 2 months later. Which children participated twice depended on their availability. Adults were tested in one session of about 30 minutes.

Materials. Probability was manipulated by varying the length of the yellow and blue paper strips on which the marble could land. The strips were composed of 5 x 3 cm segments, surrounded in black pen. There were 3 different strips, made up of 4 yellow and 1 blue, 1 yellow and 1 blue, 1 yellow and 4 blue segments, corresponding to .2, .5, and .8 probability of landing on each color. Thus strips with different probability were 25 cm long, but the equal probability strip measured only 10 cm.

Prizes were laid out by the corresponding tube segment on a blue and yellow sheet of paper. Lucy could win bundles of crayons in mixed colors, either 1 (very small), 3, (small), 6 (large), or 10 crayons (very large prize). The blue outcome was associated with the two intermediate bundles, the yellow outcome with the very small and very large bundle. There were 12 different games altogether.

During instruction children saw Lucy shake the marble in 3 tubes with 1 blue and 5 yellow, 3 blue and 3 yellow, and 5 blue and 1 yellow segments. During experimental trials, however, children saw the paper strips without the marble to avoid the actual outcome influencing children’s judgments. Also, in these practice tubes, probability of landing on each color was confounded with the length of the corresponding color segment. In the experimental strips, 1, 1.4, 1.1, and 4.1 this was not so.

To make their judgments, children used a graphic “stick” scale. This consisted of 17 wooden dowels, increasing in height from 2.5 to 18.5 cm, with each stick 1 cm taller than the previous one. For each game, the child pointed to a stick to show how happy Lucy would be, with bigger sticks corresponding to better games. Even 4-year-olds can use this scale successfully in a linear fashion (Anderson & Schlottmann, 1991; Schlottmann & Anderson, 1994). Appropriate scale usage was elicited in the standard way by instruction with end anchors (Anderson, 1982, chapter 1; and see below).

Design. The main design was a 3 (probability) x 2 (yellow prize) x 2 (blue prize) within subjects factorial. Age (younger, older, adult) was an additional between subjects factor. All children judged two replications of the design in one session; half of the children at each age judged two more replications in a second session given to increase power. Adults judged 4 replications of the design in one session. Each replication was individually randomized for each subject.

Procedure. Initially, the puppet Lucy who liked guessing games was introduced, and asked the child to help her play. Lucy showed the child the tube and marble game, using 51, 3-3, and 1.5 tubes
(all games are listed as blue:yellow proportion). It was explained and demonstrated that it would be easier for the marble to land on yellow if there was more yellow; easier to land on blue if there was more blue. Thus one tube was best for landing on blue, one was best for landing on yellow, and the third tube was ok for both. This instruction unfolded in question-and-answer style. It was clear from this that even the youngest children had the right idea about where the marble would land before the experimenter made this explicit.

Next children were told that Lucy would win one prize if the marble landed on yellow, but another if it landed on blue. Then the different prizes were introduced and children sorted them from the worst to the best prize, that would make Lucy most happy. Children had no difficulty with this.

Next tubes and prizies were put together. Initially, a 5:1 tube was used with the largest 10 crayon prize by the long segment, and the smallest 1 crayon prize by the short segment. Children generally thought this was a good game, and the experimenter confirmed that this was because Lucy would win a big prize if the marble landed on blue, and there was lots of that color. The tube was replaced by the 3:3, then 1:5 tube, with the game becoming progressively worse. Then the experimenter switched the prizes around, and all children knew that now it was a great game again. Children were told that Lucy wanted to know just how good each game was, but because this was clearly a little tricky she needed a smart child to help.

Next the stick scale was introduced, with long sticks for good games, short sticks for bad games and medium sticks for ok games. The largest stick was made to correspond to the very best game there was, namely a 1:5 game with 10 crayon prizes on both segments, whereas the smallest stick was made to correspond to the very worst game, namely a 5:1 game with single crayons on both segments. These games were left by the ends of the scale throughout as a reminder of scale orientation. A 5:3 tube with a single crayon on blue and a 10 crayon bundle on yellow was used to illustrate a "so-so" game.

Children made 10 judgments for practice games, using strips of paper only ("they're just the same as the tubes but they don't roll around as much"). During practice, only one aspect of the game, probability or one prize was changed from one trial to the next. Practice games began with the 3:3 configuration with 1 crayon on blue, 10 on yellow. Games first improved, then decreased on successive trials, with the last trials involving a switch from a very bad game (5:1 with 3 crayons on blue and 1 on yellow) to a good game (5:1 with 3 crayons on blue and 10 on yellow). Practice used minimal directions, with verbalization mainly on the first trials ("what do you think of this game? Is this better than before or not so good? How happy would Lucy be?"). If initial adjustments were to the end of the scale, children were told that they would run out of sticks, as demonstrated by the next trial. Counting sticks as occurred initially with a few older children was discouraged. They were told to just go by their feeling, and that there wasn't one right answer, because "these aren't numbers, they're just sticks". Other than that, no feedback was given. After the practice, lazy Lucy went for a nap, and the child continued with two randomized replications of the set of experimental stimuli. The initial instruction-practice took about 15 to 20 minutes, the experimental trials another 5 to 10 minutes.

Children participating in a second session were reminded of the task and given a few additional practice trials. To explain the set-up to adults, they were told that they served as controls in a study designed for small children. They were also told not to pretend child-like behavior, but to make intuitive judgments according to their own feeling about the games.

Results

The main findings were clear: All ages took both alternative outcomes of the game into account. For the within-outcome integration, all ages used multiplication. The across-outcome integration, however, showed individual differences and often did not conform to the normative addition rule. Accordingly, the group results will be discussed first, followed by the individual analyses.

Figure 3 shows the mean judgments for the 3 age groups. The normative pattern is shown again on the top right. Data patterns at all ages appear generally similar to the normative predictions.
The slope of the curves also suggests that children see relative, not absolute quantity as a cue to probability. The tubes used for 1:4 or 4:1 probability the prizes had 5 units length, whereas the 1:1 tube had only 2 units length. If children simply considered the absolute space for winning (e.g., Hoemann & Ross, 1971, 1982), one might expect the curves to bow down in the middle. This was not observed. Together with the curve reversal this suggests that children as young as 5 years possess a fairly abstract representation of probability as an expectancy about the likelihood of a target event. This confirms prior studies (Anderson & Schlottmann, 1991; Anderson & Wilkening, 1991; Schlottmann & Anderson, 1994).

Statistical Analysis. Of main interest in the overall 3 age x 3 probability x 2 blue prize x 2 yellow prize mixed model ANOVA are the interaction patterns. The reversal of slope described above was reflected in a significant probability x yellow prize interaction, F(2, 140) = 299.87 (all p < .05 unless noted). The interaction differed between the ages, F(2, 140) = 4.50, with reversal slightly more pronounced in older than younger children or adults, but it was significant for each age group, F(2,54) = 61.63 and 182.15 for younger and older children, and F(2,32) = 144.00 for adults. The other interactions from the overall ANOVA bear on the questions of within- and across-outcome integration and are deferred to the relevant sections below.

The overall ANOVA also showed main effects for the blue and yellow prizes, F(1,70) = 216.78 and 1338.91 and for probability, F(2,140) = 42.06, indicating that children took account of both possible outcomes of the game. All three main effects differed across the age range, F(2,70) = 10.06 and 5.88 and F(4,140) = 3.42, respectively. These age effects appear to reflect minor differences in scale usage, and significant main effects of all three factors were found when the data for each age group were analyzed separately.

As is customary, the main analysis simply averaged the data across replications. However, because children participated in one or two sessions, preliminary ANOVAs established that this was appropriate. First, when data from the second session were excluded, the same pattern of significances was found as when all the data were included. Second, the performance of children participating once versus twice did not differ in the first session: When this group factor was added to the analysis, only one of its 16 effects was marginally significant, F(2,104) = 3.44, p = .036, but it merely reflected a slightly more extreme reversal pattern for older children participating once. Thus attrition does not seem related to the variables under study. Third, when first versus second session performance was
compared for children participating twice, none of the 16 effects involving the additional session factor was significant, with the largest reaching $F(1,26) = 1.66$. Thus, there do not seem to be practice effects. In sum, no major differences were apparent between children participating once or twice or between first and second session data. Accordingly, group and session factors were not considered further.

**Within-Outcome Integration: Multiplication of Probability and Prize on Blue.** In each panel of Figure 3, the top and bottom pairs of curves diverge towards the right (this may be seen clearly by considering the vertical bars, which have the same size on the right and left). In the bottom pair, expected value increases faster when the blue prize consists of 6 rather than only 3 crayons. In the top pair, expected value decreases faster when the blue prize consists of only 3 rather than 6 crayons. These fan patterns indicate multiplicative integration of probability and prize on blue.

Multiplication is statistically reflected in a probability x blue prize interaction with a bilinear component. Both were observed, $F(2,140) = 20.72$ for the interaction in the overall ANOVA, and $F(1,70) = 32.82$ for the bilinear component. There was no change across the age range in either of these, $F < 1$ for both. The interaction was significant for each individual age group, $F(2,54) = 8.80$ and 5.86, for younger and older children, and $F(2,32) = 7.42$, for adults, as was the bilinear component, $F(1,27) = 13.36$ and 9.38 for the children, and $F(1,16) = 12.50$ for adults.

Both the visual and statistical analysis show multiplicative within-outcome integration of probability and value at all ages, a result which agrees with Anderson (1980) and Schlozmann & Anderson (1994, session 2). These data suggest that children as young as 5 or 6 are capable of integrating probability and value as appropriate under the expected value model for an individual outcome. Multiplication was not built into the physical display or instruction, and hence indicates conceptual understanding that children brought to the experiment. Apparently, children can apply their expected value concept even in a complex 2-outcome situation.

**Across-Outcome Integration: Additivity of the Two Outcomes?** Under the expected value model of Equation 2, both outcomes contribute independently to overall expected value. This would be reflected in the absence of a blue x yellow prize interaction, despite significant main effects. The interaction was nonsignificant in the overall ANOVA, $F(1,70) = 1.91, p = .171$, as was the blue x yellow prize x probability interaction, $F < 1$. The interactions with age were nonsignificant as well, $F < 1$, and $F(2,140) = 1.41, p = .232$. However, the single subject analysis below shows that this impression of additivity at the group level is misleading.

**Individual Analysis.** Single child analyses are desirable because group averages may misrepresent individual performance. However, children have high error variability and will not sit still for many replications. Thus, their data patterns are often not stable enough for individual analyses to be informative, and single subject statistics are hampered by lack of power.

This difficulty also appeared in single subject ANOVAs for the present study (with 12 or 36 df for error depending on whether children had judged 2 or 4 replications): Only the theoretically largest probability x yellow interaction, reflecting the curve reversal in Figure 3 was significant for most participants, 20 younger children (71%), 27 older children (96%) and all 17 adults (using a criterion of $p < .1$ to increase power). However, probability x blue prize interactions, to reflect fanning and multiplicative within-outcome integration, were rare and significant only for 1 (4%), 4 (14%) and 9 (53%) participants, respectively. Even blue x yellow outcome interactions, to indicate deviations from across-outcome additivity, appeared more frequently, for 8 (29%), 5 (18%), and 6 (35%) participants. Not too surprisingly, children with significant effects typically had judged four replications, as had the adults. Therefore, the usefulness of the single subject statistics was limited.

More informative was an analysis inspecting individuals' patterns of means for the relevant interactions. This was done by an objective procedure comparing the simple effects observed (averaged across replications) with those predicted by the EV model; the predictions below can be easily verified by reference to the normative data in Figure 3. 2

**Multiplication and Reversal as the Majority Pattern for Individuals.** For the within-outcome integration, multiplication (as the predicted form of the probability x blue prize interaction) was diagnosed if the simple effect of the 6 versus 3 crayon prize on blue was larger with high than with low probability of this outcome occurring. This pattern appeared for the majority at all ages, 24 younger children (86%), 19 older children (68%), and 14 adults (82%). There was no age difference in the frequency of this pattern, $\chi^2 (2) = 2.84$. The individual and group analysis thus agrees that for the within-outcome integration, multiplication is the majority pattern at all ages.
The probability reversal in the group data was also the majority pattern for individuals. Reversal (the predicted form of the probability x yellow prize interaction) was diagnosed if games with high probability of the yellow outcome had higher means than low probability games provided yellow carried the larger of the two prizes, but the opposite appeared when yellow carried the smaller prize. This was found for 19 younger children (68%), 27 older children (96%), and 15 adults (88%). The small age difference in frequency of the reversal pattern is significant, $\chi^2(2) = 8.68$. Children without reversal tended to simplify the task by ignoring probability in some conditions, but for 3 younger children judgments increased with probability of the yellow outcome even when it carried the smallest prize — these children might have responded to physical changes in tube length rather than to probability. The main point, at any rate, is that the means-based analyses found good agreement between group and individual data for the probability reversal and the multiplicative within-outcome integration.

*Individual Differences in the Across-Outcome Integration.* For the across-outcome integration, however, the individual analyses revealed that the group data were misleading. Individuals showed different types of yellow x blue outcome interactions in the means pattern. These cancelled each other out in the group analysis, and so there was little sign of non-normative effects at the group level. However, deviations from additivity appeared clearly in an individual difference classification, with many children showing risk seeking or risk averse patterns.

Theoretically, additivity appears in the pattern of means if the simple effects of the blue prize do not depend on what the yellow prize is. In practice, of course, some variability is unavoidable and the simple effects will not be exactly equal. A small difference of 1 point or less (on the 17 point scale) was thus taken as meaningless and additivity rejected only if the simple effects differed by more than this. (Smaller but statistically significant blue x yellow interactions also had to be excluded, of course, but there were only two cases.) Accordingly, participants were considered additive if they showed some effect of the blue prize, but no blue x yellow interaction.

The other participants fell into 3 groups: Many did show such blue x yellow interactions and these configural patterns were classified as either risk seeking or risk averse (see below). A few children, however, did not fit the classification, because they had no effect of the blue prize or because their data showed large ordinal violations. As would be expected, such unstable data patterns appeared mainly for children judging only two replications. Figure 4 shows the mean judgments for the resulting groups.

![Figure 4](image-url)
The third column shows an opposite risk averse pattern, with more separation in the top than bottom pair of curves. Here the same blue prize has more effect when it is smaller than the yellow prize, and hence the riskless option. Finally, the rightmost column shows data for the remaining unclassifiable children. This pattern shows little effect of blue overall and thus provides no evidence on the across-outcome integration.

The homogeneity of participants within these groups is supported by separate ANOVAs for the different groups. These found significant across-outcome, blue × yellow interactions for the two configural groups ($F > 34.84$ -- with df depending on group size in Figure 4), but not in the additive group ($F < 1$). The additive and both configural groups, however, showed significant probability × blue and probability × yellow outcome interactions ($F > 8.27$) for the multiplicative within-outcome integration and the probability reversal. The probability, blue and yellow main effects were all significant as well ($F > 5.20$). The unclassifiable children showed neither a main effect or interactions involving blue. None of the groups showed age differences in the interaction patterns ($F < 2.12$).

When the distribution of participants across these individual difference groups is considered, half of the participants at each age, 14 (50%), 14 (50%) and 8 (47%), show one of the two configural pattern, with risk seeking (second column) occurring twice as often as risk aversion (third column) at all ages. In contrast, only 5 younger children (18%), 9 older children (32%) and 9 adults (53%) show the additive pattern. Finally, 9 younger (32%) and 5 older children (18%) are unclassifiable.

The small increase with age in participants showing the additive pattern relative to those who do not is significant, $\chi^2 (2) = 6.04$. There appears to be a more pronounced decrease with age in unclassifiable rather than configural participants, but this difference does not reach significance, $\chi^2 (2) = 4.55$. The main point, at any rate, is that failures of the additive model for across-outcome integration as previously reported for adults also appear for children.

Experiment 2

The first study adds to the mounting evidence that children conform to the EV model for the within-outcome integration; thus even 5-year-olds used a multiplication strategy and appeared to understand that probability acts as a modifier of value. However, many children deviated from the EV model for the across-outcome integration, using configural approaches rather than the normative addition strategy; similar to results with adults (Shanteau, 1975).

While this was the first study allowing a test of children's across-outcome strategy, previous work had considered their within-outcome strategy. These studies, however, had not always found support for the multiplicative within-outcome strategy apparent here. A limitation of this study lies in its peculiar sampling: Not all the children had been available for a second session, and for those who were there was often a considerable delay between first and second session. Although no effects of subject selection or practice/development were discernible in the analyses, such effects can, of course, not be ruled out definitely. Finally, the present evidence on children's risk seeking or risk averse across-outcome strategy emerged only from a post-hoc classification. In light of these points, it seems important to replicate the present study in a new sample.
Method

Experiment 2 had the same design, materials and procedure as Experiment 1, except that all children judged four replications in two sessions which took place 4 to 6 days apart.

Participants. Twenty-eight children participated. The younger age group included 6 girls and 8 boys between 6-4 and 7-5, with a mean age of 6.6, the older age group included 8 girls and 6 boys between 8-11 and 10-4 with a mean age of 9.4. The children were predominantly caucasian, volunteers at a primary school of middle class character.

Results

Children’s mean judgments are shown in Figure 5. The group data closely resemble both the normative predictions and the data obtained in Experiment 1.

![Graph showing mean judgments for Younger and Older Children](image)

**Figure 5** Mean judgments of happiness as a function of probability in Experiment 2. The two graphs show the integration of probability and value for younger and older children, with lines for each probability and value.

As in Experiment 1, judgments at both ages show a probability reversal: When yellow carried the larger of the two prizes judgments increased with the probability of obtaining the yellow outcome, but when yellow carried the smaller of the two prizes judgments decreased with the same probability. This reversal was reflected in the probability x yellow interaction, F(2,52) = 230.17, in the age x probability x blue x yellow mixed model ANOVA.

As in Experiment 1, the fanning in the bottom and top pair of curves in each panel indicates that children at both ages integrated probability and prize on the blue outcome multiplicatively, as prescribed by the EV model for the within-outcome integration. This was reflected in the probability x blue interaction, F(2,52) = 42.51, with a significant bilinear component, F(1,26) = 63.91.

The EV model for the across-outcome integration predicts a nonsignificant blue x yellow outcome interaction with the two outcomes contributing independently to overall expected value. As in Experiment 1 this was obtained, F(1,26) = 2.35, p = .14.

The main effects of probability, blue and yellow outcome were significant as well, with Fs > 10.17. There were no significant effects involving age, and the same interactions were significant when the data for each age group in each session were analyzed separately. In sum, these group results replicate those of Experiment 1 in all respects.

Individual Analysis. Single subject ANOVAs were largely uninformative, as in Experiment 1. All but one older child showed significant probability x yellow interactions, reflecting the probability reversal. However, only 8 younger and 3 older children showed significant probability x blue interactions, as expected under the multiplication rule. Also, only 2 older and 2 younger children showed blue x yellow interactions, statistically indicating deviations from across-outcome additivity.

The means-based analysis was more informative. It was done in the same manner as Experiment 1 and showed very similar results. The only difference was that in Experiment 2 slightly more children were classifiable. This could be due to all children judging four replications in two sessions and/or to testing taking place during the schoolday rather than outside the school environment during playtime. Accordingly, the results of the individual analyses were even clearer than in Experiment 1.

In Experiment 2, probability reversals appeared for all but one older child, whose judgments increased with the probability of the yellow outcome regardless of whether this carried the larger or smaller prize (100% and 93%, compared to 68% and 96% in Experiment 1). A diverging fan, indicating multiplication of probability and prize on the blue outcome, appeared for all 14 younger and 11 older children (100% and 79%, compared to 86% and 68% in Experiment 1). These results strengthen the conclusions previously drawn: Abstract probability understanding and multiplicative within-outcome integration, as predicted by the normative EV model, appear for both age groups, in both samples, at both the group and individual level.

As in Experiment 1, children differed in how they treated the across-outcome integration. The classification developed for Experiment 1 was applied, and mean judgments for the resulting
individual difference groups are shown in Figure 6 averaged over the two ages. As previously, additive children show similar separation within the top and bottom pair of curves (left). The other two patterns deviate from the normative model: Risk seekers show more effect for the larger, risky outcome, seen as less separation in the top pair of curves for games sharing the larger prize than the bottom pair for games sharing the small prize (middle). Risk averse children, in contrast, show more effect for the smaller, riskless outcome, seen as less separation in the bottom pair for games that share the smaller prize (right).

![Figure 6](image)

**Figure 6** Individual difference classification for the across-outcome integration in Experiment 3. The figure shows mean judgments of happiness for participants employing the additive (left column), risk seeking (middle column), or risk averse strategy (right).

Regarding the distribution of participants, in Experiment 2 6 younger and 7 older children showed an additive pattern for the across-outcome integration (43% and 50%) compared to 18% and 32% in Experiment 1. Thus, slightly more children performed in accordance with the EV model, while there were fewer unclassifiable children — only one older child did not fit the classification, as opposed to 25% in Experiment 1.

More importantly, however, as in Experiment 1 half of the children deviated from the additive across-outcome expectation (57% and 43% as opposed to 50% at both ages in Experiment 1). Also as in Experiment 1, these deviations were more often of the risk seeking than risk averse type. Five (36%) younger and 4 (29%) older children showed more effect of the risky outcome, whereas 3 (21%) younger and 2 (14%) older children showed more effect of the riskless outcome. Overall, the pattern of nonadditivities for the across-outcome integration appeared similar in both experiments.

**DISCUSSION**

The present study demonstrates that children have good intuitive understanding of expected value even in a complex situation with alternative winning outcomes. Like adults, children multiplied values and probabilities, and also like adults, they did not add component EVs precisely, but often overemphasized one or the other outcome. This section first discusses some processes that might underlie children's within- and across-outcome integration, then goes on to consider more general implications of the findings.

**Expected Value Model: Multiplicative Within-Outcome Integration**

At all ages, in both experiments, and at both the group and individual level, probability and value of a single outcome were combined multiplicatively, in line with the normative EV model. This confirms Anderson (1980) and results from Schlottmann & Anderson's (1994) "tube" task. Why children did not multiply in the "spinner" task of the latter study is unclear — it does not seem to be a peculiarity of spinners, because Anderson (1980) used a spinner task as well. With this exception, at any rate, the studies agree on multiplication. Because this integration rule was not part of the display or instruction it reflects children's pre-existing understanding. It seems that children as young as 5 or 6 years have conceptual understanding of expected value and can apply it to demanding situations with two alternative prizes.

That multiplication appears this early in the expected value domain contrasts with results on judgments of physical quantities, where additive integration is typical below 8 years even if the normative physical rule is multiplicative (e.g., Anderson & Cuneo, 1978; Cuneo, 1982; Wilkening, 1981; Wolf & Algom, 1987). This difference could stem from differential familiarity with the domains, if children have more experience with evaluating uncertain goals than with evaluating quantities. However, children have many occasions to contemplate how much of something there is, as well as to consider how worthwhile a course of action might be. It does not seem entirely obvious that differential amounts of experience would account for this task difference.

**Multiplication as Weighting Operation in Expected Value Judgment.** A different possibility is that EV is special because probability acts as a modifier of value: when probability and value are combined, the resultant impression is still one of value. Thus the integration involves only two distinct dimensions, with their multiplication reflecting that one variable (probability) weights the other (value). In many physical concepts, in contrast, multiplication is a conjunctive operation: Two dimensions are combined to form a third, qualitatively different dimension as a result. This holds in those cases where
addition is best documented in children, e.g., area judgments as a function of width and length of an object (e.g., Anderson & Cuneo, 1978; Wolf & Algom, 1987), judgments of numerosity as a function of length and density of an array (Cuneo, 1982), or time judgments as a function of distance traveled and speed of the object (e.g., Wilkening, 1981). The change of perspective involved in moving between components and resultant may make multiplication as a conjunction operation harder to grasp for children than multiplication as a weighting operation.

Task variability, as just discussed for different multiplication tasks, is commonly observed in development. Often it due to surface features inadvertently influencing processing (e.g., Reyna & Brainerd, 1995), but in other cases it occurs because formally isomorph tasks need not be cognitively isomorph: Here it is suggested that the same compensatory, multiplicative structure may map onto different schematic situations. This is analogous to arguments made, for instance, in work on conditional reasoning (e.g., Cheng & Holyoak, 1985) and arithmetic (e.g., Nunes & Bryant, 1996).

More work is required, however, to study whether the distinction between a conjunction or weighting operation is useful for understanding children's multiplicative reasoning.

Expected Value Model: Additive Across-Outcome Integration

In contrast to most participants using the normative, multiplicative within-outcome integration, the normatively expected additive across-outcome integration appeared far less often in both experiments. The present results with children agree with previous work on adults. Shanteau (1974, 1975) found failures of additivity using an information integration approach and also reviewed similar findings obtained with different methods. The nonadditivities here could, of course, simply be due to nonlinearities in children's use of the response scale. However, there is extensive evidence for response linearity with the present methodology (see Anderson & Wilkening, 1991; Schloßmann & Anderson, 1994; Surber, 1984), and Schloßmann (2000) found nonadditivity of EV in children's judgments that cannot be explained as nonlinearity. Nonadditivities thus do not appear to be artifacts of method, although their cause remains unknown (see also Anderson, 1996).

A Compromise Strategy for Expected Value Judgment: To cope with the complexity of the present task children may use some form of serial strategy. The data rule out, however, that they use a simple addition-only, outcome-by-outcome strategy (Schloßmann & Anderson, 1994). Instead, children's response may be seen as a compromise between the two values, weighted for their probabilities. Thus children locate the values of alternative outcomes on the response dimension, then fractionate the distance between them.3

This compromise strategy for two outcome events is an extension of the fractionation strategy previously proposed for one outcome events and supported by reaction time data from adults (Anderson, 1980; Lopes & Ekberg, 1981). In the single outcome case, the value is located on the response dimension, then an adjustment is made in proportion to the probability, with a lower bound of zero. The present two-outcome case differs only in that the lower bound is variable and non-zero. The functional procedure for simple and complex EV judgment may thus be similar. This makes the repeated finding of young children's sophisticated performance in the complex case less perplexing.

Anecdotal support for the compromise strategy comes from the observation that some children during practice initially used both hands to point to two sticks for the two different valued outcomes. After the experimenter stressed again that they could only point to one stick to show "how good the whole game is, both together, not just the blue or the yellow", they would often choose a response in between. These children seemed initially to conceptualize the task in terms of competing responses. A compromise then may be constructed in a second step.

Children have two ways to find this compromise between the components: They could start from the larger value and make a downward adjustment, or they could start from the smaller value and make an upward adjustment. The model of Equation 2 may be reformulated to capture these approaches.

\[
EV = V_L - P_S (V_L - V_S) \quad (3)
\]

\[
EV = V_S + P_L (V_L - V_S) \quad (4)
\]

Equation 3 expresses downward adjustment from the larger value, \(V_L\). The size of this adjustment is a proportion of the possible change, \(V_L - V_S\), weighted for the probability, \(P_S\), of only winning the smaller value. Similarly, Equation 4 expresses upward adjustment from the smaller value for the probability, \(P_L\), of winning the larger value. Equations 2, 3 and 4 are formally, but not cognitively equivalent: Equations 3 and 4 imply that children compare alternatives and represent them as the larger and smaller value. This ordinal representation can be related to the outcome configurabilities in the data.

Risk Attitudes in Children: Risk seekers attend more to the risky option than the smaller value. If they adopt the approach of Equation 3, for example, they would start from the more exciting larger value, \(V_L\), then make an insufficient adjustment for the less interesting smaller value. This would
result in similar judgments for games that share the large 10 crayon prize but differ in the smaller prize, as seen in the second column of Figure 4 and 6.

Risk averse children, in contrast, attend more to the riskless option than the larger value. If they adopt the approach of Equation 4, they would start from the point of certain gain, \( V_s \), then make an insufficient adjustment for the uncertain larger value outcome. This would result in similar judgments for games which share the certain single crayon prize but differ in the larger prize, as seen in the third column of Figure 4 and 6. In general, risk seekers may give more weight to the risky option, \( V_{1r} \), than to the sure thing, \( V_s \), while risk averse children do the opposite. The normative addition pattern, in contrast, occurs with equal weights for both.\(^4\)

The notions of risk seeking and risk aversion originate in the literature on risk attitudes in adult judgment/decision (e.g., Bronmili & Curley, 1992; Lopes, 1987). Their standard operationalisation is somewhat different (i.e., risk aversion is usually diagnosed if a person prefers a riskless outcome to a risky gamble of equal/higher EV), but the underlying definition seems comparable. The present view fits especially well with Lopes’ (1983, 1984; 1987; Schneider & Lopes, 1986) account of motivational differences in adults’ judgment/decision. Lopes argues that individuals have dispositions towards security and potential. A stronger motive for security – to avoid losses – involves a tendency to overweight the worst outcomes of risky situations, whereas a stronger motive for potential – to achieve gains – involves overweighting the best outcomes.

Individuals differ in their motives to seek or avoid risk, but how they act, according to Lopes, also depends on the current aspiration level. This, in turn, depends on situational factors as well, for instance, whether the options are framed as gains or losses: Adults are often risk averse and choose a sure gain over a gamble of equal EV, but become more risk seeking when the gamble allows them to avoid a sure loss (Tversky & Kahneman, 1981).

The present task presented a gain frame, but in contrast to the finding just cited, more participants were risk seeking than risk averse. Two accounts, not mutually exclusive, may be given of this. One appeals to other situational factors: Schneider (1992) also found risk seeking for gains when gambles had no zero outcome, as in the present study, but risk aversion with a zero component. Moreover, in standard tasks people choose between different gambles, while here they merely evaluate a single gamble. Choice and evaluation paradigms may promote different forms of processing (see Payne, 1982).

The second account focuses on developmental influences: Children and adolescents are often described as being more oblivious to the risks in their lives than adults (see discussion in Furby & Beyth-Marom, 1992). There are few experimental results on this, but Reyna & Ellis (1994) found that preschoolers to fifth graders chose the gamble more often than the sure gain, even in a choice task with zero outcome gambles. Individual performance was not reported in that study, but on the group level children appeared more risk seeking than adults usually are in similar tasks.

In the two experiments here, risk seeking was more frequent than risk aversion at all ages, even for adults. Tentatively, this suggests that the risk attitudes found here depended in part on situational factors. This, of course, does not preclude developmental effects. To understand the origins of risk attitudes in children, work with a wider range of ages and tasks is needed. The present study merely reports initial evidence that individual differences in such risk attitudes may appear from a young age.

Intuitive versus Formal Probability Understanding

Early Competence versus Late Biases. This study, as many other recent ones, highlights young children’s competence. The bulk of work with adults, in contrast, emphasizes shortcomings in their judgment/decision (e.g., Kuhneman, Slovic & Tversky, 1982). In part, the discrepancy is due to different tasks. Adults, like children, do well when approximate, intuitive solutions apply. Children, like adults, have difficulty with achieving computational and logical accuracy. However, the biases and heuristics that produce systematically “irrational” responses in adult judgment/decision do not just involve weak computational skills, but often occur because incorrect intuitive solutions are given instead of correct computational ones (e.g., Reyna, 1991; Reyna & Brainerd, 1991). Intuition can both help or hinder reasoning.

Computational reasoning is traditionally seen as the “higher” and ultimately more adaptive form of thought, but the contrast of tasks in which people fail or excel suggests that adaptive success often lies in the match between situational demands and cognitive strategy. In many circumstances, computation and intuition lead to similar solutions -- one being more precise, the other a cognitively more efficient shortcut -- and dissociations may occur mainly in ecologically invalid, misleading situations (e.g., Giggenenzer & Murray, 1987; Anderson, 1991b). Regardless of why dissociations occur, however, their existence points towards the potential independence of intuitive and computational processing. Under Reyna & Brainerd’s (1994, 1995) Fuzzy Trace theory this is possible
because people form multiple, parallel encodings of a given problem, with separate representations of literal features and intuitive "gist" providing alternative bases for reasoning. According to Reyna & Brainerd, adults prefer intuitive processing unless the task is clearly perceived to require a computational approach.

The consideration of intuitive biases in adult judgment suggests that it is not sufficient for developmentalists to replace the traditional emphasis on children's computational skills with a new focus on their intuitive competence -- in addition, we must also consider the interplay of intuition and computation. The two are rarely considered together, however, and so little is known about how development in one influences advances in the other (see discussion in Ahl et al., 1992). Nevertheless, it is clear that causality works in both directions: Ahl et al (1992) found that performing an intuitive task improved subsequent performance on a numerical version. Conversely, Agnoli (1991) showed that logic training decreased use of a nonnormative heuristic. More generally, it is often assumed that the acquisition of expertise, including appropriate computational procedures, sharpens experts' intuition (e.g., Fischbein, 1987). Conversely, it is also widely believed that to help children overcome difficulties with computation we need to give the procedures intuitive meaning (e.g., Resnick, 1995). The mapping of intuition and computation is complex, but has many implications for education in the probability domain and beyond.

Even if a narrower issue is considered, that of the developmental origins of the biases and heuristics in judgment/decision, the findings are diverse. In the present study, children and adults seemed to take similar nonnormative risk attitudes. In a study of the sure thing/dominance principle, children violated it more often than adults (Schlottmann, 2000). On the other hand, framing effects (Reyna & Ellis, 1994) and sunk cost effects (Webley & Plaisier, 1997; cited in Arkes & Aton, 1999) may increase with age. And of two studies on the representativeness heuristic, one also found some increase with age (Jacobs & Potenza, 1991), the other found strong effects at all ages that could be suppressed after logic training (Agnoli, 1991).

The direction of development in reasoning is usually taken to be from intuition to computation (e.g., Piaget, 1967; Surber & Haines, 1987) and this would imply that the biases found with adults should be even stronger with children -- but from the data cited above this does not hold generally. Reyna & Ellis (1994), in contrast, argue that development may often be from computation to intuition. Children, like adults, are seen as capable of both intuitive and (rudimentary) computational processing, but children are less able to resist surface interference. Accordingly, if gist and literal problem features do not coincide, children's approach mainly depends on surface aspects. In many standard developmental tasks (e.g., conservation), this means children are distracted from underlying correct solutions, while in the biasing scenarios of the judgment/decision field they are distracted from underlying incorrect solutions. This intriguing hypothesis can account for some biases increasing with age -- but again it does not seem to account for all of them. More than one type of developmental link between intuition and computation may exist even within a single domain.

Frequentist versus Subjectivist Conceptions of Probability. The basic metaphor of the standard frequentist approach to probability is that of a lottery. In such systems, probability can be determined over repeated events as the relative frequency with which the target occurs (or a correlate, e.g., length of the tube segments or number of lottery tickets). Thus we observe properties that covary with probability over time and space, even though the actual probability of the individual event is unobservable -- probability is an abstract event feature. To select informative cues to it depends crucially on our knowledge of the generative system and its possible states.

Children, in the traditional view, do not appreciate the abstract concept, failing to differentiate probability from physical features of the target event (e.g., Hoemann & Ross, 1982). Difficulties with the logic of combinations and the computation of proportions prevent them from understanding that probability is a relative quantity that requires attention to both targets and nontarget alternatives. In contrast to this view, however, recent work has shown that even 5-year-olds' intuitive probability understanding is reasonably abstract. This also appears in two aspects of the present data.

First, children understand probability dependence: As one outcome becomes more likely, the probability of an alternative decreases. The blue and yellow tube segments varied independently, with 5 units length for 4:1 and 1:4 tubes, but only two for the 1:1 tube. If children simply consider magnitude of the winning area (Hoemann & Ross, 1982), judgments should bow down for the shorter tube, but they didn't. Similarly, Schlottmann & Anderson (1994) found judgments were not bowed up when a longer tube represented even odds. Instead, in both studies judgments increased appropriately with probability. Children's understanding of dependence may not be precise, but they clearly appreciate that relative, not absolute quantity signals probability.

Second, children are selective in their interpretation of the physical cues and flexibly adapt it to the value structure of the situation. In this study, the same physical manipulation reversed its effect.
depending on the values of the alternatives. Similarly, in Schollmann & Anderson (1994), identical manipulations had an effect in risky, but not riskless games. Clearly, children do not just learn during instruction to associate changes in stimulus and response: Instead they seem to know that the stimulus represents the likelihood of a more or less desirable event, and that its meaning depends both on this target and the alternatives to it.

That even young children intuitively separate probability from the physical cues is surprising under the frequentist view of probability development, but it is expected under a subjectivist view. Extending recent approaches in mathematical probability (e.g., Fishburn, 1986), Anderson (1991a) argued that subjective probability should be defined as degree of belief or uncertainty. In this view, attention to relative frequency is unnecessary for early probability understanding. Two other aspects are basic instead: First, children must grasp that different outcomes are possible. Second, they must grasp that each outcome is more or less certain. Both features appear in children's approach here: They realize that Lucy can win only one of the prizes, and that the actual outcome is uncertain, hence the need for a compromise. Children's estimate of the blue-yellow proportion is co-opted into this approach and informs how the alternatives are weighted. But in the subjectivist view, this is not the crucial feature – other sources of information about uncertainty could be equally used.

Such information is available in many circumstances: Children distinguish, for instance, between situations in which inferences can be made with certainty or not (Fabricius, Sophian, & Wellman, 1987; Sophian & Somerville, 1988). They often give definite answers even in indeterminate situations and this was initially taken to indicate difficulty with recognizing uncertainty (e.g., Pierrault-Le Bonniec, 1980), but it later appeared that children's problem lies mainly with withholding judgment (e.g., Byrne & Bellin, 1991; Horobin & Acredolo, 1989). Acredolo & O'Connor (1991) argued this occurs because choice procedures discourage the expression and measurement of uncertainty. In contrast, with continuous judgment procedures, as used here, children express feelings of uncertainty in a natural manner and this enables them to display a higher level of competence.

In general, when we consider our chances of goal attainment in everyday life, this is not necessarily based on direct consideration of relative frequencies – instead feelings of uncertainty can be related, for instance, to knowledge about task difficulty, effort and ability level. These concepts are stressed in theories of motivation (e.g., Heckhausen, 1982; Schneider, Hanne & Lehmann, 1989; Weiner, 1991). Such theories also build on the EV conception, but the link with judgment/decision is rarely considered, perhaps because subjective uncertainty about successful action is difficult to map onto schoolbook, lottery-type probability. If subjective probability is seen as degree of belief, leaving open how it is instantiated, it is easier to appreciate its many different facets in cognitive functioning. It is also easier to appreciate the many different opportunities even young children have to learn about it.

CONCLUSIONS

In the everyday world, we rarely judge probabilities per se, instead we judge expected value of our goals. In such a context, even the youngest children in this study showed conceptual understanding of expected value. They knew that probability acts to weight the goal value, even in a complex situation with nonzero alternatives. It appears that children acquire the EV intuition in their everyday life, a concept functional by the time they start school and before formal instruction with probabilities.

In the present study, all ages used similar intuitive operations. This would seem to suggest that there is similar intuitive potential for the teaching of judgment/decision in children and adults.

Children's expected value reasoning seems advanced compared to other areas of multiplicative reasoning. Much recent research, however, has emphasized the domain and task specificity of cognition. The present finding agrees with other recent work on probability development in stressing early competence. This is surprising if probability is seen as an abstract mathematical concept and in light of the difficulties that adults still have with this. It makes much sense, however, if probability is seen as an ingredient important for effective reasoning about one's own and other people's goals and desires in an uncertain world. Probabilistic reasoning, like memory, becomes salient mainly when it fails, but the many situations in which it serves adults -- and small children as well -- should not be overlooked.
ACKNOWLEDGEMENTS

Many thanks are due to the children and staff at Coram's Fields, London, and at the Parish C. of E. Primary School, London Lane, Bromley, for their participation and support, and to Christine Brooks and Jane Tring for help with data collection. I also thank Norman Anderson, Bill Gaser and Lotella Lepore for their comments. A preliminary report of some of these data was given at SRCD 1997, Washington, DC.

ADDRESS AND AFFILIATION

Anne Schlottmann
Department of Psychology
University College London
Gower Street
London WC1E 6BT UK
+44 20 7679 5383
a.schlottmann@psychol.ucl.ac.uk
REFERENCES


FOOTNOTES

1 As a side issue, the pattern in Figure 2b also indicates understanding that if one probability increases, the other decreases. If children understand precisely that alternative probabilities must sum to 1, the curves are flat under an addition rule. Opposing slopes, to reflect that overall expected value increases with the probability of obtaining the larger prize, as shown here, would be obtained if children overemphasize this probability.

2 In essence, these means-based individual analyses use a more sensitive interaction test than the ANOVA to avoid accepting the statistical null-hypothesis of additivity by default. The multiplication and reversal tests were directional because the normative model made exact predictions about the form of these interactions. For the across-outcome integration, in contrast, the normative model predicts additivity and so the interaction test was nondirectional.

3 Five-year-olds have been shown to use a simple equal weight compromise strategy in a serial judgment task before (Schloßmann & Anderson, 1995), but the present case involves differential weighting in proportion to the probabilities. Note that children do not blindly divide the distance between the values on the response scale in the same way as in the tube representing the probability, i.e. by picking a stick corresponding in relative position to the dividing line between the blue and yellow segments. If they did this, they would, for instance, give higher ratings to games with a high probability of winning the medium prize and a low probability of winning the very large prize (top right point in Figure 3) than if the probabilities were reversed (top left point), just because in the first case the dividing line would be closer to the large value than in the second case. The fractionation thus occurs according to probability, not according to the physical implementation.

4 To reflect such individual differences, the probability variable in Equations 3 and 4 could be replaced with a general weight parameter, to depend on probability as well as on attentional/motivational factors (see Anderson, 1996), or preference weights could be attached to the values.

FIGURE CAPTIONS:

Figure 1. Schematic of two sample games (A marble could land on either tube segment, and the puppet would win the prize placed by that segment. The two games illustrate that the same physical cue has different meaning in the context of different games: In the top example, the one unit yellow segment represents .2, in the bottom example .5 probability. In the top example, the six crayon prize for blue makes it the higher value, risky alternative, but in the bottom example this is the lower value sure thing.)

Figure 2. Predicted patterns. (2a illustrates the normative expected value model; 2b illustrates nonnormative within-outcome integration under the addition-only rule; 2c illustrates nonnormative across-outcome integration with a risk seeking deviation from additivity.)

Figure 3. Mean judgments of happiness as a function of probability in Experiment 1. (Each graph shows the integration of probability and value for one age group. The vertical bars have equal size at the left and right. Diverging linear fan patterns appear for all ages.)

Figure 4. Individual difference classification for the across-outcome integration in Experiment 1. (The figure shows mean judgments of happiness for participants at the three ages employing an additive strategy (left column), two configural strategies (middle columns) and for unclassifiable children (right).)

Figure 5 Mean judgments of happiness as a function of probability in Experiment 2. (The two graphs show the integration of probability and value for younger and older children, with linear fan patterns appearing at both ages.)

Figure 6 Individual difference classification for the across-outcome integration in Experiment 2. (The figure shows mean judgments of happiness for participants employing an additive (left column), risk seeking (middle column) or risk averse strategy (right).)